

Jiří Průcha



JAK TESTOVAT VÝSLEDKY VÝUKY

v malých skupinách studujících



Národní centrum distančního vzdělávání
Praha 2003



O B S A H

ÚVOD / 3

1 Východiska testování / 4

2 Funkce, druhy a základní požadavky na didaktické testy / 6

3 Zásady konstrukce testů / 9

4 Faktory ovlivňující tvorbu testu / 12

5 Hlavní zásady konstrukce testových položek / 14

5.1 Položky s tvořenou odpovědí / 14

5.2 Položky s výběrovými odpověďmi / 19

6 Hodnocení výsledků testu / 22

6.1 Skórování výsledků / 22

6.2 Analýza testu a interpretace výsledků / 23

6.3 Převod skóre na klasifikační stupnici / 26

7 Testy a efektivita vyučování (učení) / 33

7.1 Příprava testování / 33

7.2 Zpracování výsledků šetření analýzou dat / 33

7.3 Účinnost výuky / 36

LITERATURA / 40

PŘÍLOHY

1 Ukázka souboru položek procvičovacího testu / 41

2 Schéma postupu při konstrukci didaktického testu / 47

3 Nejčastější chyby autorů při sestavování a používání didaktických testů / 48

4 Přehled užitých veličin a jejich výpočtů 49

5 Příklady 51

VĚCNÝ REJSTŘÍK / 52



ÚVOD



Testování je jednou z metod, jak získávat co nejspolehlivěji poznatky o určitém objektu. U složitých objektů lze samozřejmě zjistit pouze dílčí údaje, a to ještě jen za určitých předpokladů. Podrobíme-li testování člověka, pak jistě neočekáváme, že se dobereme komplexního poznání. Přesto dobře sestavené testy poskytují zpravidla užitečné informace. Velmi zhruba můžeme rozlišit 3 kategorie: testy osobnosti; testy schopností; testy didaktické.

Testy **osobnostní** zkoumají především rysy charakteru a temperamentu. Testy **schopností** senzorickou, mentální a psychomotorickou stránku osobnosti. Konečně **didaktickými testy** se zjišťují osvojené soubory vědomostí a dovedností. I když je někdy dost obtížné oddělovat výsledky získané testy schopností od výsledků didaktických testů, má toto rozlišení smysl. Výsledky didaktických testů totiž vypovídají o účinku vzdělávacího působení, což je jeden z nejdůležitějších procesů, který jedinec absolvuje (a to v průběhu prakticky celého svého života).

Je víceméně známo, že nejspolehlivější výsledky lze očekávat od **standardizovaných testů**, tj. těch, u nichž byla na vzorku reprezentujícím příslušnou populaci zajištěna platnost, spolehlivost, citlivost a objektivnost na požadované úrovni a zároveň byly stanoveny normy výkonů i přesná administrace testu. Jde o nákladný a dlouhodobý postup, ve své úplnosti přesahující možnosti učitelů či většiny autorů učebnic. Určité standardizované didaktické testy jsou k dispozici a v takovém případě lze samozřejmě jejich používání jen doporučit. My se však budeme věnovat tvorbě **nestandardizovaných testů** vědomostí a dovedností, běžně používaných učiteli ve vlastních výukových skupinách, aniž by jejich výsledky mohly být zobecnovány. Vystačíme při tom se zjednodušenými postupy a odhady, které nicméně mohou v daných podmínkách vést rovněž k velmi dobrým výsledkům.

Ačkoli sestavování didaktických testů tvoří významnou součást odborné kvalifikace učitelů, ve vysokoškolské přípravě se tomuto tématu nevěnuje přiměřená pozornost. To je jeden z důvodů, proč jsou publikace na toto téma účelné. Většina z nich je staršího data (viz např. /1/; /2/; /4/; /5/; /6/; /8/;) a od vydání novější monografie /3/ uplynulo už více než 10 let. Je tedy vhodná doba obnovit zájem o toto téma a pojednat je jak z hlediska nezbytné teorie, tak zejména s ohledem na možnosti využití v pedagogické praxi.

Návodné pojetí této příručky vychází z potřeb a charakteristiky cílové skupiny, kterou jsou především **autoři** učebnic (skript) a učitelé **všech typů** škol. Je určena k samostatnému studiu, což vedlo ke zcela určité grafické i obsahové úpravě. Pro usnadnění orientace v textu je použito těchto symbolů:



vyzývá k **písemnému zpracování** zadání;



naznačuje, že správné řešení je uvedeno v **klíči** na konci příslušného oddílu.

Text je proložen prázdnými rámečky, v jejichž záhlaví je otázka nebo zadání určitého problému. Úkolem v těchto případech je písemně odpovědět nebo zaznamenat řešení. Potlačte své rozladění, pokud to nebude možné splnit na základě prostého nalistování několika předcházejících stran. Má Vás to přimět k využití dosavadních vědomostí či zkušeností a k přemýšlení o problémech. Není vyloučeno, že tak získáte i zajímavé prožitky. Řazení jednotlivých obsahových částí má svou logiku. Lze proto doporučit sledovat uvedenou posloupnost textu. Přesto si nejprve celý text zběžně prolustujte a pokud Vás některé téma zaujme, můžete se mu věnovat přednostně.

Testování se neobejde bez určitých výpočtů založených na poměrně složité teorii. Účelem této příručky je však poskytnout jen ty **nejprostší nástroje** pro zpracování získaných údajů. Značné úsilí bylo věnováno výběru co nejmenšího počtu měřitelných veličin. Jejich definice jsou matematicky jednoduché a konkrétní hodnoty dostupné pomocí elementárních operací, zahrnujících ve dvou případech pouze použití druhé mocniny a odmocniny. Nikdo ze čtenářů proto nemusí mít obavy, že se setká s pojmy, kterým by neporozuměl nebo s výpočty, které svou složitostí odradí. I za těchto okolností je možné dopracovat se spolehlivých výsledků, jež umožní nahlédnout nejen „do hlubin študákovy duše“, ale zároveň do některých zákrutů pedagogického procesu.

Cíle: Poznat druhy didaktických testů i jejich položek. Umět samostatně vypracovat nestandardizované testy vědomostí i dovedností pro diagnostické, klasifikační a didaktické účely, posoudit jejich kvalitu, vyhodnotit dosažené výsledky a naučit se používat testy i pro zjišťování účinnosti výuky a učení.

Problémové okruhy: Psychologické základy testování; funkce, druhy a základní požadavky na testy; zásady konstrukce testů a faktory, které ji ovlivňují; hodnocení a interpretace dosažených výsledků v závislosti na účelu testu.

Klíčová slova: test; testová položka; profil absolventa; uzavřené a otevřené položky; platnost, spolehlivost, objektivnost; citlivost; kritéria efektivity výuky.

Odhad času na studium: 20 hodin (včetně písemného vypracování všech úkolů).

1 Východiska testování



Anglické slovo test znamená zkoušku, přičemž didaktický test zahrnuje takovou (zpravidla písemnou) zkoušku vědomostí a dovedností, která vyžaduje řešení stejných zadání pro všechny zkoušené a u níž je stanovena přesná technika hodnocení úspěchu nebo neúspěchu, popř. technika vyčíslení výsledku.

Poznámka: v dalším textu budeme pod pojmem test rozumět didaktický test.

Zkoumání dosahovaných výkonů při učení má dva hlavní aspekty. Především jde o bezprostřední informace pro studující o aktuálním stavu jejich přípravy. Tato **vnitřní** zpětná vazba má **motivující** účinek pro další práci a zároveň dokumentuje úspěšnost nebo neúspěšnost v učení. Druhým aspektem je doložit komplexnější výsledky učení (především učitelé), což ovlivňuje prospěch resp. získání vysvědčení apod. (**vnější** zpětná vazba). Zcela subjektivní posuzování výsledků má řadu nepříznivých důsledků, neboť je vždy určitou projekcí víceméně složitých vztahů mezi subjekty. Ty v některých případech souvisejí s povahou, druhem a kvalitou posuzovaného výkonu jen velmi volně. Vzpomeňme na obecně přijímané hodnocení některých sportovních výkonů (gymnastika, krasobruslení, skoky do vody), realizované skupinou nezávislých rozhodčích. I když jsou dost přesně vymezena kritéria hodnocení, i když jsou extrémní hodnocení anulována, i když jsou rozhodčí specializováni a zkušení, dochází přesto ke značnému rozptylu v počtech přidělovaných bodů za tentýž výkon. Hodnocení výsledků učení bývá ještě méně precizováno a je komplikovanější (protože i hodnocená situace je komplikovanější).

Na rozdíl od toho, každý test (ať „dobrý“ či „špatný“) v sobě nese rysy **objektivity**, protože měří každému z testovaných stejně. Bez hrubého porušení zásad testování nemůže docházet ke zkreslování výsledků. Přitom v rukou skutečně kvalifikovaného učitele bude vždy převažovat **diagnostická** funkce testu nad funkcí „klasifikační“, vedoucí často až k postihům (trestům).

Studující by měl již na počátku studia (kurzu, programu) znát nejen požadavky na zkoušku, ale i způsob testování. Ten je stejně důležitý jako samotný obsah testování. Může být např. rozdíl v tom, zda studující při testování může, anebo nemůže používat různé pomůcky. Jistě v tomto ohledu existují určité meze, avšak argumentem pro užívání pomůcek může být fakt, že v řadě situací pozdější praxe se mohou kvalifikovaní pracovníci opírat nejen o publikace, ale i další pomůcky, popř. expertízy.

Studující by rovněž měl vědět, jaké druhy pomoci k testování se mu dostane a jaké cvičné testy (které může sám vyhodnotit) mu poslouží k ověření výsledků vlastního studia předtím, než se podrobí výstupnímu testu, hodnoceného vyučujícím. Za těchto okolností přestává být testování hrozbou a stává se konstruktivním prostředkem studia.

Testování studujících je užitečné chápat v rámci integrovaného systému výuky a/nebo samostatného studia, které zahrnuje řadu článků, sloužících osvojení požadovaných činností po stránce teoretické i praktické. Struktura osvojované činnosti vyžaduje různé formy prezentace informací i různé formy praxe potřebné k osvojení učiva. Návaznost jednotlivých článků výuky vyžaduje kontrolu stupně osvojení jak specifických a dílčích, tak i celkových cílů. Sepětí celého programu výuky a kontroly osvojení jeho hierarchicky organizovaných částí může:

- zcela chybět;
- být v různé míře uvolněné;
- z větší části odpovídat cílům výuky.

Tato rozmanitost integrace výuky je lépe prokazatelná u výukových produktů (textů pro samostatné studium nebo pomůcek) než u frontálního vyučování, kdy postup výkladu je určován převážně vyučujícím. Samozřejmý požadavek, aby testování úzce navazovalo na **cíle výuky** a stanovený **profil absolventa**, nemusí být při různých formách výuky plněn na (mlčky) předpokládané úrovni. Kontrola testy, jimiž studující musí během studia projít, by měla též být prubířským kamenem připravenosti absolventů na povolání a základem pro akreditaci studijních programů, kurzů apod.



1/ Ze svých dosavadních zkušeností uveďte přednosti testování



2 Funkce, druhy a základní požadavky na didaktické testy



Při použití testů a testových položek se využívá dále uvedených hlavních funkcí: kontrolní; klasifikační; diagnostické; motivační; poznávací.



2 a/ Napište stručně ke každé z uvedených funkcí testu, co znamená.



Pro posouzení kvality testu jsou rozhodující odpovědi na následující otázky.

- Jak zajistit, aby test **skutečně měřil** to, co měřit má?
- Jak jej sestavit, aby při **opakovaném použití** za stejných podmínek **poskytoval tytéž výsledky** (byl spolehlivý)?
- Jak docílit **přiměřené obtížnosti** položek?
- Jak dosáhnout toho, aby test **citlivě rozlišoval** úroveň jednotlivých výkonů?
- Jak **vyloučit subjektivitu** při posuzování výsledků?

Pokusme se na tyto otázky odpovědět. Základním je požadavek, **platnosti testu** (validity). Čtyři základní kritéria stanoví:

- testovaný jev nesmí být zkreslován jinými jevy, pro testování nepodstatnými (například rozsáhlé testy nemusí být platné už proto, že výkony posléze nepříznivě ovlivňuje únava nebo nedostatek času věnovat se všem položkám, hrozby atd.);
- testovací technika musí být vhodná pro kontrolu osvojení příslušného jevu (fakta a konkrétní údaje měří verbálně, úroveň rozumových dovedností měří prostřednictvím myšlenkových operací, úroveň psychomotorických dovedností měří při manipulacích atd.);
- úlohy v položkách mají pokud možno rovnoměrně pokrývat celý testovaný úsek učiva;
- obtížnost zadávaných úkolů má být přiměřená úrovni znalostí (např. didaktický test z učiva fyziky platný pro gymnazist/k/y není pravděpodobně validní pro žáky/ně/ 8. tříd ZŠ).

Platnost testu lze zajišťovat (zjišťovat) především jasnou **formulací** jeho **cíle**, podrobnou **analýzou učiva**, konstrukcí přiměřeného počtu položek, prověřujících všechny analýzou vytipované partie a **nezávislou expertízou**.

Snadněji a přesněji lze stanovit **spolehlivost** testu (**reliabilitu**). Ta se vyjadřuje pomocí koeficientu **r**, jehož hodnoty leží v uzavřeném intervalu od **0** do **1**. V podstatě měří úroveň náhodných jevů, překrývajících testovaný jev. Za náhodné jevy se považují jak dispozice zkoušeného, tak vnější podmínky a vlivy. Jsou-li splněny základní požadavky, závisí spolehlivost testu na počtu položek **p**; rozptylu **s²** a na mediánu **m** dosažených bodových výsledků. Koeficient spolehlivosti je dán:

$$r = m(p - m) / p \cdot s^2$$

Poznámka: základními požadavky se rozumí symetrické rozložení skutečně dosažených výkonů v testu kolem průměrné hodnoty a strmý pokles výskytu hodnot, které se od průměru odchylují.

Od testů se očekává spolehlivost vyjádřená **r = 0,6** a vyšší. Bylo totiž zjištěno, že teprve testy vyhovující tomuto požadavku poskytují stejné výsledky i při opakovaném testování. Spolehlivost testu lze nejjednodušší zlepšit **zvýšením počtu položek**. O kolik? to zjistíme pomocí faktoru **z**:

$$z = r_p (1 - r_d) / r_d (1 - r_p)$$

kde **r_p** je požadovaný a **r_d** dosažený koeficient spolehlivosti. Vyšla-li hodnota koeficientu **r_d = 0,33** a požadujeme-li **r_p = 0,6**, vychází multiplikativní faktor **z**:

$$z = 0,6 (1 - 0,33) / 0,33 (1 - 0,66) = 3$$

Tzn. počet testových položek by bylo třeba zvýšit třikrát(!). Rozsah testu by se tím pravděpodobně dostal do rozporu s požadavkem platnosti. Testy s nízkou spolehlivostí je třeba zcela **přepracovat**.

Teprve blíží-li se index spolehlivosti hodnotě **0,6** má smysl se zabývat jejím event. zlepšením. V zásadě jsou dvě další možnosti:

- Rozdělit test na sadu dílčích testů, v nichž je potom zvýšení počtu položek únosné.
- Nahradit příliš snadné a příliš obtížné položky, přičemž v obou případech je měřítkem obtížnosti dosažený počet správných odpovědí. Položka je považována za **snadnou**, splní-li ji **85 %** a více testovaných a za **obtížnou**, splní-li ji **25 %** a méně testovaných.

Poznámka: Opravu příliš obtížných resp. příliš snadných položek není vždy účelné uskutečnit. Jde o případy testů **základních vědomostí**, nebo testů **naučení**, u nichž se předpokládá nesymetrické rozdělení dosažených skóre. (U testů základních vědomostí musí převážná část testovaných splnit převážnou část položek.) Proto postup při zvyšování spolehlivosti závisí také na **účelu** testu.



2 b/ Test má **35** položek, medián dosažených bodů je **21** při rozptylu **16** (směrodatná odchylka **s = 4**). Vyjádřete koeficient spolehlivosti **r** tohoto testu.



2 c/ Na jaký počet položek by bylo třeba zvětšit rozsah testu, kdybychom chtěli dosáhnout spolehlivosti **r = 0,65** ?

Obtížnost testu se vyjadřuje indexem obtížnosti **I_o = 100 $\bar{X}$ / X_{max}**, kde $\bar{X}$ je aritmetický průměr dosažených bodových skóre a **X_{max}**, je nejvýše dosažitelný skóre v testu.

I_o udává průměrné procento správně vyřešených testových položek a jeho optimální hodnota je **50 %**. Při podstatně nižších nebo vyšších hodnotách jde o test příliš obtížný, resp. příliš snadný.

Poznámka: hodnoty indexu obtížnosti podstatně vyšší než **50 %** jsou obvyklé v testech **základních vědomostí**; hodnoty podstatně nižší jsou oprávněné jen v testech sloužících k výběru nejlepších studujících.



2 d/ Jaký je index obtížnosti testu uvedeného pod zadáním 2 b/, jestliže průměr testového skóre dosáhl hodnoty $\bar{X} = 19,8$ bodů a v testu je za každou správnou odpověď přidělován **1** b? Jak lze zjištěnou hodnotu komentovat?



Citlivost určuje, do jaké míry lze rozlišovat mezi jednotlivými výkony. O testu řekneme, že je citlivý, jestliže:

- výkony testovaných osob dosahují co největšího počtu možných bodových hodnot (od nejnižšího do nejvyššího),
- jsou symetricky rozloženy kolem aritmetického průměru, přičemž
- výskytu extrémů (nejnižších i nejvyšších bodových hodnot) strmě ubývá.

Citlivost testu se pak vyjadřuje indexem:

$$i_t = s_x / \bar{X},$$

kde $\bar{X}$ je aritmetický průměr dosaženého skóru v testu a s_x je směrodatná odchylka skóre. Index citlivosti má být roven nebo větší než **0,25**.



2 e/ Určete ještě citlivost testu, kterým jsme se v předchozích výpočtech zabývali. Vyjádřete se ke zjištěné hodnotě i_t .



Objektivnost testu se týká především hodnocení dosažených výsledků, jež má být propracováno tak, aby každý hodnotitel přiřadil stejnému výkonu tentýž bodový skóre. Objektivnost testu závisí na jeho obsahu a technice hodnocení. Každý test by měl být sestaven tak, aby:

- existovala proporcionalita mezi závažností a rozsahem učiva a **počtem i obtížností** položek;
- byla stanovena přesná kritéria pro hodnocení správnosti, popř. neúplnosti odpovědi (výkonů).

Poznámka: Z tohoto hlediska je nejjednodušší situace při zadávání **výběrových** odpovědí. Je-li jednoznačně a srozumitelně formulováno zadání (otázka), je také možné jednoznačně určit, která z nabízených odpovědí je (zcela) správná, chybná popř. neúplná. Objektivnost posuzování výkonu je tak maximalizována a je prakticky nezávislá na osobě hodnotitele.

Objektivnost testu také úzce souvisí s jeho **platností**. Má-li být např. vypracován test zjišťující vědomosti a dovednosti z učiva matematiky, mohou být značné rozdíly už v jeho **koncepci**. Jeden autor může přikládat větší význam dovednosti **sestavovat** rovnice, jiný zdůrazní význam jejich numerickému **řešení**. Skladba testu bude v obou případech jiná, neboť řešení sestavených rovnic je časově méně náročné, takže test může obsahovat více položek. I v hodnocení výsledků mohou být značné rozdíly, protože první z autorů bude patrně tolerantnější k numerickým chybám, zatímco druhý bude chyby tohoto typu posuzovat přísněji, což se projeví v kritériích hodnocení. Jak ještě uvidíme dále, je nejpříjemnější ocenit **každou operaci** určitým počtem bodů a tím sjednotit hlediska pro posuzování dosažených výkonů. Nelze tím nicméně docílit objektivizaci **obsahu** testu, protože koncepce je zde konečnou dána pojetím učiva (např. v rámci určitého modulu nebo předmětu).

3 Zásady konstrukce testů



Východiskem tvorby testu je podrobná **analýza učiva**, které bylo předloženo k osvojení (a má být testováno) i kritické posouzení **obsahových proporcí** jeho jednotlivých částí (souborů). Co nejpřesnější vymezení **cíle(ů) zkoušky** a zaměření celé práce na testu tímto směrem, je nejen východním, ale patrně také nejdůležitějším úkolem, který před autorem testu leží. Jednou z prvních otázek, na kterou si musí autor testu odpovědět, je **účel** testu. Půjde o test klasifikační, diagnostický, procvičovací, test základního učiva nebo dokonce o určení nejlepších (nejhorších) studujících?

Důležitou je také odpověď na otázku po **počtu položek**. Stanovuje se u každého vytčeného cíle, a to s ohledem na množství času, které je pro osvojení příslušného učiva plánováno (podle jeho rozsahu, významu, obtížnosti apod.). Vždy je účelné počítat asi s **dvojnásobným počtem** položek, protože při zkoušení testu a analýze výsledků dojde obecně k jejich snížení (pro malou citlivost, přílišnou obtížnost, nevhodnou formulaci zadání apod.). Není také vyloučeno, že bude navržen test s nízkou spolehlivostí a dodatečné zařazení „nadbytečných“ položek ji může - jak už víme - zlepšit.

Limitujícím faktorem je však i **čas**. Testování by nemělo přesáhnout dvě spojené vyučovací jednotky (90 až 100 minut), přičemž určitou dobu vždy zaberou pokyny, popř. zácvik a distribuce formulářů. Někdy se mluví i o vizuálním zadávání položek (na tabuli, projekcí apod.). To však nelze doporučit. U testů cizojazyčné způsobilosti jsou nicméně běžné hlasité přehrávky zvukových záznamů, určené všem testovaným osobám společně. Patrně nepříjemnější je individuální elektronická prezentace testových položek z disket, CD resp. DVD nosičů nebo zprostředkovaná počítačovou sítí a spojená s automatickým vyhodnocováním jednotlivých odpovědí i celkového výsledku.

Při „tradičním“ hromadném nasazení tištěných formulářů je vhodné použít alespoň dvou ekvivalentních testových forem, které zpracovávají studující, sedící vedle sebe.



3 a/ Jak se dosahuje toho, aby více testových forem bylo rovnocenných?



Jedním z důležitých cílů při osvojování jakéhokoli učiva je nejen zapamatovat si početnou množinu izolovaných fakt, nýbrž pochopit řadu souvislostí mezi nimi, osvojit si příslušné aplikace základních myšlenkových operací a také zajistit maximální **transfer** vědomostí i dovedností. Proto by každý didaktický test měl zahrnovat také úlohy, při jejichž řešení má studující prokázat dovednosti:

- mnohostranné a správné **manipulace** s nově osvojenými poznatky;
- zařazování nových poznatků do systému dříve osvojených.

Používají se různé typy položek, tříděné podle **druhu odpovědi**:

- **tvořené** odpovědi jednoslovné či grafické;
- odpovědi doplňovací (s vynechaným klíčovým slovem, číslem nebo symbolem);
- rozsáhlé tvořené odpovědi;
- odpovědi **výběrové** z předložených variant (formulací);
- přiřazovací odpovědi;
- rozdělovací odpovědi;
- seřazovací odpovědi.



Seznamte se s příklady uvedenými v odd. 5 a pak *uved'te vhodný příklad na každý z druhů testových otázek. Jde o základní dovednosti při tvorbě testu, a proto si dejte záležet na výběru, ale zejména na **jednoznačných** formulacích.*

Relativně nejmenší obtíže jsou spojeny s posuzováním správnosti **tvořených odpovědí doplňovacích**. I zde samozřejmě závisí na vhodné stylizaci věty (tvrzení, definice), v níž má být doplněno klíčové slovo (číslo, symbol). Obecněji řečeno - na jednoznačném „navedení“ studujících k doplňku chybějící části (ale bez nápovědi!). Doplněk se má utvořit na základě získaných vědomostí, aniž je řešení v explicitní formě „skryto“ v předchozím textu.

Tvořené (volně formulované) odpovědi mají některé přednosti. Mezi ty největší patří právě ta okolnost, že jsou tvořené, takže někdy dávají nahlédnout do myšlenkových pochodů testované osoby. To usnadňuje poznat kvalitu případné chyby a spolehlivěji rozhodnout o možnosti, jak neznalost nebo nepřesnost napravit. Jako výsledek **znovuvybavení** (na rozdíl od snazšího znovupoznání, které se uplatňuje u výběrových odpovědí), jsou z didaktického hlediska považovány za cennější.

Druhou velkou skupinu testových položek představují zadání s připojeným souborem **výběrových** odpovědí. Jejich výhodou je objektivita při hodnocení výkonu, protože zvolenou odpověď lze jednoznačně posoudit z hlediska správnosti a přidělit jí bez problémů určitou váhu (bodové ocenění). Zavedení výběrových odpovědí rozvířilo diskuse o jejich přiměřenosti a vhodnosti. Ačkoli názorové výměny o jejich významu a vhodnosti ztratily časem mnoho ze své vyhraněnosti, dodnes zcela neutichly.



3 b/ Uveďte obvyklé výhrady pedagogů k zařazování položek s výběrovými odpověďmi do testů.



- **Správnou odpověď** lze (na rozdíl od tvořené) **uhádnout**. To je samozřejmě pravda. Ze statistického hlediska je pravděpodobnost náhodné volby správné odpovědi nepřímo úměrná počtu alternativ. Jsou-li nabízeny 4 možnosti, z nichž pouze jedna je správná, je pravděpodobnost náhodné volby správné odpovědi $\frac{1}{4}$.
- **Celkový počet** správných odpovědí náhodně zvolených závisí nejen na počtu alternativ, ale i na počtu položek. Proto platí, že výsledky dosažené v testu s nabízenými 2 alternativami za každou otázkou mohou být při **větším počtu položek spolehlivější**, než výsledky dosažené v testu nabízejícím sice za každou otázkou více alternativ, který je však krátký.

Je tedy třeba opravit výsledek testu s ohledem na možnost náhodné volby správných odpovědí? V každém případě je to možné. Následující vztah vyjadřuje opravený počet správných odpovědí A_k , stanovený s ohledem na pravděpodobnost volby správné odpovědi (je-li celkový počet správných odpovědí zvolený testovanou osobou A_s , A_n celkový počet zvolených nesprávných odpovědí a y počet alternativ předložených k volbě správné odpovědi).

$$A_k = A_s - \frac{A_n}{(y - 1)}$$



3 c/ Ze všech testových položek zodpověděla testovaná osoba 11 a z toho 8 správně. Volilo se vždy ze 4 výběrových odpovědí. Kolik správných odpovědí započtete jako výsledek testu po opravě na možnost náhodné volby správné odpovědi?



Poznámka: Ze vztahu pro opravený počet správných odpovědí je patrné, že se vzrůstajícím počtem alternativ (y) klesá hodnota opravného členu, takže oprava na „hádní“ se ve shodě s teoretickým předpokladem snižuje.

Jak se tento vztah uplatňuje v praxi lze ukázat na dalším příkladu. Určitý test má 30 položek a v každé z nich jsou předkládány na zadanou otázku 3 výběrové odpovědi. Zkoušený(á), který(á) zodpověděl(a) všech 30 otázek a z toho pouze 10 správně, **mohl(a)** všechny správné odpovědi uhádnout, neboť podle počtu pravděpodobnosti by v takovém případě náhodně zvolil(a) každou třetí odpověď správně. Opravený počet správných odpovědí vychází roven 0, jak vyplývá z výpočtu:

$$A_k = (10 - 20)/(3-1) \div 10 - 10 = 0$$

(Prostý skóre testu X je tedy 0.)

Poznámka: „oprava na hádní“ má ovšem tu nevýhodu, že nerozliší testovanou osobu, která (v daném případě) 1/3 správných odpovědí vybrala náhodně od osob, které znaly správné odpovědi pouze na 1/3 testových položek. V prvním případě se opravou skóre eliminuje náhodně dosažený výsledek, ale ve druhém se bohužel testování neoprávněně poškozuje (aniž můžeme objektivně zjistit, o který z případů jde).



3 d/ Jaký počet možností při předkládání výběrového souboru odpovědí v testu považujete za optimální a proč?



Pokud je testová položka doplněna výběrovým souborem odpovědí, může už jejich formulace **napovědět nejvhodnější volbu** nebo snížit počet možností, které přicházejí v úvahu jako správné řešení (aniž testovaná osoba správnou odpověď skutečně zná). Tato výhrada se nejčastěji uvádí u číselně vyjadřovaných nabídek, které jsou často výsledkem matematických operací. Při numerických výpočtech může být totiž nesprávná odpověď signalizována neshodou vypočítaného výsledku s předloženými nabídkami. Zkoušená osoba může pak hledat různé postupy tak dlouho, až dosáhne shody s jednou z alternativ. Tomuto nebezpečí lze čelit tak, že autor testu zařadí jako alternativy taková čísla, která vznikají při **typicky chybných** postupech výpočtu, takže zkoušená osoba může dosáhnout shody i s některými chybnými nabídkami a na výsledky víceméně náhodných manipulací se nemůže spolehnout. Další „obranou“ je zařazení univerzální odpovědi typu „žádná z uvedených odpovědí není správná“ a jako ostatní alternativy se pak uvedou jen chybné odpovědi.

4 Faktory ovlivňující tvorbu testu



• Faktor času

Čas pro vypracování testu může být podle okolností několik minut až několik hodin. Už zmíněné rozdělení testu na subtesty umožňuje při celkově vysoké spolehlivosti prověřit i rozsáhlé úseky učiva, ovšem v oddělených časových intervalech. U didaktických testů je třeba stanovit dobu pro vypracování, sdělit ji před zahájením práce a časový limit dodržet.



4 a/ Jak odhadujete potřebný čas pro vypracování testu studujícími, před jeho prvním nasazením a jak je možné zabránit nekorektním postupům při řešení testových zadání studujícími?



- **Průkaznost osvojení učiva**

Autor testu musí dobře uvážit, **kolik a jakých** testových položek postačí k prověření toho, zda vytčeného didaktického cíle bylo skutečně dosaženo.

Např. má-li test prokázat osvojení **základů** názvosloví binárních anorganických sloučenin, postačí plně 20 položek, z nichž deset požaduje **napsat** nebo **zvolit názvy** oxidů, sulfidů a halogenidů podle uvedeného **vzorce** a deset zbývajících položek pak **napsat** nebo **zvolit vzorce** analogických (v prvních položkách nezastoupených) binárních anorganických sloučenin, je-li uveden jejich **název**. Požadavek platnosti testu ještě stanoví, aby pro kationtovou část sloučenin byly vybrány běžné prvky, ale v různém oxidačním stupni (různého mocenství).

Podobně je tomu u testu, který zkoumá znalost pravopisně správného používání vybraných slov např. po obojetné souhlásce „p“. V zásadě se lze spokojit s formulací 20 položek, kde se všechna tato vybraná slova v různých větách opakovaně vyskytují (např. s vynecháním samohlásek „i“ a „y“ v kritických místech).

Konečně třetím příkladem může být test pro předmět hudební výchova, zjišťující základní vědomosti o tóninách. Jednotlivé položky zkoumají znalost názvů durových a mollových tónin podle předznamenání a naopak - dovednost určit předznamenání podle názvu tóniny. Přitom lze samozřejmě využít jak tvořených tak výběrových odpovědí (volitelných, seřazovacích i přiřazovacích). Počet položek závisí na tom, kolik testovaných jeví se v rámci položky zároveň zkoumá, ale ani zde nemusí překročit číslo 20.

Chce-li autor testu u důležitých partií učiva získat jistotu, že správná odpověď nebyla zvolena náhodou, má v zásadě dvě možnosti. Buď příslušnou položku přeformuluje a zařadí opakovaně, anebo pod výběrový soubor uvede ještě otázku typu: „Jste si svou odpovědí jist/a/?“ Za kladnou odpověď přidá v případě správné volby bod navíc. Vybere-li studující nesprávnou odpověď a potvrdí svoji jistotu, je penalizován/a/ mínus 1 bodem (místo nulového hodnocení za pouze chybnou volbu).

- **Uspořádání testových položek**

Pořadí průběžných (dílčích) testů má v zásadě odpovídat metodické posloupnosti, již bylo prezentováno učivo. Všeobecně se ještě doporučuje, aby v každém dílčím testu byly položky uváděny s proměnnou obtížností (i když toto hledisko je poněkud problematické). Existují zadání (otázky), o nichž se dá celkem objektivně říci, že jsou snadnější či obtížnější než jiné, ale obtížnost je často **subjektivní** záležitostí. V každém případě je vhodné, aby alespoň první položky byly objektivně snadnější a přispěly tak k úspěšnému „otevření“ testu. Další důležitou podmínkou je, aby testové položky byly pokud možno **vzájemně nezávislé** v tom smyslu, že správnost odpovědi na kteroukoli z nich **není podmíněna** správností odpovědi v jiné položce.

- **Instruktaž testovaných osob**

Vzhledem k poměrně značně rozdílným možnostem v zadávání testů, různým hlediskům uplatňovaným při posuzování správných a chybných výkonů, časovým limitům vymezeným pro práci a dalším víceméně závažným okolnostem, je zpravidla nutné stručně, ale zcela přesně informovat zkoušené osoby o způsobu vypracování testu. Hlavní snahou je zde zachovat standardní podmínky i pro případy opakovaného testování, zadávání testu jiné skupině, jiným učitelem apod.



4 b/ Co vše považujete za nezbytné sdělit studujícím **před** zahájením testování?



5 Hlavní zásady konstrukce testových položek



Sestavení souboru testových položek je vždy problémem, jemuž je věnována ještě podstatná část dalšího textu. V následujících odstavcích jsou shrnuty a na příkladech předvedeny **hlavní zásady**.

Nejdůležitější požadavky na sdělovanou informaci jako na kvantum nového učiva, resp. jako na určitý (po)krok v učení jsou:

- informace musí být srozumitelná a názorná;
- ke kontrole jejího pochopení mají postačit 1 - 2 testové položky.

V zásadě se rozlišují dva druhy položek - **otevřené** a **uzavřené**. Otevřené vedou k tvořeným odpovědím, v uzavřených položkách se nabízejí přímo různé formulace odpovědí. Jak už bylo řečeno, uzavřené položky mohou být výběrové, přiřazovací, seřazovací a rozdělovací.

5.1 Položky s tvořenou odpovědí



5.1 a/ Jak rozumíte pojmu „tvořená odpověď“?



Uveďme dva jednoduché příklady, vycházející z malých částí učiva sdělovaných v učebnici:

A: Přirozená jsou všechna čísla, která jsou zároveň celá i kladná. Např. 2 je přirozené číslo, protože je zároveň celé i kladné.

B: Podle toho, zda údaje byly zjištěny přímo, či byly vypočteny teprve vzájemným porovnáváním zjištěných údajů, rozlišují se ukazatelé ABSOLUTNÍ a RELATIVNÍ (poměrní).

Např.: a) město má letos 10 852 obyvatel;

b) velikost produkce podniku vzrostla za 10 let 2,5 krát.

Je-li nyní třeba **prověřit**, zda byla pochopena **podstata** každé ze zcela elementárních informací, pak existuje prakticky nepřehledné množství možností, jak to udělat. Formulace testové položky (zadání úkolu) vždy závisí na tom, v jaké souvislosti (proč; za jakým účelem) bylo učivo prezentováno. Tj., jak zdůraznit jeho podstatu v návaznosti na předešlou „historii“ učení i v závislosti na následně předkládaných sekvencích učiva v rámci učebních cílů. Uveďme si teď ukázky typických možností při prověřování výsledků učení.

***Poznámka:** Pojem otázky je zde i dále v textu myšleno jakékoli zadání v testové položce, vyzývající k samostatné činnosti.*

A1: Rovněž číslo 786 je číslem.

A2: Rovněž číslo 786 je přirozeným číslem, protože je zároveň i kladné.

A3: Rovněž číslo 786 je přirozeným číslem, protože je zároveň celé i

A4: Rovněž číslo 786 je přirozeným číslem, protože je zároveň i

A5: Rovněž číslo 786 je přirozeným číslem, protože je celé i kladné.

A6: Je také 786 číslem přirozeným?

A7: Číslo 786 je celé a kladné. Je tedy přirozeným číslem?

A8: Číslo 78,6 je kladným číslem, ale není číslem celým. Je 78,6 číslem přirozeným?

A9: Číslo - 786 je celým číslem, ale není kladné. Je - 786 číslem přirozeným?

A10: Doplněte v řadě 1; 2; 3;; 5; 6; 7; 8; 9; 10; ...; 12; 13; 14; ...; 16 chybějící přirozená čísla.

A11: Vypište z dále uvedených čísel všechna přirozená čísla

15; 173,8; - 21,3; 3; 1 601 385; 2/3; 3 . 2,5

A12: Patří lichá čísla k číslům přirozeným?

A13: Jsou sudá čísla čísla přirozenými?

A14: Patří nula k přirozeným číslům?

Poznámka: chceme-li zdůraznit **definici** přirozených čísel, použijeme některé z možností A1 až A5; chceme-li posílit dovednost rozlišovat (vyhledávat; určovat) přirozená čísla, můžeme použít některé z formulací A6 až A10; otázka A11 je náročnější, neboť rozšiřuje použití definice i na výsledek některých aritmetických operací; podobně i otázky A12 až A14 vyžadují pro správnou odpověď další úvahy.

A15: Co jsou přirozená čísla?

Poznámka: Na tomto místě otázka není vhodná, protože vede k pouhému opsání části učebního textu.

A16: Dokážete rozhodnout, které z následujících čísel jsou přirozená ? (následuje řada čísel).

Poznámka: Otázka je nevhodně formulována, neboť se de facto ptá na něco úplně jiného, než je pochopení učiva. Formálně správná odpověď by totiž mohla znít **ANO**.

A17: Vyberte z následující skupiny čísel všechna přirozená čísla (následuje řada čísel).

Poznámka: v otázce není jednoznačně řečeno, jakým způsobem se má výběr uskutečnit (podtržením, škrtnutím, vypsáním apod.).

A18: Vypište z následující řady přirozená čísla: - 11; 35; 15; - 126

Poznámka: mezi čísla by měla být nejen čísla celá, ale také desetinná, popř. zlomky a další racionální čísla. Jen tak se prověří správnost pochopení definice přirozených čísel v širším rozsahu.

A19: V následujícím seskupení označte čísla:

a) celá; b) kladná; c) záporná; d) celá záporná; e) kladná desetinná; f) přirozená.
(následuje seskupení různých čísel).

Poznámka: Otázka je příliš rozsáhlá, odpověď vyžaduje mnoho operací.

A20: Uveďte další příklady přirozených čísel.

Poznámka: Aby odpověď byla kontrolovatelná, je nutné při automatickém vyhodnocení testu vložit speciální program pro určování přirozených čísel.

A21: Můžete uvést další příklady přirozených čísel?

Poznámka: Testovaný/á/ může odpovědět také jen **ANO** nebo **NE**.

A22: Je každé číslo číslem přirozeným?

Poznámka: Otázka je nevhodná, protože ani správná odpověď **NE** neprověřuje úplné pochopení učební dávky.

A23: Jaká jsou přirozená čísla?

Poznámka: Nevhodná formulace - existuje mnoho neúplných a nepřesných odpovědí, k nimž otázka přímo vybízí.

A24: Kolik přirozených čísel je v následující řadě? (následuje řada čísel, z nichž je 5 přirozených).

Poznámka: Odpověď 5 může být výsledkem chybné úvahy; studující může zvolit 5 čísel a přesto se při jejich výběru dopustit (nezjistitelné) chyby.

K ukázce B.:

B1: Údaj 10852 je absolutní ukazatel. Vzrůst produkce podniku na 2,5 násobek je ukazatel.

B2: Údaj 10 852 je ukazatel. Vzrůst produkce podniku na 2,5 násobek je ukazatel.

B3: Hovoříme-li o **relativních** ukazatelích, znamená to, že jde o ukazatele, vzniklé zjištěných hodnot. Jejich opakem jsou ukazatelé, zjišťované přímým měřením (počítáním).

B4: Údaj 10 852 obyvatel je a..... ukazatelem, kdežto vzrůst produkce podniku na 2,5 násobek je p..... ukazatelem.

Poznámka: v ukázce je použito nápovědi, využívající počátečních písmen příslušných termínů.

B5: Hovoříme-li o ukazatelích vzniklých porovnáváním údajů, jde o ukazatele

B6: Počet obyvatel byl zjišťován absolutně, tj. bez porovnáváním s nějakým jiným údajem.
Jde tedy o ukazatel

B7: ukazatelé vznikají porovnáváním (tj. poměrem dvou údajů).

B8: Jaký ukazatel vznikne porovnáváním (poměrem) dvou údajů?

Poznámka: V ukázkách B6 až B8 je nápověď skryta v stejnojmenném názvu postupu při zjišťování ukazatelů.

Další ukázky již nejsou komentované a měly by sloužit jako inspirační zdroj pro tvorbu prověřovacích otázek.

C: Ukázka zadání definice:

C1: Zkuste charakterizovat pojem **tvůrčivost**.....

C2: Moderní společnost je plná nejrůznějších sdružení, spolků, svazů i dalších komunit, které se utvářely a dále vznikají v rámci nejrůznějších sociálních kontaktů. Uveďte co nejvíce dalších seskupení:

.....

(Řešení: klany, družstva, náboženské společnosti, politické strany, sportovní kluby, hnutí, nadace, gangy, aliance, trusty).

C3: Charakterizujte pojem **pracovní tým**, řešící nějaký technický, organizační či pedagogický problém. Jak by se dal výstižně a stručně definovat pojem **tým**? (uvažte podstatné rysy týmu, jakými jsou cílová orientace, malý počet členů, organizovanost, zodpovědnost za řešení cíle a dočasnost).

.....

(Řešení: cílově zaměřená, vnitřně organizovaná skupina s relativně malým počtem členů odpovědných za řešení určitého problému a existující v časově vymezeném období.)

D: Ukázky položek na doplňování chybějících klíčových slov:

D1: Charakter i úroveň komunikace vždy závisí také na sociálním v němž probíhá;

(Řešení: kontextu, pozadí, prostředí).

D2: Každý z nás už v nějakém týmu pracoval a ví leccos o potížích s tím spojených. Pokuste se proto sepsat **hlavní okruhy** problémů, které týmová práce představuje:

.....;;;;;

D3: Sepište typické výhody (+) a nevýhody (-) týmové práce:

přednosti (+)

nedostatky (-)

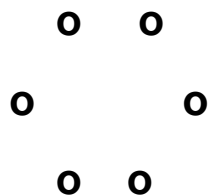
.....;;;;;

D4: Vytýčením pravidel chování a kontroly jejich dodržování přechází tým do (které?) fáze své existence.

D5: V následujících charakteristikách zatrhněte vždy tu alternativu (*uvedenou v závorce*), kterou považujete za správnou:

- vedoucí týmu je uznáván především podle (*formální - neformální*) autority, kterou získal.
- názorové rozdíly jsou chápány především jako (*zdroj bohatství podnětů - rušivý faktor*).
- řešení problémů mají být (*optimální - ideální*), vybraná z alternativ podle (*co nejjasněji - alespoň přibližně*) určených hodnotících kritérií.

D6: V následujícím obrázku doplňte v představě **všechny** komunikační vazby mezi členy týmu (kteří jsou schematizováni kroužky) a pak odvoďte a napište obecný vzorec pro jejich výpočet:



Vzorec pro zjištění počtu komunikačních vazeb **k** mezi **n** členy týmu:

k =

E: Ukázky opakovacích (shrnujících) zkušebních otázek:

E1: Na závěr této části uveďte přehled typů konfliktových situací a u každého typu uveďte **podstatu** konstruktivního řešení:

střet:

řeší se (jak) ?

představ;

názorů;

postojů;

zájmů

E2: Určitá energetická hladina v atomu má **n = 3** a **l = 0**. Do které elektronové vrstvy patří a kterým písmenem se tato vrstva označuje?

E3: Jsou v následující posloupnosti seřazeny elektronové vrstvy podle klesající nebo podle stoupající energie elektronů? (napište): **O; N; M; L; K**

E4: Do kterého orbitalu patří elektron s charakteristikou: **n = 3; l = 2; m = -2; s = +1/2** ?

E5: Vyhledejte v Mendělejevově periodické tabulce všechny prvky patřící:

a) do II. hlavní podskupiny; ;

b) do 2. periody;

E6: Jaký převládající typ chemické vazby se pravděpodobně vytvoří mezi následujícími dvojicemi atomů, jestliže dojde k jejich interakci (použijte Mendělejevovu periodickou tabulku):

Br a Br; Ca a O; H a S; Na a O; Mg a Cl; H a F

.....

E7: Napište srážecí reakci probíhající při smíchání roztoků uhličitanu sodného a chloridu vápenatého:

..... →

E8. Zachyťte chemickou reakcí tepelný rozklad uhličitanu vápenatého: →

E9: Představte si, že na jedné misce tárovací váhy leží prázdná duše do nafukovacího míče a na druhé misce vyvažující závaží. Co se stane po nafouknutí duše a jejím vrácení na miskou vah? (označte správnou alternativu):

- a) rovnováha na miskách váhy zůstane neporušena;
- b) poklesne miska se závažím;
- c) poklesne miska s nafouknutou duší.

Svou volbu zdůvodněte :

Poznámka: jde o kombinaci výběrové a tvořené odpovědi.

E10: Z následujících údajů sestavte graf; zvolte vhodně závislou a nezávislou proměnnou a navrhnete výstižný název grafu; (následuje zadání údajů.)

E11: V následujících dvou sloupcích spojte úsečkou vzorec s příslušným názvem sloučeniny (některé vzorce zůstanou bez spojnice s názvem).

NH₃
K₂O
Mn₂O₇
H₂O₂
NaH
HCl
ClO₂
H₂O
As₂S₃
SnS₂
AgBr

voda
čpavek
oxid draselný
oxid chloričitý
chlorovodík
oxid manganistý
sulfid ciničitý
hydrid sodný
peroxid vodíku

Řešení:

NH₃
K₂O
Mn₂O₇
H₂O₂
NaH
HCl
ClO₂
H₂O
As₂S₃
SnS₂
AgBr

voda
čpavek
oxid draselný
oxid chloričitý
chlorovodík
oxid manganistý
sulfid ciničitý
hydrid sodný
peroxid vodíku

E12: K uvedeným obrázkům přiřadte správné názvy:

graf; diagram; nomogram; schéma; (následují příslušná vyobrazení.)

E13: V připojené tabulce přiřaďte k jejím částem dále uvedené názvy: hlavička; legenda; sloupce; řádky; nadpis (následuje tabulka).

Poznámka: E11 až E13 jsou ukázkami tzv. přiřazovacích odpovědí; další příklad uvádí seřazovací odpovědi.

E14: Seřadte následující operace technologického postupu výroby litiny do správné posloupnosti tím, že před každou operací napíšete její pořadové číslo (následuje výčet operací):

.....

E15 Zařadte následující výrazy pod 10 slovních druhů; (následuje netříděná množina podstatných jmen, přídavných jmen, sloves, číslovek, předložek, příslovcí atd.)

E16 Vyberte z dále uvedených vlastností síry vlastnosti fyzikální a chemické a sepište je:

fyzikální vlastnosti

chemické vlastnosti

(Uvedena je neroztříděná množina vlastností prvku S.)

Poznámka: E15 a E16 jsou příklady rozdělovacích odpovědí.

F: Ukázky tvořivých problémových zadání

F1: Vyjmenujte co nejvíce objektů, které jsou zároveň:

a) hranaté a měkké;

b) umělé a pohyblivé.

.....

.....

F2: Uvedte co nejvíce všech myslitelných použití pro:

a) dřevěné 10 cm pravítko;

b) nahuštěný kopací míč.

.....

.....

F3: Uvedte co nejvíce problémů, které jsou spojeny s použitím:

a) ruční pily;

b) stearinové svíčky.....

F4: Sepište co nejvíce společných vlastností pro každou z následujících dvojic objektů:

a) povidla - noty;

b) déšť- piliny;

c) krtek - lokomotiva.

.....

.....

.....

F5: Jak byste zlepšil(a) dále uvedené předměty?

a) kapesní nůž;

b) jízdenka MHD;

c) propiska.

F6: Jak byste pomocí jediné zápalky sestrojil(a) čtverec (bez jejího lámání či ohýbání)? Popište:

F7: Vymyslete a popište co nejvíce způsobů, jak:

a) oddělit od klubka kousek provázku bez použití nože či nůžek

b) udělat do dubového prkénka otvor (průměr 0,5 cm) bez použití nebozezu nebo vrtačky

 **5.1 b/** O jaký typ odpovědi jde v položkách F1 až F7 ?



Převážná většina tvořených odpovědí vyžaduje doplnění jednoho nebo několika málo znaků. Se vzrůstajícím počtem požadovaných znaků roste možnost jejich kombinací (zvláště pokud se spojují do složitějších soustav (např. slov). Odpověď pak může být kontrolována automaticky jen s obtížemi. Čím přesněji souhlasí utvořená odpověď s odpovědí vloženou do kontrolního mechanismu pro vyhodnocování testu, tím jednoznačnější je rozhodnutí o její správnosti.

Vyžaduje-li otázka odpověď o několika slovech nebo dokonce odpověď ve formě krátké věty, pak je v rámci kontrolního mechanismu nutné uvést co nejvíce obměn odpovědi při zachování jejího smyslu. Např. na otázku: „Co je protonové číslo prvku?“ lze odpovědět těmito více či méně přesnými formulacemi, které můžeme považovat za správné:

- počet protonů;
- součet protonů;
- počet protonů v jádře;
- součet protonů v jádře;
- počet protonů v jádře (jeho) atomu(ů);
- součet protonů v jádře (jeho) atomu(ů);
- počet protonů v jádře atomů prvku;
- součet protonů v jádře atomů prvku;
- počet protonů v jádře atomů tohoto prvku;
- součet protonů v jádře atomů tohoto prvku atd.

Některé odpovědi vyžadují respektovat **pořadí slov** (termínů, operací). Pokud nejsou kontrolovány automaticky, je nutné uvést i skutečnost, že na tomto pořadí záleží.

5.2 Položky s výběrovými odpověďmi

Příklad:

A: Pro krystalickou látku je charakteristické:

a) pravidelné uspořádání atomů;

b) vytváření pravidelných geometrických tvarů;

c) vytváření jediného krystalu.

(Testovaná osoba má označit jednu z předložených nabídek odpovědí.)

Poznámky: - Pokud jde o charakter nabídek, žádná z nich není vysloveně chybná. Možnosti b) a c) jsou pouze **nepřesné** (resp. odpověď b/ je nepřesná a odpověď c/ není obecně správná).

- Existuje-li v nabídce více správných odpovědí, musí na to autor testu upozornit.

Nabízí-li autor testu odpovědi, místo toho, aby vyžadoval jejich tvorbu, omezuje tím často varietu možností. Podstatou problému v takovém případě sice nepřestává být **formulace otázky**, ale těžiště se přesouvá do oblasti sestavení vhodného souboru **nabízených odpovědí**. Jaká jsou v tom případě kritéria „vhodnosti“ pro soubor výběrových odpovědí a jaké jsou možnosti jejich zadávání?

Především je nutné vyloučit z nabízených odpovědí ty, které **nemají vůbec žádný vztah** k prezentovanému učivu (jako jsou např. odpovědi v kvízech) a tvoří jen jakousi „vatu“ k doplnění počtu. Předkládané možnosti mají tvořit **smysluplné kognitivní pole**, v němž se testovaná osoba pohybuje.

Problému konstrukce výběrových odpovědí spočívá právě v jejich **výběru**, tj. v anticipaci (typických) chyb, jichž se studující dopouštějí a v začlenění těchto chyb do stručně stylizovaných nabídnutých odpovědí. I když vyloučení prvku zjevné nesmyslnosti zužuje pole pro výběr chybných odpovědí, zůstává i tak k dispozici poměrně široký okruh možností. Proto se doporučuje respektovat další požadavek: **zahrnout do souboru odpovědí typické chyby**.

Z toho je patrné, že zatímco úskalím činnosti studujících je nalézt **správnou** odpověď, (odpovědi; řešení) představuje jedno z hlavních úskalí **autorovy** práce nalézt přiměřeně **chybné** odpovědi. Tím se také dostáváme k otázce **počtu** předkládaných možností.

Obecně lze říci, že optimální počet je dán počtem typických chyb. V praxi orientačně platí, že nejvhodnější je zařazovat 3 až 4 nabídky odpovědí. Tím není řečeno, že autor nemůže nabízet 2 nebo 5 odpovědí. Pokud by studující volil odpověď zcela náhodně, platí, že větší počet nabídek snižuje možnost náhodně správné volby - to mluví ve prospěch co největšího počtu variant.

Při výkladu vzniku substitučních tuhých roztoků s omezenou rozpustností složek byly probírány okolnosti, které ovlivňují stupeň rozpustnosti (velikost atomů, elektrochemické vlastnosti, krystalická stavba). Příslušnou testovou položku uvádí další příklad:

Vzájemná rozpustnost kovů jako složek ve slitinách, je tím větší, čím:

- a) je menší rozdíl ve velikosti jejich atomů;
 - b) jsou si bližší elektrochemickými vlastnostmi;
 - c) je poměr valenčních elektronů k počtu atomů blíže k číslu 1,4.
- Zakroužkujte odpověď nebo jejich kombinace, jež považujete za správnou (správné):

- | | | |
|-------|----------|-------|
| a) | b) | c) |
| a);b) | a);c) | b);c) |
| | a);b);c) | |

Poznámka: Protože všechna tvrzení jsou správná, jsou správné i jejich kombinace. Výběrový soubor je tedy tvořen pouze správnými odpověďmi.

Jak mnoho závisí správnost volby odpovědi na formulaci otázky (způsobu zadání problému) lze dokumentovat na tomto příkladu. Stačí totiž, aby autor testu změnil zadání a ihned se změnil také možnost výběru správné odpovědi. Místo: „Určete, která tvrzení jsou správná“ může použít pokyny:

Určete, které (resp. která) z tvrzení a); b); c) platí:

- platí pouze tvrzení a);
- platí pouze tvrzení b);
- platí pouze tvrzení c);
- platí současně tvrzení a); b);
- platí současně tvrzení a); c);
- platí současně tvrzení b); c);
- platí současně všechna tři tvrzení a); b); c).

Poznámka: Za těchto okolností jsou správné pouze 4 poslední odpovědi.

Dalším zpřísněním pokynu v testové položce může být např.:

Zvolte tvrzení **nejpřesněji** vystihující podmínky vzniku omezených tuhých roztoků.

Poznámka: *Tim se dosáhne toho, že je správná pouze poslední varianta z nabízených odpovědí. Všechna ostatní tvrzení jsou neúplná.*

Jednou z hlavních výhod volených odpovědí je (kromě snadné kontroly správnosti) i skutečnost, že je lze libovolně přesně formulovat. Předkládaný soubor tak může obsahovat škálu odpovědí od té s hrubou chybou přes částečně chybné nabídky až k odpovědím s jemnými nuancemi v (ne)přesnostech formulace.

Rozhodnutí o tom, které odpovědi autor do výběrového souboru zařadí, závisí na konkrétním didaktickém záměru a na analýze charakteristických chyb. V každém případě je však vhodnější dát přednost nepřesnosti (neúplnosti) před hrubou („hloupou“ nebo zcela náhodně zařazenou) chybou. Pro výběrový soubor odpovědí by však vždy mělo platit, že žádná z nabídek nemá sugestivně navozovat chybnou odpověď (nemá být záludným „chytákem“).

Do souboru výběrových odpovědí lze zařadit i „univerzální“ odpovědi, které mohou všude tam, kde je to pocitováno jako účelné, rozšířit pole alternativ. Jde o nabídku typu: „žádná odpověď z uvedených není správná“;

Příklady

Který z následujících kovů je při normální teplotě tekutý? Zakroužkujte příslušnou položku.

a) olovo; b) rtuť; c) cín; d) žádný z uvedených; e) žádný z kovů.

Poznámka: *Tekutost rtuti je natolik výjimečnou vlastností mezi kovy, že je v tomto případě vhodné rozšířit paletu nabídek o ty, které jsou uvedeny sub d) a e).*

(Málo známým faktem, že galium taje při 29,7 °C a může zůstat po roztavení a zpětném ochlazení tekuté i několik měsíců, není třeba komplikovat základní kurz chemie.)

V uvedeném případě představují odpovědi a); c); d); e) **chybné alternativy**. Stejně důležité je zařazovat takovou nabídku i jako **správnou odpověď** (potom je jedinou správnou odpovědí). Uvedený příklad lze modifikovat také tak, že kovy nabídnuté v jednotlivých alternativách by neobsahovaly rtuť.

Který z kovů je při normální teplotě tekutý?

a) olovo; b) cín; c) vizmut; d) žádný z uvedených; e) žádný z kovů.

Poznámka: *Jedinou správnou odpovědí je nyní nabídka d).*

V testech, které nejsou určeny pro klasifikaci, může autor zařadit i odpověď typu „nevím“. Obecně platí, že taková reakce by měla být posuzována z psychologického hlediska odlišně od zcela chybné volby. Vyjadřuje totiž nejistotu, opatrnost, snahu vyhnout se chybě bez zastírání nebo hádání.

Je důležité, aby správná odpověď byla v souboru nabízených variant umístěována **náhodně**. Její pořadí v zadávání by se tedy mělo měnit. Jinak by testovaná osoba mohla v tomto „zakódování“ odhalit nějaký systém.

Další vodítko pro volbu správné odpovědi může autor **nechtěně** poskytnout, jestliže nabízené správné odpovědi jsou textově **nejdelšími alternativami**. Přesné, resp. nejsprávnější (nejúplnější) odpovědi bývají výrazově složitější, než neúplné či nesprávné. Autor tedy musí dbát, aby více nabízených odpovědí obsahovalo přibližně tentýž počet slov, čímž přestane být délka vět vodítkem pro volbu správné z nich.

6 Hodnocení výsledků testu



6.1 Skórování výsledků

Nejčastěji používaným základem pro hodnocení každého testového výkonu i celkového výsledku je **systém bodování**. Proto je nutné nejprve rozhodnout, jak bodově ocenit výkon v každé testové položce. Nejjednodušší je přidělovat za každou správnou odpověď **1 bod** a za každou chybnou odpověď **0 bodů**. Výsledný součet dosažených bodů se v takovém případě označujeme jako **prostý skór**.



6.1 a/ Test má 25 položek a za každou správnou odpověď se přiděluje 1 bod.



Testovaná osoba odpověděla 19 položek správně. Jakou hodnotu má dosažený prostý skór X ?

Chceme-li ještě postihnout **kvalitu** výkonu v položce a rozlišit tak zcela správnou odpověď od odpovědi neúplné (a chybné), přidělujeme **různý počet bodů**. Na otázku kolik? není snadná odpověď. Může to záviset na závažnosti testovaného učiva, jeho obtížnosti, počtu předpokládaných operací (vedoucích k výsledku) apod. Tomuto problému je věnována řada teorií, které nebudeme rozvádět.

U **rozsáhlých tvořených odpovědí** můžeme zcela správnou odpověď ocenit 2 až 3 body, částečně správnou 1 bodem a chybnou 0 body.

U **přiřazovacích** odpovědí oceníme správné splnění tolika body, kolik párů prvků má být utvořeno. Zkoušený(á) tak získá tolik bodů, kolik dvojic prvků k sobě správně přiřadí (např. přiřazovat k názvům sloučenin chemické vzorce, k chemickým prvkům značky, k fyzikálním veličinám jednotky atd.). U **seřazovacích** odpovědí můžeme zcela správně stanovené pořadí nějakých (např. technologických) operací (výrobních metod, pracovních postupů) ohodnotit 3 body, nepodstatné přehození pořadí 1 bodem a závažně chybné v seřazení 0 body.

Podobně u **rozdělovacích** odpovědí (rozdělení elementů do dvou nebo více ohraničených skupin - např. rozdělení vlastností objektu na fyzikální a chemické) můžeme zcela správné rozdělení hodnotit 3 body, nepodstatnou záměnu 1 bodem a rozdělení s hrubou chybou (chybami) 0 body.

Při technických, fyzikálních, chemických a matematických výpočtech narážíme na problém rozlišování obtížnosti řešení, neboť výpočet zpravidla zahrnuje postupné uplatnění řady dovedností a vědomostí. Pokud je výsledek **jednoduchého** výpočtu součástí nabídnuté množiny odpovědí nebo vyústí v jednoznačnou tvořenou odpověď, je bodování jednoduché: 1; 0 popř. 2; 1; 0. U složitějších výpočtů však je třeba požadovat zaznamenání postupu řešení a podle počtu požadovaných operací (dílčích výsledků) přiřadit testové položce příslušný počet bodů (viz také přílohu 1).

I zde je celkový výkon v testu oceněn součtem bodů dosažených ve všech položkách. Jde však o **vážený skór** (protože jednotlivým odpovědím byla přidělena různá „váha“).



6.1 b/ Jaký je tedy rozdíl mezi prostým a váženým skórem?



6.1 c/ Jaké podstatné úskalí skrývá přidělování váženého skóru?

Při hodnocení dosažených výsledků (prostých skórů) mějme na paměti, že každý jednotlivý výsledek byl zjištěn s určitou chybou S_{EM} , jejíž hrubý odhad je dán výrazem:

$$S_{EM} = 0,43\sqrt{p}$$

kde k je počet položek. To znamená, že např. při 25 položkách a dosaženém hrubém skóru 15 bodů je $S_{EM} = \pm 5,43 \approx 2$ a výsledek de facto leží v rozmezí 13 až 17 bodů.



6.1 d/ Zjistěte rozmezí výsledku testované osoby, která v testu o 36 položkách dosáhla prostého skóru 20 bodů.



6.2 Analýza testu a interpretace výsledků

Citlivým problémem je interpretace výsledků a převod testového skóru na klasifikační stupnici. Řeší se přitom dvě důležité otázky:

- Jak rozdělit dosažitelný skór do několika tříd (odpovídajících klasifikačním stupňům) pro věrohodné srovnání výkonů studujících?
- Jak známkovat na rozhraní takto vzniklých intervalů?

Interpretace výsledků by měla vycházet z **analýzy testu**. Pokud je uskutečněna, stírá do určité míry rozdíly mezi standardizovanými a nestandardizovanými testy a zaručuje vyšší spolehlivost rozhodování.

Zopakujme si: Podrobné vymezení cílů testu a proporcionality počtu a obsahu položek ve vztahu k rozsahu a závažnosti učiva, které je předmětem testování, jsou zárukou **platnosti** testu. Analýza testu pak zahrnuje zjišťování:

- **obtížnosti** testu;
- **citlivosti** testu;
- **spolehlivosti**.

Jak už víme, **obtížnost** testu se vyjadřuje indexem obtížnosti: $I_o = 100 \cdot \bar{X} / X_{\max}$, kde $\bar{X}$ je průměrný skór a $X_{\max}$ je maximálně dosažitelný skór. Index obtížnosti I_o pak udává průměrné procento správně řešených testových položek. Za optimální hodnotu považujeme $I_o = 50$ (u testů s výběrovými odpověďmi o něco více, a to v nepřímé závislosti na počtu alternativ).

 **6.2 a/** Maximálně dosažitelný prostý skór v testu s výběrovými odpověďmi je 30 bodů.



Průměrný výkon vyšel $\bar{X} = 17,5$ bodů. Jaký je index obtížnosti testu?
Je to přijatelná hodnota?

Citlivost testu se měří koeficientem i_t , tj. poměrem mezi směrodatnou odchylkou s_x prostého (nebo váženého) skóru a průměrným skórem $\bar{X}$: $i_t = s_x / \bar{X}$.
Přijatelné hodnoty citlivosti testu jsou nad $i_t = 0,25$.

 **6.2 b/** směrodatná odchylka prostého skóru je $s_x = 4,6$. Průměrný skór $\bar{X} = 17,5$ bodů.
Vyjádřete citlivost testu: $i_t = \dots\dots$. Je to přijatelná hodnota?



Citlivost jednotlivých **položek** testu je definována odlišně. Index $i_p = \frac{2(X_l - X_h)}{n}$,

kde X_l je počet správných odpovědí na příslušnou testovou položku dosažený **osobami s lepším než průměrným výsledkem** v testu, X_h je počet správných odpovědí na tutéž položku dosažený **osobami s horším než průměrným výsledkem** v testu a n je počet zahrnutých osob. U dobře setavených testů má také v tomto případě dosahovat i_p hodnoty větší než **0,25** (pro každou položku).

Poznámka: Položky, které správně vyřeší většina studujících i položky, které naopak většina (či všichni) nevyřeší, jsou necitlivé a z testů se obvykle vyřazují.

 **6.2 c/** Testu se podrobilo 22 studujících. Průměrný skóre je 15,7 bodů. Žádný ze studujících přesně takového výsledku nedosáhl. Studujících s lepším než průměrným výsledkem je 10; 12 studujících dosáhlo horšího než průměrného výsledku. Na položku č. 11 odpovědělo správně z „lepších“ studujících 9 správně, kdežto z horších studujících jen 5.



Jaký je index citlivosti této položky? Co o můžete o jeho hodnotě říci?

Příklad analýzy testu

Analýze byl podroben test o 20 položkách, který vypracovalo 29 studujících. Většina položek byla výběrových, se 3 až 4 alternativami odpovědí. Individuálně dosažené bodové výsledky byly tyto:

5; 6; 7; 7; 8; 8; 9; 9; 9; 10; 10; 10; 11; 11; 11; 11; 12; 12; 12; 12; 13; 13; 13; 14; 14; 15; 15; 16; 17.

Sestavíme tabulku, usnadňující výpočty. Sloupec „Intervalová četnost“ je zařazen proto, že **n** je pro jednotlivé položky nízké. Intervaly vzniknou spojením několika sousedních skóre - vždy ovšem stejného počtu - dohromady a graf rozložení četností pak lépe vystihuje statistické rozložení.

Poznámka: tabulka je zpracována proto, aby byl co nejlépe znázorněn postup výpočtů; jinak lze samozřejmě všechny veličiny rychleji zjistit pomocí kalkulatoru nebo dokonce pomocí speciálního počítačového programu.

Prostý skóre X	Četnost f	Intervalová četnost	nX	X ²	nX ²
0	0	0			
1	0				
2	0				
3	0				
4	0	1	0	0	0
5	1		5	25	25
6	1		6	36	36
7	2	5	14	49	98
8	2		16	64	128
9	3		27	81	243
10	3	10	30	100	300
11	4		44	121	484
12	4		48	144	576
13	3	9	39	169	507
14	2		28	196	392
15	2		30	225	450
16	1	4	16	256	256
17	1		17	289	289
18	0				
19	0	0			
20	0				
Σ	29	29	320		3784

Nyní poměrně snadno zjistíme parametry testu:

Průměrný prostý skóre je $\bar{X} = 320/29 \div 11$

Index obtížnosti $I_0 = 100 \cdot \bar{X} / X_{\max} = 55$ (tato hodnota je blízká optimu).

Směrodatná odchylka prostého skóre $s_x = \sqrt{(3784 - 29 \cdot 11^2) / 29} = 3,07$

Citlivost testu vyjadřuje $i_t = s_x / \bar{X} = 3,07 / 11 = 0,28$. (má být větší než 0,25)

Spolehlivost testu je dána indexem $r = m(p - m) / p \cdot s_x^2$

kde **m** je medián (střední hodnota skóre) - v tomto případě je jeho hodnota shodná s aritmetickým průměrem:

$m = 11$; počet položek $p = 20$ a rozptyl skóre (druhá mocnina směrodatné odchylky) $s_x^2 = 9,4$.
 $r = 11 \cdot (20 - 11) / 20 \cdot 9,4 = 0,53$

(u standardizovaných testů se požaduje spolehlivost $r = 0,6$ nebo vyšší; pro nestandardizovaný test je hodnota $r = 0,53$ na hranici přijatelnosti).

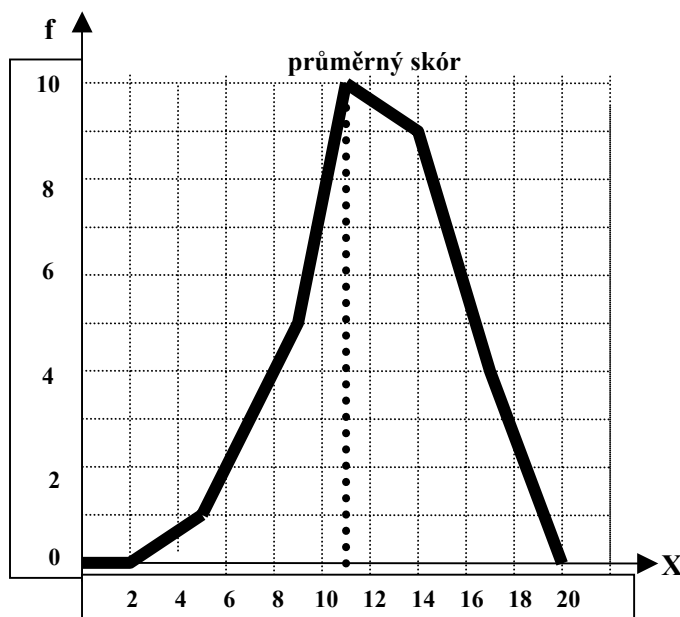
Poznámka: pro zvýšení spolehlivosti by bylo nutno rozšířit počet testových položek podle vztahu:

$$z = r_p (1 - r_d) / r_d (1 - r_p)$$

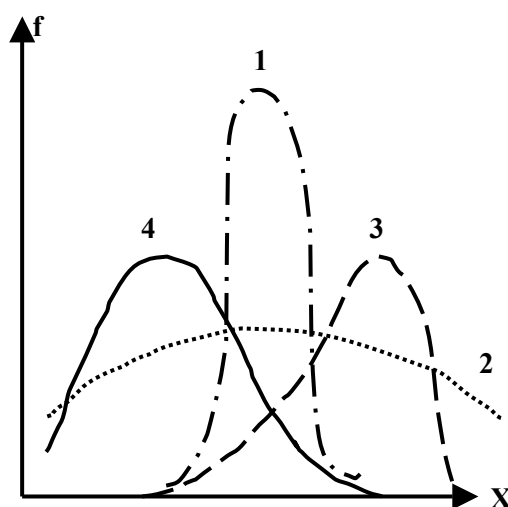
kde r_p je požadovaný a r_d dosažený koeficient spolehlivosti. Vyšla-li hodnota koeficientu $r_d = 0,53$ a požadujeme $r_p = 0,6$, vychází multiplikativní faktor $z = 1,33$. To by znamenalo zvýšit počet položek testu o 1/3; z 20 na 26 resp. 27.

Nakonec utvoříme polygon rozložení dosažených skóre a využijeme intervalových četností.

Graf závislosti dosažených bodových skóre na četnosti jejich výskytu



Graf závislosti bodových skóre na různých četnostech jejich výskytu



Vychází-li rozložení skóre příliš „úzké“ (1), potom test nerozlišuje mezi kvalitou výkonů (neboť dosažené výsledky se hromadí kolem průměru). Při příliš „širokém“ rozložení (2), rozlišuje test naopak jen v rámci průměrných výkonů. Posun vrcholu křivky směrem k lepším výkonům (3) znamená, že test dobře rozlišuje jen mezi podprůměrnými výsledky a naopak (4). V našem případě vyšlo četnostní rozložení skóre vcelku dobře a dosažené výkony pokrývají podstatnou část intervalu bodování (od 5 do 17 bodů) s minimálním výskytem extrémů.

Shrnutí výsledků:

Podařilo se sestavit test velmi přijatelné obtížnosti, spolehlivosti i citlivosti, který dobře využívá variační rozpětí výkonů (od 5 do 17 bodů) vcelku symetricky rozložených kolem průměru, bez extrémních hodnot na obou stranách stupnice (interval $<0; 20>$). Pro zvýšení spolehlivosti testu je vhodné rozšířit počet položek asi o 1/3.

6.2 d/ Uskutečňte analýzu testu o 15 položkách s tvořenými i výběrovými odpověďmi, který zpracovalo 22 studujících s těmito celkovými výsledky:

3; 4; 5; 6; 7; 7; 8; 8; 8; 8; 9; 9; 9; 10; 10; 10; 11; 11; 11; 12; 12; 13; 13.

Zjistěte průměrný skóre, jeho směrodatnou odchylku; index obtížnosti i citlivost testu a spolehlivost; nakonec slovně popište dosažené parametry testu.

6.2 e/ sestrojte graf (polygon) rozložení četností dosažených skóre (přitom slučujte vždy dva sousední skóre do jednoho intervalu); co můžete o získané křivce říci?

6.3 Převod skóre do klasifikační stupnice

Převod individuálních skóre na klasifikační stupně řeší učitelé různě, a to v závislosti na druhu testu, jeho účelu a přísnosti požadavků na výsledky. Někdy se dodržuje symetrické rozložení podle dosaženého průměrného výsledku, někdy nikoli. Při častějším testování se mohou dosažené výsledky i u jednoho a téhož studujícího dost lišit. Proto je účelné registrovat pouze body a klasifikaci provést až v závěru výukového období (tedy nepřevádět výsledky každého testu na známky). Zároveň je tak vcelku uspokojivě řešen i problém klasifikace studujících, kteří se při bodování ocitli na **rozhraní intervalů**, protože v takovém případě se problém klasifikace na rozhraní řeší jen jednou. Sestavování **pořadí** podle dosažených testových skóre (při němž se postupně načítají dosažené výsledky v testech) poskytuje učiteli **nejobjektivnější** pohled. Pololetní či závěrečná klasifikace pak může být uskutečněna buď v rámci „skupinové normy“ reprezentované školní třídou (studijní skupinou), anebo v rámci standardizované normy platné pro příslušnou populaci. U nás je druhý případ neuskutečnitelný, protože standardizované normy se teprve připravují, a to pouze pro vybrané vyučovací předměty. V řadě evropských států však takové normy již existují a umožňují porovnávat úroveň nejen individuálních výkonů, ale usuzovat i na kvalitu škol.

Rozdělit bodový skóre do 4 či 5 klasifikačních intervalů lze mnoha způsoby. Uvažte je a věnujte pozornost následujícímu úkolu.



6.3 Navrhněte čtyřstupňovou a pětistupňovou klasifikační škálu pro různá procentní rozdělení správných odpovědí v testu.



Při úvahách v oblasti skórování testovaných výkonů nebývá často brán zřetel na **důležitost** určitých částí učiva. Zkoušený, který např. získal méně než 25 % dosažitelného skóre je ohodnocen stupněm nesplnil, bez ohledu na to, **v čem** jeho znalosti a neznalosti spočívají. Právě odpověď na tuto otázku by však mohla být rozhodující. Je totiž rozumné předpokládat, že vždy existuje určitá část učiva, jehož neznalost vylučuje ovládnutí vztahů a souvislostí mezi jevy, ohrožuje či znemožňuje **postup dalšího učení** a je vážnou překážkou strukturalizace navazujícího souboru poznatků i jejich praktických aplikací. Proto by testy měly sestávat ze dvou podmíněných část. První část - spíše orientační - by prověřovala vědomosti, dovednosti a manipulace vytypované jako **základní**, s předepsaným vysokým bodovým skórem úspěšnosti (90 % a více). Splnění limitu by jednoznačně rozhodlo otázku prezentace další části, zkoumajícího **podrobnosti, hlubší souvislosti, méně významné učivo** apod. Splnění testu základního učiva znamená, že studující prospěl **dostatečně**. O lepší klasifikaci rozhodne až míra splnění druhé části testu. Takto koncipované testy by také lépe vyhověly požadavkům platnosti a znamenaly by „spravedlivější“ hodnocení.

Poznámka: Vhodným příkladem realizace záměru přísně posuzovat základní vědomosti jsou celostátně používané testy pro řidiče motorových vozidel. Je vypracováno a standardizováno mnoho desítek rovnocenných testových variant, předkládajících řadu otázek s výběrovými odpověďmi. Hodnocení **neprospěl** je uděleno už při neznalosti odpovědi na 10 % položek(!). To je nicméně zcela legitimní, protože znalosti dopravních předpisů a dopravních značek i dovednosti řešit dopravní situace mají být úplné.

Didaktický test může při správné manipulaci s výsledky poskytnout cenné **diagnostické** informace, i když byl původně zaměřen jinak. Autor testu (učitel) si má alespoň čas od času všimnout

souhrnných výsledků, dosahovaných více studujícími. Sledovat frekvenci chyb, vhodnost navrženého skórování, skóre dosažených nejlepšími i nejhoršími studujícími apod. Vyskytuje-li se např. určitá chybná odpověď velmi často, nehovoří tento fakt ani tak o nedostacích v práci studujících, jako spíše o nedostacích výuky (učebnice, skripta). Velký rozptyl výsledků u některých položek s výběrovými odpověďmi může svědčit o tom, že studující hádají (tj. nezvládli určité učivo) apod. Získané poznatky mohou totiž pomoci zlepšit řízení učebního procesu, pokud testujeme každou ucelenou, relativně samostatnou část výuky.

SHRNUTÍ

Didaktický test je objektivovaná (písemná) zkouška vědomostí a dovedností, která může mít funkci:

- kontrolní;
- klasifikační;
- diagnostickou;
- motivační;
- poznávací,

a je proto velmi důležitou součástí výuky. Musí však splňovat základní požadavky:

- platnosti;
- spolehlivosti měření;
- objektivnosti a
- citlivosti při rozlišování výkonů.

Také didaktické testy se sestavují až po důkladné analýze učiva a tvoří kontrolní i diagnostický doplněk plnění dílčích a specifických cílů ve výuce.

Testové položky jsou různého druhu a jsou v zásadě buď otevřené (tvořené), anebo uzavřené, přičemž obě kategorie lze ještě dále členit. Mají své výhody i nevýhody a velmi záleží na formulaci otázek i nabídnutém souboru odpovědí, který by měl zahrnovat typické chyby, jichž se studující dopouštějí. Formulaci otázek (testových zadání) je nutné věnovat mimořádnou pozornost - hlavně s ohledem na jednoznačnost a kontrolovatelnost odpovědí (výsledků). Určité nevýhody výběrových odpovědí lze do jisté míry kompenzovat, stejně jako nekorektnost postupu při testování.

Efektivita testu závisí také na uspořádání testových položek, instruktaži testovaných osob, správně stanoveném časovém intervalu ke splnění testu a vhodně navržené klasifikační stupnici vzhledem k procentnímu rozložení správných odpovědí.

Obsah testu by měl respektovat rozdělení učiva na základní, prohlubující, rozšiřující popř. nepovinné a učitel by tomu měl přizpůsobit také hodnocení výkonů v jednotlivých položkách i příslušných částech testu.

Častější sledování výsledků testů a jejich analýza mohou poskytnout učiteli cenné informace o výuce a zlepšit jeho pedagogickou práci.

SOUHRNNÉ CVIČENÍ 1



S a/ Test o **25** položkách byl absolvován s těmito výsledky:

medián bodových skóre **m = 15**; směrodatná odchylka **s = 3**.

Jaká je spolehlivost testu vyjádřená koeficientem reliability **r**?

S b/ Kolik položek by musel mít tentýž test, aby jeho spolehlivost stoupla na **r = 0,8**?

S c/ Určitý test má **28** položek a v každé jsou nabízeny **4 alternativy** odpovědí. Kolik správných odpovědí lze v testu vybrat zcela náhodně?

S d/ Jedna testovaná osoba zodpověděla v testu popsaném v předcházejícím zadání všechny položky a z toho **22** správně. Jaký je opravený počet správných odpovědí, chceme-li při hodnocení výsledku pokud možno vyloučit náhodné volby správné odpovědi?

S e/ Jaká je pravděpodobnost, že zkoušená osoba v testu o **40** položkách a **5** variantách odpovědi v každé z nich uhodne správnou odpověď v 16. položce?
Závisí hodnota takové pravděpodobnosti na počtu položek? Proč?

KLÍČ SPRÁVNÝCH ODPOVĚDÍ

1 K výhodám testování se počítá:

- možnost pečlivé přípravy a standardizace;
- sjednocení podmínek zpětného přenosu informací;
- rychlá informace o průběhu učení;
- možnost automatizace hodnocení výsledků;
- možnost alespoň dodatečného srovnávání výkonů různých subjektů a analýzy výsledků;
- možnost předem stanoveného způsobu hodnocení výsledků.

Testování má i další výhody, týkající se jen určitých osob. Někteří lidé mu dávají přednost, protože jim např. ústní vyjadřování činí větší obtíže, zejména v přítomnosti examinátora(ů).

2 a/ **Kontrolní** funkci plní testové položky tím, že poskytují **zpětnovazební informaci** (vnitřní i vnější). Tak, jako řešení podnětové situace signalizuje studujícímu osvojení aktuálního (specifického event. dílčího) cíle, má dosažení nebo překročení bodového limitu navenek indikovat splnění globálního (obecného, konečného) cíle.

Diagnostická funkce se uplatňuje, jsou-li testové položky splněny jen částečně nebo vůbec ne. Dobře konstruované testové položky mohou odhalit **příčinu** neúspěchu, **původ** chyby a ukázat tak, zda nezvládnutí určitého souboru poznatků je zaviněno nedostatečnou úrovní vstupních vědomostí, nevhodným uspořádáním či nepřiměřeností učiva nebo naopak nepozorností apod.

Motivační funkce je nepopíratelná - testováním se studující opakovaně a okamžitě přesvědčují o svých úspěších v učení a to má velký vliv na jejich další práci.

Klasifikační funkce je nejznámější a v pedagogické praxi nejpoužívanější.

Poznávací funkce je u testů jednou z nejméně využívaných. Tzv. „procvičovací“ testy mohou studující vyplňovat a následně vyhodnocovat sami, podle pokynů učitele nebo klíče správných odpovědí. Tak zjistí v čem chybovali a mohou si dodatečně doplnit své (ne)znalosti.

2 b/ Má-li test 35 položek ($p = 35$), medián dosažených bodů $m = 21$ při rozptylu $s^2 = 16$ (směrodatná odchylka $s = 4$), je jeho spolehlivost vyjádřena:

$$r = \frac{m(p - m)}{p \cdot s^2}$$

$$r = 21 \cdot (35 - 21) / 35 \cdot 25 \sim \underline{0,53}$$

2 c/ Zlepšení spolehlivosti testu na hodnotu $r = 0,65$ by znamenalo násobit počet položek číslem **1,6**, což by vedlo k rozšíření testu až na **56** položek.

2 d/ $I_0 = 57 \%$; to znamená, že je blízká optimální hodnotě (**50 %**).

2 e/ Hodnota $i_t = 0,2$ je příliš nízká; tzn., že test není příliš citlivý k výkonům testovaných osob.

3 a/ Ekvivalence testových forem lze dosáhnout volbou různých, ale přiměřených položek, změnou pořadí položek a/nebo malými změnami zadání, odlišnými numerickými hodnotami jinak identických příkladů atd.

- 3 b/**
- Správnou odpověď lze (na rozdíl od tvořené) uhádnout.
 - Často jsou problémy s formulacemi nesprávných odpovědí.
 - Někdy může být zvolena správná odpověď na základě nesprávné úvahy.
 - Vybírat odpovědi je zpravidla snadnější, než je tvořit.

- Formulace výběrových odpovědí může **napovědět nejvhodnější volbu** nebo snížit počet možností, které přicházejí v úvahu jako správné řešení (aniž testovaná osoba správnou odpověď skutečně zná).

3 c/ Opravený počet správných odpovědí je: $A_k = A_s - A_n / (y - 1) \sim A_k = 8 - 3 / (4 - 1) = 8 - 1 = 7$
Tzn., že do hodnocení výsledku zahrneme jen 7 z nich.

3 d/ Při stanovování počtu alternativ výběrových odpovědí se spokojme konstatováním, že se ustálila praxe zadávat až **čtyři** výběrové odpovědi. Při pěti a více možnostech se sice (obecně) snižuje pravděpodobnost náhodné volby správné odpovědi, ale zvyšuje se doba potřebná k prostudování alternativ a zpravidla také rostou potíže, které má autor testu s formulací vhodných výběrových prvků. Při menším počtu alternativ roste riziko náhodné volby správné odpovědi, ale lze je snížit počtem položek (viz zmínku o přiměřené spolehlivosti testů s pouhými dvěma alternativami odpovědí, pokud mají větší počet položek). Jinak závisí počet alternativ především na počtu **typických chyb**, kterých se testované osoby zpravidla dopouštějí.

***Poznámka:** Při pracném sestavování mnoha alternativních odpovědí pro jediné zadání dochází často k tomu, že u některých nabídek je jejich nesprávnost zcela zřejmá. Varianty pak neplní funkci **distraktorů** při rozhodování a jejich eliminace snižuje rozsah výběrového pole. Dodejme, že pokud si je testovaná osoba správnou odpovědí zcela jista, žádný počet alternativ ji nezviklá k chybné odpovědi.*

4 a/ Při odhadování času na vypracování testu studujícími přibližně platí, že čas, který **sám autor** testu potřebuje na jeho vypracování, se pro studující násobí **5** resp. čas, který potřebuje na vypracování **jiný odborník** než autor, se násobí **3**.

Velmi častou chybou učitele je, že **podcení časový faktor** a ponechává studujícím na vypracování testu příliš **málo času**. Vede to k oboustrannému a zcela zbytečnému zklamání a výsledky jsou pak méně spolehlivé.

I když nekorektní postup studujících při vyplňování testu nelze vyloučit úplně (snad kromě případů extrémního snížení počtu testovaných), lze mu účelně čelit. Např. zadáním několika ekvivalentních testových forem, náhodnou generací údajů v jinak stejných zadáních, zpřeházením pořadí testových položek, zvětšením vzdáleností mezi testovanými osobami a pozorným dohledem.

4 b/ K nezbytným informacím patří:

- rozsah časového intervalu určeného pro práci ;
- srozumitelný popis postupu při plnění testových zadání doplněný tolika praktickými příklady zácvikových úkolů, kolik různých typů zadání test obsahuje; (popis by měl zahrnout i potřebné pomůcky a doporučení, týkající se celého postupu práce včetně toho, jak zapisovat tvořené odpovědi, popř. jak označovat vybranou odpověď při volbě apod.);
- hlavní zásady, podle nichž budou posuzovány jednotlivé výkony - tj. jak bude hodnocena vynechaná odpověď, jak chybná, jak bude skórována dvojí odpověď (pokus o volbu dvou nebo více možností místo jedné), jak dodatečně změnit jednou napsanou, anebo vybranou odpověď atd.

***Poznámka:** Korektní by také bylo objasnění toho, zda je při hodnocení výsledků prováděna oprava na hádání správných odpovědí, a zda je tedy pro zkoušenou osobu výhodnější v případě pochyb či neznalosti volit jednu z alternativ náhodně nebo raději odpověď vynechat.*

5.1 a/ Tvořenou je odpověď, formulovaná přímo testovanou osobou. Může být doplňovací, tj. vzniknout pouhým doplněním vynechané části textu položky (dopsáním symbolu, písmene, zkratky, čísla, historického data, klíčového slova, názvu objektu či operace, jména), anebo rozsáhlejší formulací (zbytkem či celou větou, definicí apod.).

5.1 b/ Jde vesměs o rozsáhlé tvořené odpovědi.

6.1 a/ 19

6.1 b/ Prostý skór vyjadřuje celkový výsledek testu, kdy výkon v každé položce byl hodnocen 1 bodem, bez ohledu na obtížnost jednotlivých úkolů (otázek). **Vážený skór** se získá tehdy, jestliže autor testu postihuje obtížnost výkonů tím, že přiděluje odpovědím (řešením) v různých položkách různý počet bodů.

6.1 c/ Posuzování obtížnosti testových položek je do určité míry subjektivní.

6.1 d/ $S_{EM} = 0,43$. $\sqrt{36} \approx \pm 2,5$; rozmezí bodového skóru tedy leží někde mezi 17,5 až 22,5 body.

6.2 a/ $I_0 = 58,3 \%$, a to je hodnota blízká optimu.

6.2 b/ $i_t = 0,26$, což je hodnota na hranici přijatelnosti.

6.2 c/ Skupina testovaných se tedy rozpadá na 10 osob s celkovým výsledkem v testu lepším než průměr a 12 osob s celkovým výsledkem horším. Z početnější skupiny vyřadíme 2 osoby, jejichž výsledek se **nejméně** liší od průměrného skóru. Ze zbývajících 10 osob jsme např. zjistili, že počet správných odpovědí v této skupině tím klesl na 4. Rozdíl v počtu správných odpovědí obou skupin je tedy 5 a počet osob zahrnutých do šetření je 20. Citlivost položky č. 11 $i_p = 0,5$, což je velmi dobrá hodnota.

Kdyby počet správných odpovědí na položku č. 11 vyřazením 2 osob z „horší“ skupiny zůstal na hodnotě 5, byla by položka méně citlivá; její index by klesl na $i_p = 0,4$. To je ve shodě s tím, že vyřazené osoby s nejlepšími výsledky v „horší“ skupině by měly odpovědět spíše správně než jiné osoby z této skupiny. Protože se tak nestalo, považujeme položku za hůře rozlišující (méně citlivou).

Kdyby naopak vyřazením zmíněných 2 osob klesl počet správných odpovědí v „horší“ skupině na 3 (obě vyřazené osoby odpověděly ve zkoumané položce správně), zvýšila by se citlivost položky až na hodnotu $i_p = 0,6$.

6.2 d/

Prostý skór X	Četnost f	Intervalová četnost	nX	X ²	nX ²
0	0	0			
1	0				
2	0				
3	1	1	3	9	9
4	1	2	4	16	16
5	1		5	25	25
6	1	3	6	36	36
7	2		14	49	98
8	3	6	24	64	192
9	3		27	81	243
10	3	6	30	100	300
11	3		33	121	363
12	2	4	24	144	288
13	2		26	169	338
14	0	0	0		0
15	0		0		0
Σ	22	22	196		1908

Průměrný prostý skór je $\bar{X} = 196 / 22 \div 8,9$

Index obtížnosti $I_0 = 100 \cdot \bar{X} / X_{\max} = 59,3 \%$

Směrodatná odchylka prostého skóru $s_x = \sqrt{(1908 - 22 \cdot 8,9^2) / 22} = 2,7$

Citlivost testu vyjadřuje: $i_t = s_x / \bar{X} = 4/12 = 0,31$.

Spolehlivost testu (při m = 9) je dána: $r = m(p - m) / p \cdot s_x^2$; $r = 54 / 15 \cdot 7,29 = 0,51$

Shrnutí:

Podarilo se sestavit test optimální obtížnosti s dobrou citlivostí (rozlišující výkony v jednotlivých položkách), který využívá variační rozpětí výkonů (od 3 do 13 bodů), bez extrémních hodnot na obou stranách stupnice (interval <0; 15>). Spolehlivost testu daná indexem r je méně přijatelná.

Poznámka: spolehlivost testu by bylo možné zvýšit podle vztahu: $z = r_p (1 - r_d) / r_d (1 - r_p)$, kde r_p a r_d jsou požadované resp. dosažené hodnoty indexu spolehlivosti a z je násobek daného počtu položek v testu; tzn., že při

$z = 1,44$ by se počet položek měl zvýšit až na 22 (což potvrzuje známou skutečnost, že testy s méně než 20 položkami bývají nespolehlivé).

6.2 e/ Vrchol křivky znázorňující rozložení jednotlivých skóre v závislosti na četnosti jejich výkytu je posunut k vyšším hodnotám prostého skóre, takže test dobře rozlišuje jen v oblasti podprůměrných výkonů. Pokud nejde o záměr (např. test s převahou otázek na základní vědomosti), bylo by třeba **zvýšit obtížnost položek**.

6.3 Navržených rozdělení podle % správných odpovědí může být celá řada - např.:

výborně:	85 - 100 %	dobře:	20 - 49 %
velmi dobře:	50 - 84 %	nevyhovující:	0 - 19 %

nebo v případě pětistupňové škály:

výborně:	90- 100%	velmi dobře:	71 - 89 %	dobře:	35 - 70 %
dostatečně :	15 - 34%	nedostatečně:	0 -14 %		

Prísnejší klasifikace (např. v testu základního učiva) může vyžadovat:

výborně:	91- 100 %	dobře:	71 - 80%
velmi dobře:	81- 90%	nevyhovující:	0 - 70 %

VÝSLEDKY SOUHRNNÉHO CVIČENÍ 1

S a/ $r = 15(25 - 15) / 25 \cdot 9 \div \underline{0,67}$

S b/ Multiplikativní faktor pro zvýšení položek testu $z = 0,8 (1 - 0,66) / 0,66 (1 - 0,8) = \underline{2}$
Ke zvýšení spolehlivosti by tedy test musel mít dvojnásobek položek (50).

S c/ Při předložení testu o 28 položkách se 4 alternativami odpovědi v každé z nich osobám, které správné odpovědi **pouze hádají**, by byly výsledky samozřejmě různé. Pokud by takových osob bylo mnoho, zjistili bychom, že se počet správně uhodnutých odpovědí blíží číslu 7. Jinými slovy průměrný počet správně uhodnutých odpovědí (v reprezentativním vzorku testovaných osob, které se chovají popsáním způsobem) je dán:

$$\bar{A} = 28/4 = \underline{7}$$

S d/ Jestliže zkoušená osoba zodpověděla správně 22 položek z celkového počtu 28 a ke každé položce přitom náležely 4 možnosti odpovědi, je její opravený počet správných odpovědí dán:

$$A_k = 22 - 6/3 = \underline{20}$$

S e/ Pravděpodobnost, že zkoušená osoba zvolí správnou odpověď náhodně v kterékoli položce je závislá **pouze na počtu alternativ** a vůbec ne na počtu položek. Proto je v 16. položce stejná jako v kterékoli jiné, pokud u každé z položek je stejný počet alternativních odpovědí. Při 5 možnostech je tato pravděpodobnost rovna **1/5**.

7 Testy a efektivita vyučování (učení)



Zvláštním případem diagnostické funkce jsou testy „procvičovací“, jejichž výsledek může posloužit jako bezprostřední zpětná informace studujícím o jejich pokrocích v učení. Učitel test nejprve zadá a po jeho vypracování spolu se studujícími opravuje případné chyby a zdůvodňuje správnost jednotlivých odpovědí. Takové testy mohou být i součástí učebnice spolu s klíčem řešení. Doporučuje se u každé odpovědi v klíči ještě uvést odkaz na tu část výukového textu, která testované učivo probírá. Příloha 1 takový test uvádí pro ilustraci.

7.1 Příprava testování

Jde o podrobnější sledování výsledků vyučování a učení. Nejprve pořídíme seznam všech **proměnných**, které chceme sledovat a zároveň vypracujeme časový plán. Návrhy je vhodné konzultovat se statistikem a pedagogem, který má potřebné experimentální zkušenosti. Mezi nejčastěji **přímo sledované** proměnné patří:

- celkový počet chyb v každé testové položce;
- druh (kvalita) jednotlivých chyb;
- počet a druh chyb každé testované osoby;
- čas potřebný k absolvování testu;

Aby bylo možné tyto proměnné zpracovat, požaduje se registrace jednotlivých odpovědí (voleb). Statistická analýza pak umožní sledovat **další proměnné**:

- průměrný počet chyb i správných reakcí v každé položce;
- průměrný počet chyb i správných odpovědí každého studujícího v testu;
- průměrný čas potřebný k absolvování celého testu;
- průměrný bodový skóre;
- rozložení bodového skóre pro jednotlivé položky testu;
- hodnoty rozptylů u zjišťovaných statistických veličin.

Pro ověřování testů pak přichází v úvahu určení platnosti, spolehlivosti, objektivnosti a citlivosti, popsané už dříve. K tomu je vhodné připravit **časový plán**.



7.1 Co zejména by měl podle Vašeho názoru zahrnout?



7.2 Zpracování výsledků šetření analýzou dat

Zahrnuje zpravidla absolvování **vstupního testu** (pretestu), který zjišťuje předchozí úroveň znalostí. Pak následuje vlastní výuka a **test zařazený po jejím skončení** (výstupní test). Je mimořádně důležité, aby šlo o **ekvivalentní formu** vstupního testu. Jen tak lze spolehlivě zjistit, jaké vědomosti a dovednosti si studující osvojili během výuky. Po 4 až 6 týdnech může být ještě zařazen **retest** (posttest), který zjišťuje vědomosti a dovednosti uložené v dlouhodobé paměti. Poslední fází je třídění a zpracování získaných údajů i jejich transformace do konečných úprav.

Pro revizi a rekonstrukci testu je nutné uskutečnit **frekvenční a typovou** analýzu chyb.

Frekvenční a typová analýza chyb

V testových položkách se studující mohou dopouštět chyb různé povahy. Při poměrně značném, ale účelném zjednodušení, můžeme např. rozlišit tyto typy:

- logické chyby;

- chyby z neznalosti;
- chyby z nepozornosti (náhodné);
- chyby bez zjevné souvislosti se zadáním (nesmyslné, „hloupé“ chyby);
- chyby spočívající v neúplnosti odpovědi.

Poznámky: - Při nepozornosti jde zejména o náhodnou chybu, jako např. přehlédnutí nebo přepsání. Je častá při úpravách matematických výrazů, numerických výpočtech nebo při vyhledávání údajů v tabulkách, grafech, mapkách, schématech, normách apod.

- Některé chyby se nedají zjistit, ani klasifikovat.



Pokuste se uvést příklady každého z výše uvedených 5 typů chybných odpovědí na Vámi zkonstruované otázky.

Kromě zcela chybných sem spadají ještě neúplné a nepřesné odpovědi. Při revizi návrhu testu můžeme na základě shromážděných odpovědí **sestavit tabulku**, zachycující odpovědi všech respondentů. Kvalitu odpovědi můžeme symbolizovat např. takto:

- + správná odpověď;
- - chybná odpověď;
- N neúplná odpověď;
- 0 odpověď chybí;

Příslušné znaky se pak postupně zanesou do tabulky. Ta má tolik řádek, kolik je v testu položek a tolik sloupců, kolik bylo testovaných osob. **Řádky** tak podávají přehled o absolvování každé položky a **sloupce** zachycují úspěšnost postupu každé testované osoby. Proto je **řádková analýza** tabulky východiskem pro revizi testu. Po sestavení tabulky je nutné sledovat každou zařazenou položku z hlediska počtu a typu chyb. Shledá-li se, že v určité položce se několik osob dopustilo **téže** chyby, jde o **konstantní hromadnou** chybu, jež je signálem pro přepracování příslušné části učebnice. Přitom podle převažujícího typu chyby lze usuzovat, zda bude efektivní pozměnit znění instrukce nebo její rozsah, popř. zvětšit počet příkladů, zařadit opakování, procvičování, shrnutí učiva apod. Shledá-li se naopak, že v určité položce se vyskytují chyby typově odlišné, jde o **variabilní chybu**, která sice svědčí o obtížnosti učiva, ale nemusí být nutně výzvou k rekonstrukci výukového textu.

Následuje hypotetický příklad části tabulky pro frekvenční a typovou analýzu chyb. Pro jednoduchost předpokládejme, že se test skládal pouze z 15 položek a vypracovalo jej 25 studujících (A - Z). Dosažené prosté skóry byly u jednotlivých studujících následující:

14; 13; 12; 12; 11; 11; 10; 10; 10; 10; 9; 9; **2**; 8; 8; 8; 7; 7; 7; 6; 6; 5; 5; 4; 4

Kromě prostých skóre jednotlivců, počtu nesprávných (neúplných) a chybějících odpovědí zaneseme do tabulky také obtížnost (I_o) a citlivost (i_p) **každé** položky.

Poznámka: Index obtížnosti I_o položek se dá vypočítat i z dále uvedených (neúplných) tabulkových údajů; pro výpočet indexu citlivosti i_p bychom potřebovali znát pro každou položku vždy počet správných odpovědí té poloviny studujících s celkově nadprůměrným skóre v testu a počet správných odpovědí druhé poloviny - té s celkově podprůměrnými skóre v testu. Předpokládejme, že takové údaje byly k dispozici.

Položky	Studenti	A	B	C	D	E	F	G	H	I	J ...	Odpovědi	+	-	N	0	I_o %	i_p
1		+	+	+	+	+	+	N	+	+	+	...	20	3	1	1	80	0
2		+	-	+	+	+	+	+	+	+	+	...	19	3		3	76	0,13
3		+	0	+	-	+	+	N	+	+	+	...	15	6	1	3	60	0,4
4		+	+	+	+	+	-	+	+	+	-	...	13	7	3	2	52	0,29
5		+	-	+	-	+	+	-	+	+	+	...	17	5		3	68	0,21
6		-	-	+	-	-	-	-	+	+	-	...	16	9			64	0,25
7		-	+	N	0	+	+	+	+	+	+	...	14	9	1	1	60	0,29
8		+	+	+	-	+	+	+	+	+	+	...	13	7	5		52	0,21
9		+	+	+	+	+	+	+	+	+	+	...	13	8	2	2	52	0,28
10		-	-	-	-	+	-	-	-	+	-	...	15	7		3	60	0,24
11		-	+	+	-	+	+	+	-	+	-	...	9	15		1	36	0,38
12		+	+	0	+	-	-	+	0	-	+	...	15	8		2	60	0,29
13		+	0	+	N	-	+	0	-	+	0	...	14	7	4		52	0,13
14		-	+	+	+	+	-	0	+	+	+	...	10	11	1	3	40	0,25
15		-	+	+	0	+	+	-	+	+	0	...	12	9	4		48	-0,08
												Σ	215	114	22	24		

Věnujme nyní pozornost údajům v tabulce. Méně než 50 % správných odpovědí v položce 11 a 14 signalizuje určité nedostatky. U obou je proto nutné udělat **typovou analýzu chyb**. Při ní se zjistilo, že v položce 11 jde o chyby **různého** typu, zatímco v položce 14 jde o jedenáct chyb **téhož** typu. Tato konstantní chyba svědčí o **nedostacích v prezentaci učiva**. V položce 11 jde o chyby variabilní, jichž se testované osoby dopouštějí pravděpodobně v důsledku přílišné obtížnosti úkolu zadaného v testu.

Míra konstantnosti chyb je závažná pro diagnózu a revizi příslušné části učebnice, míra variability chyb zase pro diagnózu činnosti jednotlivců a je zároveň měřítkem homogenity studujících.

Dalšími „podezřelými“ jsou položky č. 1; 2; 5; 8; 13 a 15. Jak je patrné, byla u nich zjištěna příliš nízká citlivost (i_p menší než **0,25**), což znamená, že dobře nerozlišují mezi výkony jednotlivých studujících. V případě prvních dvou položek to nemusí vadit, pokud jsou **záměrně** zařazeny na samém počátku testu právě pro svou nízkou obtížnost. U položky č. 15 dokonce vyšel index citlivosti záporný. Znamená to, že ji o něco (málo) lépe řešili studující s podprůměrnými celkovými skóre než ti s nadprůměrnými skóre. Pro další použití testu proto lze doporučit náhradu položek 5; 8; 13 a 15 jinými (týkajícími se ovšem téhož učiva). Položku č. 14 může učitel v testu ponechat s tím, že se ve výkladu příslušného učiva příště sám „polepší“.

Poznámky: - Je třeba zvlášť upozornit na to, že častější výskyt nesprávných a zejména chybějících odpovědí u položek s nejvyššími pořadovými čísly, může svědčit nikoli o neznalostech, ale o **nedostatku času** na vypracování testu.
 - Pokud je vzdělávací instituce vybavena počítačovou sítí (popř. má-li k ní přístup), lze výsledky hromadně ukládat a pak statisticky zpracovat speciálním programem.

Ačkoli jsme uvedený příklad použili především pro účely analýzy testových položek a k podrobnějšímu nahlédnutí do vlastní výuky, pokusme se ještě posoudit i kvalitu testu samotného.



7.2 Vypočítejte parametry testu, nakreslete polygon rozdělení jednotlivých skóre v závislosti na četnosti jejich výskytu a připojte shrnující komentář.



7.3 Účinnost výuky

Pedagogická účinnost může být zkoumána v zásadě dvojím způsobem. Buď srovnáváním výsledků učení dvou skupin, z nichž každá pracuje za standardních podmínek z jiných (obsahově stejně zaměřených) materiálů, anebo je výsledek studia vyjadřován porovnáním **výchozí a konečné fáze** učení (tedy bez referenční skupiny).

Poznámka: Konečnou fázi může být přitom míněn nejen stav dosažený bezprostředně **po** ukončení studia, ale i stav ještě pozdější, měřitelný obvykle **retenčními testy** vědomostí a dovednost (tj. testováním znalostí uložených v dlouhodobé paměti).

V případě srovnávacího experimentu s referenční skupinou se dost obtížně zajišťují podmínky studia (stejný vzorek studujících, stejné učivo apod.). Proto se spíše používá porovnávání výchozího a konečného stavu téže skupiny studujících. Vhodným nástrojem jsou **Huberova kritéria G** /8/:

$$\% G(I) = 100 (X_{vt} / X_{max})$$

kde X_{vt} je studujícím dosažené bodové skóre ve výstupním testu a X_{max} je **maximálně** dosažitelné skóre v tomto testu. Například v případě studentky, která dosáhla 42 bodů z 50 možných, je efektivita vyjádřena 84 %. Chceme-li měřit (lépe řečeno kvalifikovaně odhadovat) efektivitu učení v celé skupině, vypočítáme **aritmetický průměr G(I)**.

Druhé Huberovo kritérium efektivity zahrnuje úroveň studujícího s ohledem na **počáteční stav** vědomostí a měří tak **relativní zisk** (anglicky **GAIN**):

$$\% G(II) = 100 (X_{vt} - X_{pre}) / X_{max}$$

kde X_{pre} je bodové skóre dosažené ve vstupním testu, jehož hodnota se **odečítá** od skóre dosaženého ve výstupním testu. Jak už bylo řečeno vstupní test musí být **ekvivalentní formou** výstupního testu. Pouze za takových okolností je zjištěný výsledek spolehlivým kritériem účinnosti učení. Obdobně jako v předcházejícím případě se skupinový efekt odhaduje jako aritmetický průměr všech individuálně dosažených hodnot **G(II)**.

Za nejspolehlivější ukazatel efektivity učení je považováno třetí Huberovo kritérium:

$$\% G(III) = 100 (X_{vt} - X_{pre}) / (X_{max} - X_{pre})$$

I zde se aritmetický průměr zjišťuje z individuálně dosažených hodnot **G(III)**. Tím, že se X_{pre} odečítá i od hodnoty X_{max} , je zavedena nejen oprava na vědomosti, které si studující osvojil již **před** zahájením příslušné části výuky, ale také změna „normy“, kterou je maximálně dosažitelný skóre. Kromě údajů o relativním efektu učení je možné - avšak mnohem složitější - získat důležité informace o výsledcích učení také pomocí **korelací**. Zkoumání vzájemných závislostí mezi výsledkem vstupního a výstupního testu, výsledků těchto testů s výsledky z průběžných testů, všemi těmito výsledky a výsledky retenčních testů, dosažených výsledků a času potřebného ke studiu, umožňuje učitelům vnikat do složitého procesu lidského učení a shromažďovat informace pro optimalizaci jeho řízení. I zde platí, že měření efektivity lze zautomatizovat a zahrnout do souboru výpočtů týkajících se hodnocení účinnosti.

SHRnutí

Výsledků didaktických textů je možné také využít pro kvalifikované odhady účinnosti výuky. Proces má dvě fáze a zahrnuje:

- analýzu položek;
- zavedení kritérií pro vlastní zjišťování efektivity.

Účelem položkové analýzy je určit charakter a četnost chyb, kterých se studující dopustili a tak odhalit především hromadné konstantní a variabilní chyby. Přitom míra konstantnosti chyb je signálem k přepracování příslušných partií učebnice, resp. ke zlepšení ústní prezentace učiva učitelem.

Další možností, jak odhadovat účinnost výuky, je zavedení Huberových kritérií G, která využívají dosažených testových výsledků buď přímo (G I), anebo po aplikaci pretestů (tj. testů vědomostí, které studující získali před zahájením výuky popř. její zkoumané části (G II a G III)).

K hlubším statistickým rozborům, využívajícím testování hypotéz, korelací, popř. analýzy rozptylu apod. je nezbytná spolupráce dalších odborníků.



SOUHRNNÉ CVIČENÍ 2



S a/ Jakých chyb (z typologického hlediska) se dopustila osoba, která doplnila nebo zvolila následující odpovědi?

- Univerzita Karlova v Praze byla založena roku**1346**
- Je-li v Praze 14,00 hodin, je ve Filadelfii (USA) o 6 hodin**více**
- Podtrhněte názvy států, s nimiž sousedí Česká republika:
Německo; Rakousko; Slovensko; Itálie; Ukrajina; Bělorusko; Francie; Maďarsko; Polsko
- Vzroste-li poloměr koule na dvojnásobek, kolikrát se zvětší její objem?
○ dvakrát ● čtyřikrát ○ osmkrát
- Teplota varu vody je za normálního tlaku**pozvolná**..... °C

S b/ Vhodnou a poměrně často používanou pomůckou empirického ověřování učebnice je anketa, zkoumající na závěr studia formou dotazů názory, postoje, připomínky a náměty pokusných osob. Uvažte, co by vás jako autory zajímalo, navrhnete otázky a nezapomeňte přitom, že anketa má být statisticky zpracovatelná.

S c/ Stanovte **průměrné** hodnoty efektu učení pomocí všech tří Huberových kritérií ve skupině 10 studujících, jsou-li výsledky vstupního a výstupního testu shrnuty v následující tabulce. Maximálně dosažitelný skóre testu byl $X_{\max} = 50$

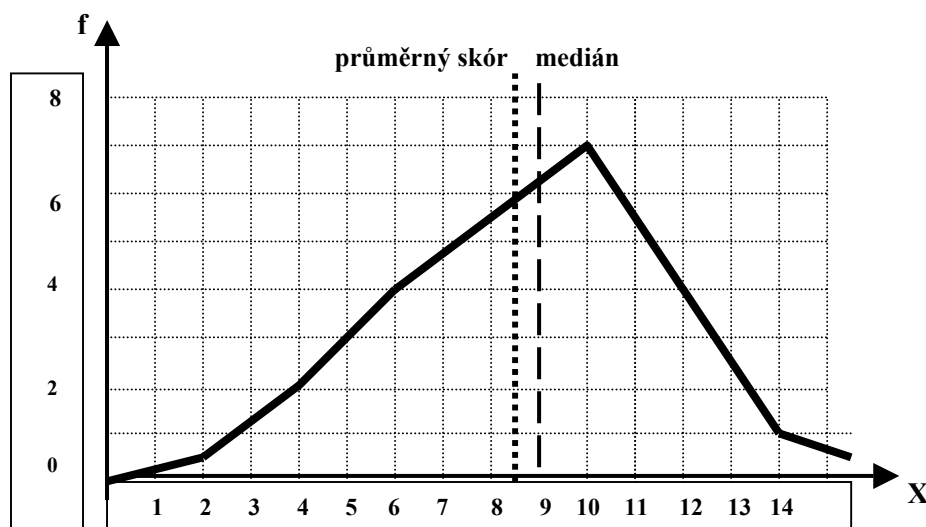
Testovaná osoba	X_{pre}	X_{vt}	Testovaná osoba	X_{pre}	X_{vt}
1	7	41	6	12	46
2	13	43	7	9	46
3	24	50	8	6	42
4	3	39	9	15	39
5	11	40	10	14	47

KLÍČ SPRÁVNÝCH ODPOVĚDÍ

7.1 Časový plán by měl zejména zahrnout:

- výběr studujících;
- rozmnožení a distribuci testu;
- předtestovou instruktáž;
- shromáždění záznamů;
- vyhodnocení získaných datových souborů;
- opravy a úpravy testu.

7.2 $I_0 = 57,3 \%$; $s_x^2 = 7,3$; $s_x = 2,7$; $r = 0,49$; $i_t = 0,31$



Shrnutí výsledků

Test je přiměřeně obtížný, citlivý, s nízkou spolehlivostí (což se u testu o 15 položkách dá předpokládat). Polygon rozdělení jednotlivých skóre v závislosti na četnosti jejich výskytu ukazuje na vcelku rovnoměrné rozložení dosažených výsledků v téměř celé šíři intervalu (od 4 do 14 bodů) a symetrický kolem průměrného skóre (s rychle klesajícím výskytem extrémů).

Spolehlivost testu vyjádřenou indexem alespoň $r = 0,55$ by bylo možné zlepšit zvýšením počtu položek na 19 ($z = 1,27$).

VÝSLEDKY SOUHRNNÉHO CVIČENÍ 2

S a/ Jde patrně o následující chyby:

- Univerzita Karlova v Praze byla založena roku ...**1346**... (neznalost)
- Je-li v Praze 14,00 hodin, je ve Filadelfii o 6 hodin ...**více**... (logická chyba)
- Podtrhněte názvy států, s nimiž sousedí Česká republika:
Německo; Rakousko; Slovensko; Itálie; Ukrajina; Bělorusko; Francie; Maďarsko; Polsko
 (neúplná resp. nepřesná odpověď)
- Vzroste-li poloměr koule na dvojnásobek, kolikrát se zvětší její objem?
☐ dvakrát ☒ čtyřikrát ☐ osmkrát (logická chyba)
- Teplota varu vody je za normálního tlaku ...**pozvolná**... °C (nesmyslná chyba)

S b/ Ukázkou možného řešení je připojený návrh znění anketního lístku (který může být také součástí počítačem prezentovaného výukového textu a po vyplnění odeslán elektronickou poštou):

Jméno Škola..... Věk. Datum

1. Cítil(a) jste při práci učebnicí větší únavu **ANO** resp. menší únavu **ANO** než při studiu jiných materiálů?

2. Připadala vám práce: **SNADNÁ - NEPOHODLNÁ - ZÁBAVNÁ - RUŠIVÁ**
(můžete vybrat i 2 odpovědi).

3. Vzpomínáte si na některé zvlášť obtížné úkoly nebo otázky? **ANO - NE**. Pokud ano, uveďte je:

.....

4. Vzpomínáte si na nějaké zvlášť snadné úkoly nebo otázky? **ANO - NE**. Pokud ano, uveďte je:

.....

5. Líbí se vám grafická úprava textu?

ANO, VELMI ANO SPÍŠE NE NE PŘÍLIŠ NA MNE NEZAPŮSOBILA

6. Považujete obrázky, schémata, grafy, diagramy, náčrtky za názorné?

ANO, VELMI ANO SPÍŠE NE NE

7. Pracoval(a) jste s textem samostatně (bez cizí pomoci)?

ANO VĚTŠINOU ANO VĚTŠINOU NE

8. Po nynějších zkušenostech dal(a) byste učení z tohoto typu učebního textu přednost před učením z jiných pomůcek?

ROZHODNĚ ANO ANO JEN V NĚKTERÝCH TÉMATECH NE

9. Považujete testování za účelnou a přijatelnou formu kontroly toho, čemu jste se naučil(a)?

ANO NE PŘÍLIŠ MI NA TOM NEZÁLEŽÍ

10. Co byste do obsahu ještě doplnil(a)?

11. Co byste změnil(a)?

12. Co byste vypustil(a)?

S c/ Pro výpočet průměrných hodnot všech tří Huberových kritérií je nejlépe sestavit tabulku, která uvádí v řádcích všechny testované osoby a ve sloupcích postupně hodnoty X_{pre} , X_{vt} i hodnoty zlomkových částí tří kritérií dosažených každou osobou.

Poznámka: postupné výpočty se v tabulce uvádějí pro názornost; v praxi by se řešily automaticky.

Testovaná osoba	X_{pre}	X_{vt}	$X_{vt} \cdot 100 / X_{max}$	$\frac{(X_{vt} - X_{pre}) \cdot 100}{X_{max}}$	$\frac{(X_{vt} - X_{pre}) \cdot 100}{X_{max} - X_{pre}}$
1	7	41	82	68	79
2	13	43	86	60	81
3	24	50	100	52	100
4	3	39	78	72	77
5	11	40	80	58	74
6	12	46	92	68	90
7	9	46	92	74	90
8	6	42	84	72	82
9	15	39	78	48	69
10	14	47	94	66	92
			$\Sigma \quad 866$	$\Sigma \quad 638$	$\Sigma \quad 834$

V daném případě pak jsou průměrné hodnoty kritéria pro 10 člennou skupinu studujících:

$$\bar{G} \text{ (I)} = 86,6 \% \quad \bar{G} \text{ (II)} = 63,8 \% \quad \bar{G} \text{ (III)} = 83,4 \%$$

Poznámky: - Jak je patrné, jsou to hodnoty odlišné, i když vycházejí z téhož souboru dat. Proto je třeba vždy uvést, jak bylo kritérium efektivity počítáno a popř. srovnávat pouze hodnoty korespondujících veličin.

- Za nejvhodnější míru efektivity výuky lze považovat kritérium $\bar{G} \text{ (III)}$.
- O účinnosti výuky by nejspolehlivěji vypovídala průběžná měření přírůstků vědomostí a dovedností, pokud by pro jejich postupné osvojování existovaly normy. Takové normy je nicméně možné alespoň zkusmo sestavit na základě analýzy obsahu a časových etap učebního plánu kteréhokoli předmětu.



LITERATURA

1. Byčkovský, P.: Tvorba didaktického testu. ČVUT, Praha 1982
2. Hniličková, J. - Josífk, M. - Tuček, A.: Didaktické testy a jejich statistické zpracování. SPN, Praha 1972
3. Hrabal, V. - Lustigová, Z. - Valentová, L.: Testy a testování ve škole. PedF UK, Praha 1992
4. Munzar, Z.: Možnosti a meze pedagogických testů ve vzdělávání dospělých. Institut pro výchovu vedoucích pracovníků ministerstva průmyslu ČSR, Praha 1987
5. Mužič, V.: Testy vědomostí. SPN, Praha 1971
6. Půlpán, Z.: Zásady sestavování didaktických testů. Ústav rozvoje vysokých škol, Praha 1986
7. Reisenauer, R.: Metody matematické statistiky a jejich aplikace. SNTL, Praha 1965
8. Řešátko, M.: Didaktické testy ve školní praxi. SNTL, Praha 1974



PŘÍLOHA 1



Ukázka souboru položek procvičovacího testu

(učební předmět Moderní vzdělávací technologie - 2. ročník bakalářského studijního programu)

1. Posuďte správnost dále uvedených tvrzení a označte křížkem ta, která považujete za **pravdivá**.

- ☐ a) Přímá prezentace učiva učitelem je při frontální výuce časově úspornější než při skupinovém vyučování, založeném na řešení problémových situací.
- ☐ b) Při frontálním vyučování má učitel lepší možnosti poznat schopnosti, vědomosti a dovednosti jednotlivých studujících.
- ☐ c) Z hlediska možností formovat osobnost studujících, poskytuje frontální výuka učiteli mnohem více příležitostí než skupinové vyučování.
- ☐ d) Skupinové vyučování zvyšuje sebereflexi studujících ve smyslu realističtějšího poznání vlastních schopností.
- ☐ e) Vyučování ve skupinách má - ve srovnání s frontální výukou - omezenější možnosti při osvojování rozumových i psychomotorických dovedností.
- ☐ f) K hlavním výhodám skupinového vyučování patří jednodušší organizace výuky a omezení výskytu konfliktových situací.
- ☐ g) Při rozdělování studujících do skupin je třeba přihlížet k jejich dosavadním výkonům.
- ☐ h) Skupiny nemají být výkonově stejnorodé, tj. mají sestávat z jedinců s rozdílnými výkony.
- ☐ i) Z hlediska pedagogických cílů výuky jsou různé postupy při tvorbě skupin rovnocenné.
- ☐ j) Uskuteční-li učitel sociometrické měření vazeb mezi studujícími, měl by s jeho výsledky studující vždy seznámit.

Za každé správné označení pravdivého výroku je 1 bod. Za označení nepravdivého výroku je -1 bod (nejvýše 4 body).

2. Sociometrická matice vztahů ochoty resp. neochoty spolupracovat mezi osobami A až J

Jméno	A	B	C	D	E	F	G	H	I	J	Σ_p	Σ_n
A		⊕	0	0	⊕	∅	0	0	0	+	3	1
B	⊕		0	+	0	0	-	-	0	0	2	2
C	0	+		0	0	-	0	0	+	0	2	1
D	+	-	-		0	0	-	0	+	+	3	3
E	⊕	0	0	-		0	0	⊕	+	-	3	2
F	∅	+	+	-	0		0	0	0	0	2	1
G	∅	0	-	0	+	+		0	-	∅	2	3
H	0	0	-	+	⊕	0	0		-	0	2	2
I	+	+	-	-	0	0	0	0		0	2	2
J	0	0	+	0	+	0	∅	0	0		2	0
Σ_p	4	4	2	2	4	1	0	1	3	2		
Σ_n	2	1	4	3	0	2	3	1	2	2		

a) Kolik **oboustranných preferencí** obsahuje tato sociometrická matice? (1 b)b) Kolik **negativních voleb** vyznačila osoba E (v této matici) a kolik jich sama dostala od ostatních? (2 b)

3. Použijte sociometrické matice a nakreslete individuální sociogram osoby B. (2 b)

B

4. Situace předložená k řešení není informačně zcela určena a skrývá více možností řešení. Jde o **úkol** nebo **problém**? Podtrhněte tu ze dvou možností, kterou považujete za správnou. (1 b)
5. Určete souhrnný počet komunikačních vazeb **K** ve skupině, která má 6 členů. **K** = (1 b)
6. Jedna skupina má řešit modelovou konfliktní situaci. Podtrhněte **dvakrát** ty z metod, které pro tento účel považujete za **nejvhodnější** a jednou podtrhněte metody, které jsou rovněž vhodné: **diskuse; situační metoda; inscenace; soutěživá hra.**
(Zcela správná odpověď **2 b**; částečně správná **1 b**)
7. Učitelem připravená inscenace má vyústit zcela určitým způsobem. K tomu účelu vhodně vybral aktéry a podrobně je seznámil s výchozí situací i jednotlivými rolemi. Jde o strukturovanou inscenaci?
☐ **ANO** ☐ **NE** (Označte křížkem do rámečku správnou odpověď.) (1 b)
8. Uveďte 3 nejdůležitější zdroje, z nichž může učitel čerpat náměty problémových situací nebo už vypracované případové studie, konfliktové či rozhodovací situace, inscenace, cvičení a hry:
a) ; b) ; c) ;
(1 nebo 2 zdroje **1 b**; 3 zdroje **2 b**)
9. Uveďte dva hlavní důvody, proč topografické uspořádání klasické třídy není vhodné pro skupinovou výuku.
a) ;
b) ; (2 b)
10. Jakými způsoby lze hodnotit výsledky jednotlivců při skupinové výuce? Napište.
.....
.....
..... (2 b)
11. Podtrhněte z dále uvedených technických prostředků a produktů **všechny** ty, které se pod označením **multimediální** využívají v moderní výuce.
a) internet (včetně webových stránek a e-mailů) b) radar,
c) projekční filmový přístroj; d) magnetofonové nahrávky,
e) video-kazety, f) mikroskop;
g) počítače h) počítačové programy (software),
i) rozhlas j) televize
k) video-kamera l) digitální fotoaparát
m) transplantáty n) pneumatická zařízení. (1 b)
12. Patří rozdělování učiva do menších celků k základní didaktické úpravě studijních materiálů určených pro distanční studium? **ANO** **NE**
(Označte křížkem do rámečku správnou odpověď.) ☐ ☐ (1 b)
13. Posuďte následující tvrzení a označte křížkem ta, která považujete za **správná**:
☐ a) Za **vnitřní zpětnou** vazbu považujeme výsledky hodnocení samostatných prací, provedených tutorem pro distančně vzdělávající instituci.

- ☐ b) Za **vnější zpětnou vazbu** považujeme odpovědi studujících na otázky položené přímo v textu výukových materiálů a řešení drobných úkolů, uvedených tamtéž.
- ☐ c) Velikost učebních dávek nelze obecně určit, protože závisejí na předchozích vědomostech studujících, povaze i významu daného učiva a dalších okolnostech.
- ☐ d) Obecně platí, že čím větší je dávka učiva, tím lépe se učivo zapamatuje a zároveň nedochází k atomizaci vědomostí.
- ☐ e) Jako h-text se označuje takový text, který umožňuje okamžitý přístup ke všem důležitým informacím, v něm obsaženým.
- ☐ f) Pokud se encyklopedie zpracovaná formou h-textu použije ve výuce, můžeme ji ztotožnit s výukovým h-textem.
- ☐ g) Při pořizování multimediálních učebních pomůcek není jejich technická kvalita tak důležitá; rozhodující je cena a oživení výuky, které přinášejí.
- ☐ h) Video-konference se ve výuce hlavně používá k přenosu důležitých jednání na konferencích, sympóziích, kongresech a seminářích.
- ☐ i) Interaktivní výuka slouží především skupinovému vyučování pro udržování nepřetržitých kontaktů.
- ☐ j) Interaktivní výuka zprostředkovaná počítačovým programem je velmi vhodná pro pamětní osvojování větších částí učiva.
- ☐ k) Distanční studium je nové označení dálkového studia.
- ☐ l) Distanční studium je nový název pro bývalé korespondenční kurzy.
- ☐ m) Existují zásadní rozdíly mezi večerním a distančním studiem.
- ☐ n) Distanční studium podnítilo psychologické výzkumy už v padesátých a šedesátých letech minulého století, ze kterých dodnes čerpá programované vyučování a učení.
- ☐ o) Mezi základní principy programovaného vyučování patří:
- kvantování informací;
 - aktivizace adresátů zadáváním otázek a úkolů;
 - zpětnovazební informace o dosahování výsledků učení;
 - individuální tempo učení;
 - účast adresátů na přípravě programu
- (Za každé správné označení pravdivého výroku je **1 bod**; za každé označení nepravdivého výroku je - 1 bod; nejvýše **4 body**).

14. Definujte pojem „distanční vzdělávání“:

(2 b)

15. Uveďte všechny základní technické jednotky nutné pro realizaci video-konference.

(2 b)

16. Hlavním **objektem**, na nějž je distanční vzdělávání zaměřeno, je **tutor** a hlavním **subjektem** je **studující**. (Označte křížkem správnou odpověď)

☐ ANO ☐ NE - je to přesně obráceně.

(1 b)

17. Co se rozumí **cílovou skupinou** při přípravě nového studijního programu nebo kurzu?

(1 b)

18. V dále uvedeném souboru technických prostředků a služeb podtrhněte **všechny** ty, které slouží ke komunikaci distančně studujících mezi sebou i s personálem vzdělávací instituce:

a) vlaková a autobusová spojení; b) telefon;

- | | | |
|-------------------------------|--|-------|
| c) fax; | c) rozhlas a televize; | |
| d) e-mail; | e) nástěnné noviny; | |
| f) taxislužba; | g) Internet; | |
| h) pošta; | i) video-konference; | |
| j) soukromé zásilkové služby; | k) technické vybavení školských úřadů. | (1 b) |

19. Z dále uvedených aktivit vyřaďte všechny ty, které v rámci DiV tutor nevykonává
(přeškrtnutím celého textu příslušné položky.)

- a) předává studujícím nové poznatky a opakuje s nimi osvojené učivo;
 - b) pomáhá vybírat studijní programy a kurzy podle potřeb zájemců o studium;
 - c) působí v psychologické poradně vzdělávací instituce;
 - d) seznamuje studující s podmínkami studia, jeho organizací a požadavky;
 - e) zpravidla také zkouší studující ve skupině, kde působí;
 - f) hodnotí samostatné práce a průběžné i celkové pokroky ve studiu;
 - g) vede tutoriály;
 - h) motivuje studující k práci a řeší s nimi i menší osobní problémy;
 - i) zaznamenává prezenci na všech tutoriálech;
 - j) shromažďuje, třídí, vyhodnocuje a odevzdává náměty studujících, týkající se návrhů změn v obsahu studia, připomínek ke studijním materiálům, testů i zkoušek včetně svých námětů.
- (2 b)

20. Součástí studijního programu jsou laboratorní práce z fyziky a chemie, konané v rámci laboratorních tutoriálů. Budou tyto tutoriály povinné, anebo nepovinné?
(Podtrhněte jednu z nabídnutých možností.)

(1 b)

21. Co jsou studijní centra?

.....

(2 b)

22. Je následující popis charakteristikou systému studia označovaného jako off-line nebo on-line, studium? (Připište na vytečkované místo na konci položky).

Zájemce se přihlásí ke studiu a zaplatí požadovaný poplatek. Dostane přiděleno přístupové heslo, umožňující vstup do softwarového studijního prostředí. Zpravidla je vyzván, aby si vytvořil vlastní webovou stránku, obsahující osobní údaje (podle uvážení) a povinně uvede svoji e-mailovou adresu. V kalendáriu sleduje harmonogram studia a termíny vypsane pro plnění studijních povinností. V čase, který si může volit, studuje předepsané materiály, prostřednictvím chat-roomu „elektronicky“ diskutuje s ostatními účastníky (reaguje na jejich příspěvky) i s tutorem, tvoří předepsané samostatné práce, které odesílá e-mailem tutorovi, vyplňuje testy atd. Nakonec splní podmínky závěrečné zkoušky, anebo se osobně dostaví k jejímu vykonání.

Jde o

(2 b)

C e l k e m 40 bodů

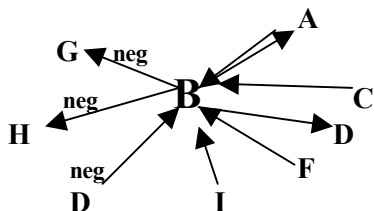
Správnost řešení testových položek zkontrolujte podle dále uvedeného „Klíče“. Promyslete své odpovědi, pokud se neshodovaly s těmi uvedenými v „Klíči“ a odstraňte neznalosti. Test zjišťuje jen **základní** vědomosti i dovednosti a ty by měly být osvojeny důkladně.

KLÍČ SPRÁVNÝCH ODPOVĚDÍ CVIČNÉHO TESTU

1. pravdivá jsou tvrzení a); d); g).

2 a) 3
b) 2; 0

3.



4. problém

5. K = 15

6. inscenace; diskuse; situační metoda

7. ANO

8. a) odborná literatura
b) jiní učitelé a další odborníci;
c) vlastní tvorba čerpající z praxe.

9. a) lepší kontakty učitele se studujícími;
b) lepší kontakty mezi studujícími navzájem.

10. - výsledky skupiny přiznat každému členu;
- ponechat na studujících hodnocení příspěvků jednotlivých členů skupiny;
- v krajních případech hodnotit některé studující jednotlivě;
- při celkovém hodnocení přihlédnout k absencím na seminářích (cvičeních) a hodnocení spojit s výsledky testu a kvalitou ročníkových či seminárních prací.

11. a) Internet (včetně webových stránek a e-mailů);
e) video-kazety;
h) počítačové programy (software);
j) televize;
l) digitální fotoaparát;

d) magnetofonové nahrávky;
g) počítače;
i) rozhlas;
k) videokamera;

12. ANO (1 b)

13. Pravdivé jsou výroky c); e); m); o).

14. Distanční vzdělávání je systém multimediálního řízeného studia, v němž jsou vyučující, konzultanti (tutoři), psychologičtí poradci i administrátoři v průběhu vzdělávání trvale nebo převážně fyzicky oddělení od vzdělávaných.

(V definici musí být DiV označeno jako **multimediální systém řízeného studia** a dále musí být v jejím textu zmínka o tom, že vzdělavatelé a vzdělávaní jsou fyzicky oddělení resp. nejsou v osobním kontaktu. Jinak je třeba považovat definici za neúplnou nebo chybnou).

15. Video-kamera připojená k počítači; mikrofon; reproduktor(y); příslušný (speciální) software.

16. NE

17. Cílová skupina je sociální nebo profesní skupina dospělých, jimž je určen studijní program, kurz, modul nebo studijní pomůcka (text); popř. v marketingovém řízení vytypovaná skupina zájemců o distanční studium určitého programu, kurzu, modulu nebo o studijní pomůcku.

(V definici musí být řečeno, že jde o skupinu osob, již je určen určitý distanční výukový produkt).

18. b) telefon; c) fax; d) e-mail; g) Internet; h) pošta; i) video-konference;

19. Přeskrtnuty musí být aktivity sub: a); c); e); i); a žádné jiné.

20. povinné

21. Studijní centra jsou zařízení distančně vzdělávací instituce technicky i personálně vybavená tak, že umožňují studujícím plnit jejich povinnosti.

(Z definice musí být zřejmé, že jde o patřičně technicky vybavená zařízení vzdělávací instituce pro podporu **distančně studujících**)

22. Jde o **on-line studium**

Maximální počet bodů: 40

Pokud jste chybovali v některé z položek:

1 a) až 1 c), prostudujte znovu oddíly 1.1 až 1.3.

1 d) až 1 i), vraťte se k textu oddílů 2.1 až 2.2.

2 a/nebo 3, seznamte se znovu s obsahem oddílu 2.3.

4 a/nebo 5, věnujte pozornost oddílu 2.4.

6 až 8, přečtěte si raději ještě jednou oddíly 2.5 až 2.7.

9 a/nebo 10, pokuste se najít správné odpovědi v oddíle 2.8.

11, prostudujte znovu odd. 3.1.1;

12, seznamte se podrobněji s obsahem odd.3.1.2;

13, přečtěte si znovu a pozorně text odd. 3.1.2. až 3.1.5;

14 a/nebo 16, vraťte se k oddílu 3.2.1;

15, věnujte ještě jednou pozornost odd. 3.1.4;

17, vraťte se k odd. 3.2.3;

18 až 20, projděte znovu text odd. 3.2.4;

21, osvěžte si informace uvedené v odd. 3.2.6;

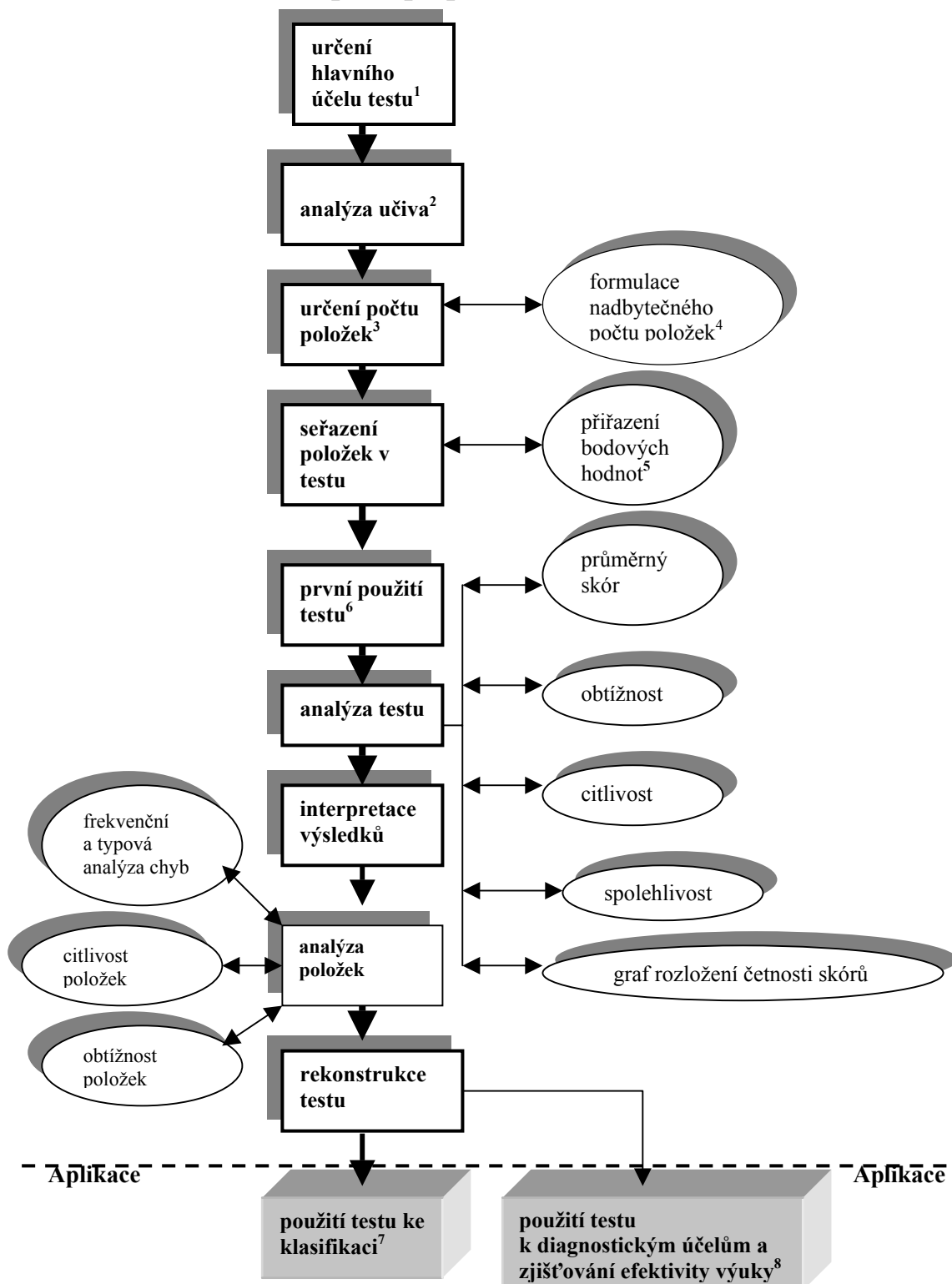
22, správnou odpověď naleznete v odd. 3.2.7.

Pokud jste dosáhli více než 30 bodů (i když to třeba nebylo napoprvé), můžete být s celkovými výsledky spokojeni a považovat prezentované učivo za osvojené. Jinak dbejte předchozích rad k opakovanému prostudování uvedených částí textu.

Poznámka: odkazy se samozřejmě týkají oddílů příslušné učebnice a nikoli tohoto textu.

PŘÍLOHA 2

Schéma postupu při konstrukci didaktického testu



Vysvětlivky

- 1/ pro kterou z diagnostických, kontrolních, procvičovacích, klasifikačních aj. funkcí bude test především určen;
- 2/ podrobné určení cílů testu vzhledem k důležitosti, rozsahu a obtížnosti testovaného učiva;
- 3/ stanovení proporcionality učiva, počtu, obsahu i druhu položek (otevřených i uzavřených);
- 4/ pro možnost konstrukce alespoň dvou ekvivalentních forem testu a také pro případné vyřazení některých;
- 5/ rovnoměrné a/nebo vážené přidělení bodů jednotlivým položkám;
- 6/ určení počtu studujících; odhad času; příprava administrace testu;
- 7/ převod prostých a vážených skóre na klasifikační stupně;
- 8/ zavedení kritérií objektivizace (např. Huberových) pro odhad účinnosti výuky.

**PŘÍLOHA 3****Nejčastější chyby autorů při sestavování a používání didaktických testů**

1. Autoři se příliš zaměřují na kontrolní a klasifikační funkce testů a nevyužívají dalších možností.
2. Podceňují význam počáteční analýzy učiva a začínají tvorbu testů přímo formulací položek.
3. Při tvorbě položek preferují zjišťování výsledků pamětního učení a zanedbávají zkoumání dovedností. Výsledkem pak jsou testy jen slovních - formálních vědomostí.
4. S tím souvisí pouze výjimečné zařazování položek formulovaných problémově, tj. položek vyžadujících větší úsilí a samostatnost při řešení (např. konstruktivní úlohy).
5. Formulace testových zadání jsou nepřesné (nejasná, složitá), odpovědi se následně neshodují s očekáváním autora (ačkoli mohou být vzhledem k zadání správné) a hodnocení výsledku není pak dostatečně „spravedlivé“.
6. V textu otázky je použito dvojího záporu nebo řešení zadaného úkolu je složité - vyžaduje příliš mnoho myšlenkových operací.
7. Není přesně určen způsob, jak správné odpovědi označovat (podtržením, přeškrtnutím, zatržením, zakroužkováním apod.).
8. Text zadání poskytuje bezděčnou nápovědu pro správnou odpověď nebo naopak nechtěně navádí k tvorbě (volbě) chybné odpovědi.
9. Některé odpovědi ve výběru nezaujmou jako možné správné nabídky a jsou testovanými osobami automaticky vyřazovány z úvah.
10. Znění testových položek je stereotypní: není využito více typů (forem) položek, i když by to bylo účelné (např. přiřazovací, rozdělovací, seřazovací odpovědi apod.).
11. Málo se využívá úkolů, jejichž řešení vyžaduje použití tabulek, grafů, diagramů, nomogramů, schémat, map, norem, technických výkresů atd.
12. Splnění úkolu položky je podmíněno znalostí správné odpovědi v předcházející položce, tzn. položky nejsou vzájemně nezávislé.
13. Zanedbává se kritické posouzení návrhu testů odborníky (např. kolegy učiteli s příslušnou aprobací).
14. Neprovádí se analýza výsledků a z ní plynoucí úpravy prvních verzí testů;
15. následkem je nepřesný popis, analýza a interpretace výsledků včetně pouze hrubě odhadnutého převodu na klasifikační stupně.
16. Odhad doby na splnění testu je příliš optimistický; u položek s nejvyššími pořadovými čísly mohou chybět odpovědi nikoli z důvodu neznalosti příslušného učiva, ale proto, že se k jejich řešení část testovaných osob nedostala.
17. Výsledky testů se jednostranně chápou jako záležitost testovaných osob a nikoli ve vztahu k možným nedostatkům výuky popř. studijních materiálů.
18. Přehlíží se možnost zavádět pomocí dosažených testových výsledků kritéria efektivity výuky a zjišťovat tak její kvalitu.



PŘÍLOHA 4



Přehled zavedených veličin a výpočtů jejich hodnot

A_k Opravený (korigovaný) počet správných odpovědí v testu. U výběrových (volených) odpovědí lze upravit konečný počet správných odpovědí v testu s ohledem na možné hádání (náhodnou volbu jedné z nabízených odpovědí). Je-li celkový počet správných odpovědí zvolený testovanou osobou **A_s**, **A_n** celkový počet zvolených nesprávných odpovědí a **y** počet možností předložených k volbě, počítáme

$$ze\ vzorce: \quad A_k = A_s - \frac{A_n}{y-1}$$

f Četnost (frekvence). Počet výskytů určitého jevu. Např. dosáhne-li skóru 12 bodů pět osob ve skupině testovaných, je četnost tohoto výsledku **f = 5**.

G (I až III) **Huberova kritéria**. V závislosti na výsledcích didaktických testů se jimi víceméně kvalifikovaně odhaduje **účinnost výuky**. Vypočítají se ze vzorců:

$$\% \bar{G}(I) = \frac{100 \bar{X}_{vt}}{X_{\max}}; \quad \% \bar{G}(II) = \frac{100(\bar{X}_{vt} - \bar{X}_{pre})}{X_{\max}}; \quad \% \bar{G}(III) = \frac{100(\bar{X}_{vt} - \bar{X}_{pre})}{X_{\max} - \bar{X}_{pre}}$$

kde **X_{pre}** je průměrný výkon v tzv. pretestu (testu, který zjišťuje znalosti **před** zahájením výuky);
X_{vt} je průměrný výkon ve výstupním testu (nasazeném **po** ukončení výuky nebo její části) a
X_{max} je maximálně dosažitelný výkon v testu.

i Index, kterým se zjišťuje **citlivost testu**, daná poměrem mezi směrodatnou odchylkou **s_x** prostého

(nebo váženého) skóru a průměrným skórem: $i_t = \frac{s_x}{\bar{X}}$

Citlivost **položek** je definována odlišně a zjišťuje se jí, jak účinně příslušná položka rozlišuje mezi

lepšími a horšími výkony: $i_p = \frac{2(X_l - X_h)}{n}$

kde **X_l** je počet správných odpovědí na příslušnou testovací položku dosažený těmi z testovaných, kteří dosáhli v celém testu lepších než průměrných výsledků, **X_h** je počet správných odpovědí na tutéž položku dosažený těmi z testovaných, kteří dosáhli v celém testu **horších** než průměrných výsledků a **n** je počet zahrnutých osob. U dobře setavených testů má dosahovat **i_t** a **i_p** hodnoty větší než **0,25**.

Položky, které správně vyřeší většina studujících i položky, které naopak většina (či všichni) nevyřeší, jsou necitlivé a obvykle se vyřazují (pokud test neslouží zvláštním účelům).

Poznámka: Spolehlivějším ukazatelem citlivosti položky je tzv. **bodově biseriální koeficient**:

$$R_{bb} = \sqrt{P'(1-P')(\bar{X}_s - \bar{X}_{ch})} / s_x$$

kde **P' = I_o/100** (index obtížnosti položky dělen 100), **s_x** je směrodatná odchylka prostých skóre dosažených v testu; **X_s** a **X_{ch}** jsou průměry skóre těch osob, které odpověděly na danou položku správně resp. chybně.

Jak je zřejmé, je tento výpočet pracný a užívá se jen u automatizovaného zpracování výsledků.

I_o Index obtížnosti položky nebo testu. Udává procent osob, které vyřešily položku nebo celý test správně. U **testu** se vypočítá jako stonásobek průměrného skóre dělený maximálně dosažitelným skórem:

$$I_o = \frac{100 \bar{X}}{X_{\max}}$$

Pro zjišťování indexu obtížnosti u **položky** se stonásobek dosaženého počtu správných řešení položky dělí počtem testovaných osob **n**.

Index obtížnosti jak u položky, tak u testu by se měl pohybovat kolem hodnoty 50 % - 60 %. Blíží-li se hodnotě 90 a vyšší, jsou test nebo položka příliš snadné. Blíží-li se hodnotě 25 či méně, jsou naopak test či položka příliš obtížné.

m Medián. Skutečně dosažená střední hodnota. Zjišťuje se tak, že se datový soubor seřazený podle vzrůstajících (nebo klesajících) hodnot rozdělí na dvě stejné početné části a jako medián se určí hodnota, které předchází i po níž následuje stejný počet údajů. U souborů s lichým **n** je to přímo prostřední údaj. U souborů se sudým **n** se medián určí jako součet dvou prostředních údajů dělený 2.

n Počet testovaných osob nebo údajů (číselných hodnot) v datovém souboru.

p Počet položek v testu.

r Index spolehlivosti (reliability) testu. V podstatě vyjadřuje to, s jakou přesností můžeme očekávat dosažení stejného (nebo velmi blízkého) výsledku při opakovaném testování tímto testem resp. jeho ekvivalentní formou. Spolehlivost závisí na počtu položek **p**, rozptylu skóre s_x^2 a střední hodnotě dosaženého výsledku (mediánu **m**),

$$\text{jak je vidět z následujícího vzorečku: } r = \frac{m(p - m)}{p \cdot s_x^2}$$

Hodnota **r** se pohybuje v rozmezí 0 až 1 a u standardizovaných testů má být vyšší (až mnohem vyšší) než **0,6**. U nestandardizovaných testů se má hodnotě 0,6 alespoň co nejvíce přibližovat. Je-li nižší, mělo by se uvažovat o **rekonstrukci** testu, spočívající především ve zvýšení počtu položek. Potřebné zvýšení zjistíme výpočtem ze vzorečku pro faktor **z**, kterým se dosavadní počet položek vynásobí:

$$z = \frac{r_p(1 - r_d)}{r_d(1 - r_p)}$$

kde **r_p** je požadovaný a **r_d** skutečně dosažený koeficient reliability. (Hodnoty **z** můžeme také odečíst z tabulky uvedené na konci této přílohy.)

s_x Směrodatná odchylka. Zachycuje „vzdálenost“ zjištěných hodnot od aritmetického průměru. Patří k důležitým charakteristikám souborů dat, o nichž toho aritmetický průměr řekne jen málo, protože sám o sobě nevypovídá nic o tom, z jak rozdílných údajů byl počítán. Odchylky jdou samozřejmě oběma směry a ty menší než aritmetický průměr nesou záporné znaménko. Chceme-li zdůraznit extrémní hodnoty (které souborné údaje nejvíce zkreslují) a zároveň se zbavit záporných znamének, zjišťujeme druhou mocninu všech odchylek od průměru.

$$s_x^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

Veličina s_x^2 se nazývá **rozptyl** a můžeme si ji pro názornost představit jako tu část **plochy terče**, na níž jsou po střelbě rozptýleny **všechny** zásahy. Odmocníme-li její hodnotu, dostaneme **s_x**, a ta pak udává to, co jsme řekli na začátku hesla - totiž vzdálenost „výstřelů“ od středu terče (v našem případě od hodnoty aritmetického průměru). Násobíme-li ji +2,5 a -2,5, získáme i ty hodnoty, které jsou od aritmetického průměru nejvzdálenější.

S_{EM} hrubý odhad chyby dosaženého prostého skóre v testu. Zjišťuje se ze vztahu: $S_{EM} = 0,43\sqrt{p}$, kde **p** je počet položek v testu.

X_i Prostý nebo vážený skóre. Hodnota udává součet bodů, které osoba **i** získala za správná řešení jednotlivých testových položek. Za **prostý** se označuje skóre v tom případě, že za každé správné řešení byl přidělován 1 bod. Za **vážený** se označuje skóre v tom případě, že různým položkám byl podle jejich obtížnosti přidělen (odhadem) různý počet bodů.

X_{max} Maximálně dosažitelný skóre v testu (ať už prostý nebo vážený). Tj. počet bodů, které by získala osoba, řešící všechny položky v testu zcela správně.

$\bar{X}$ Průměrný skóre (prostý nebo vážený). Vypočítá se tak, že součet bodů získaný všemi testovanými osobami $\sum X_i$ se vydělí počtem těchto osob **n**: $\bar{X} = \sum X_i / n$

z Multiplikativní koeficient k výpočtu dodatečného počtu položek, jimiž lze zvýšit spolehlivost testu - viz index **r**. Připojená tabulka ukazuje, jakým číslem je třeba násobit počet položek testu, jehož zjištěný **r_d**, chceme zvýšit na **r_p**.

Tabulka zaokrouhlených hodnot koeficientu z

$r_d \backslash r_p$	0,60	0,65	0,70	0,75	0,80
0,45	1,8	2,3	3	3,7	5
0,50	1,5	2	2,3	3	4
0,55	1,2	1,5	2	2,5	3,3
0,60	1	1,2	1,6	2	2,2
0,65		1	1,3	1,6	2,2
0,70			1	1,3	1,7
0,75				1	1,3

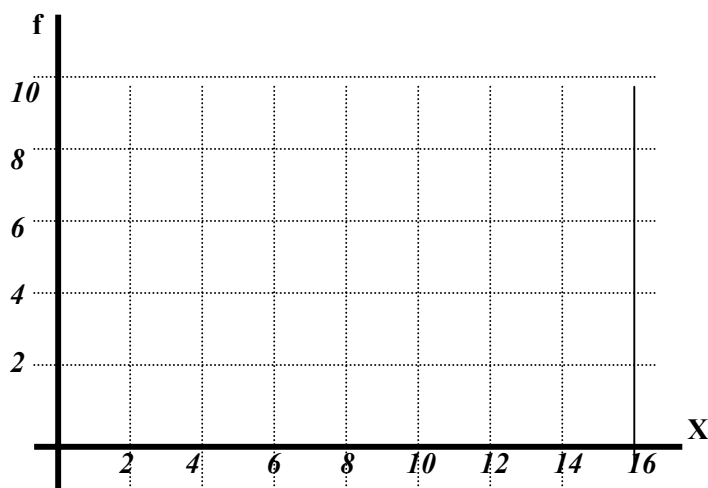
Šedý podtisk vymezuje jen ty hodnoty koeficientu **z**, které má smysl použít pro zlepšení spolehlivosti testu zvýšením počtu položek. Hodnoty uvedené v políčkách bez podtisku (větší než 1) jsou spíše teoretické.



PŘÍLOHA 5

**PŘÍKLADY**

1. Test má 35 položek, medián dosažených bodů je 21 při rozptylu 16 (směrodatná odchylka $s = 4$). Vyjádřete koeficient spolehlivosti r tohoto testu.
2. Na jaký počet položek by bylo třeba zvětšit rozsah testu, kdybychom chtěli dosáhnout spolehlivosti $r = 0,65$?
3. Jaký je index obtížnosti testu uvedeného pod zadáním 1., jestliže průměr testového skóru dosáhl hodnoty $\bar{X} = 19,8$ bodů a v testu je za každou správnou odpověď přidělován 1 b?
Jak lze zjištěnou hodnotu komentovat?
4. Určete ještě citlivost testu (vyjádřenou indexem i_t), kterým jsme se v předchozích výpočtech zabývali.
Vyjádřete se ke zjištěné hodnotě.
5. Ze všech testových položek zodpověděla testovaná osoba 11 a z toho 8 správně. Volilo se vždy ze 4 výběrových odpovědí. Kolik správných odpovědí započtete jako výsledek testu po opravě na možnost náhodné volby správné odpovědi?
6. Uskutečňte analýzu testu o 15 položkách s tvořenými i výběrovými odpověďmi, který zpracovalo 22 studujících s těmito celkovými výsledky:
3; 4; 5; 6; 7; 7; 8; 8; 8; 9; 9; 9; 10; 10; 10; 11; 11; 11; 12; 12; 13; 13.
Zjistěte průměrný skór, jeho směrodatnou odchylku; index obtížnosti i citlivost testu a spolehlivost; nakonec slovně popište dosažené parametry testu.
7. Sestrojte graf (polygon) rozložení četností dosažených skóre (přitom slučujte vždy dva sousední skóre do jednoho intervalu); slovně charakterizujte výslednou křivku.

**Komentář:**



Věcný rejstřík



alternativa 10; 11; 16; 17; 21; 23; 24; 28; 30; 32
- chybná 21

analýza - dat 27
- chyb 33; 34; 35; 47;
- položek 47
- řádková 34
- statistická 33
- testu 23; 47
- učiva 9; 47
- výsledků 27; 48

četnost 31; 49
- intervalová 24; 25
- chyb 37
- rozložení 26
- skóre 38; 47
- výskytu 31; 36

efekt - skupinový 36
- učení 36; 37

efektivita - kritéria 36; 40
- testu 27
- učení 36
- výuky 3; 33; 36; 47; 48

forma - (ekvivalentní) testu 33; 36
- prezentace 4
- výuky 4

frekvence 49
- chyb 27

graf 34; 39; 48
- rozložení četností 47

chyba - hromadná 34
- hrubá 21; 22; 34
- charakteristická 21
- konstantní 34; 35; 37
- logická 33; 38
- náhodná 34
- nejčastější 48
- nesmyslná 38
- nezjistitelná 15
- numerická 8
- počet 33
- typická 20; 27; 30
- učitele 30
- variabilní 34; 35; 37
- z nepozornosti 34
- z neznalosti 34

index - citlivosti testu 7; 49
položky 31; 35; 49
- obtížnosti testu 7; 23; 24; 26; 31; 35; 49
položky 23; 24; 35; 49
- spolehlivosti testu 24; 31; 38; 50; 51

instruktáž 13; 27; 38

interval - časový 12; 27; 30
- hodnot 6; 24; 25;
- rozhraní 23; 26

klasifikace 26; 27; 32; 47; 48

- chyb 34
- intervaly 26
- studujících 2
- stupnice 23; 26; 27; 47; 48
- účel 3
- závěrečná 26

kritérium - hodnocení 4; 8; 16
- Huberovo 36; 37; 39; 40; 47; 49
- platnosti 6
- účinnosti výuky 36; 37; 48

medián 6; 7; 24; 28; 29; 49; 50

nápověda 10; 15; 16; 48

odchylka směrodatná 7; 23; 24; 25; 26; 28; 29;
31; 49; 50

odpověď - alternativní 24; 30; 32
- doplňovací 9
- chybná (nesprávná) 11; 12; 20; 21; 22;
27; 29; 30; 34; 48; 49;
- chybějící (vynechaná) 30; 34; 35
- nabízená 20; 21; 49
- nepřesná 15; 19; 34; 38
- neúplná 8; 22; 34
- otevřená 14
- otevřená rozsáhlá 9; 15; 22; 30
- přiřazovací 9; 18; 22; 48
- rozdělovací 9; 18; 22; 48
- seřazovací 9; 18; 22; 48
- správná 7; 10; 11; 12; 15; 19; 20; 21;
22; 23; 24; 27; 28; 29; 30; 31; 32; 33;
34; 35; 48; 49
- tvořená 9; 10; 17; 19; 22; 30
- volitelná (volená) 21; 48
- výběrová 8; 9; 10; 11; 17; 19; 20; 21;
23; 24; 27; 29; 30; 48
- výběrová s opravou na hádání 10; 11;
12; 29; 49
- uhodnutá 32
- univerzální 11; 21
- uzavřená

plán časový 33; 48

položka testová 3; 6; 7; 9; 10; 11; 12; 14; 23; 25;
30; 35; 44; 48; 50

- analýza 37; 47
- citlivost 47; 49
- druhy (typy) 9; 14; 47
- obsah 23
- obtížnost 6; 8; 30; 31; 35
- počet 6; 7; 9; 10; 12; 25; 30; 31; 32;
38; 47; 50
- pořadí 29
- tvorba 48
- ukázky 16
- uspořádání 12; 25

polygon 25; 36; 38

průměr aritmetický 36; 50

rozptyl 6; 7; 25; 29; 33; 50

- skóre 25

- výsledků 27; 29; 33; 50

skór - bodový 7; 8; 25; 26; 28; 33; 36

- celkový 7; 35

- dosažený 7; 25; 26; 27

- dosažitelný 26; 36; 37; 49; 50

- maximální 7; 23; 50

- nadprůměrný 35

- podprůměrný 35

- prostý 11; 22; 23; 24; 30; 31; 34; 35; 47; 49; 50

- průměrný 23; 24; 25; 26; 31; 33; 38; 49; 50

- rozložení 25

- rozptyl 25

- testový 23; 26

- vážený 22; 23; 30; 47; 49; 50

test - analýza 23; 24; 26; 35

- cíle 9; 23; 47

- citlivost 7; 23; 24; 25; 26; 29; 31; 49

- didaktický 3; 4; 6; 9; 12; 27; 47; 49

- dílčí 7; 12

- efektivita (efektivnost) 27; 33

- forma 9; 29; 30; 47

- funkce 4; 6; 29; 47; 48

- hodnocení 15; 19; 22

- interpretace (výsledků) 23

- klasifikační 9; 21

- konstrukce (tvorba) 9; 10; 12; 14; 48

- kvalita 6

- naučení 7

- nestandardizovaný 3; 23; 25; 50

- objektivnost 8

- obsah 8; 27

- obtížnost 7; 23; 25; 35; 38; 47; 49

- osobnostní 3

- parametry 24; 26; 36

- platnost 6; 12; 23

- posttest 33

- použití 35

- pretest 33; 37; 49

- procvičovací (cvičný) 4; 29; 33; 41; 45

- průběžný 12; 36

- retest 33

- retenční 36

- revize (úpravy) 34; 38; 50

- rozsah 7

- schopností 3

- spolehlivost 6; 7; 23; 24; 25; 29; 30; 31; 38; 50

- standardizovaný 3; 23; 25; 50

- účel 7; 9; 26

- vstupní 33; 36

- výhody (přednosti) 5; 29

- výsledek 10; 11; 23; 30; 31; 37; 48; 49

- výstupní 4; 33; 36; 37; 49

- základního učiva (vědomostí) 7; 27; 32; 44

vazba - zpětná 29

vnější 4; 42

vnitřní 4; 42

volba - odpovědi 17; 30; 49

- chybná 12; 21; 48

- náhodná 10; 11; 28; 30; 49

- nejvhodnější 29

- správná 12; 20; 21; 28

© Jiří Průcha, Praha 2003

80 normovaných stran; 4 AA