

DIDAKTICKÉ TESTOVÁNÍ

Jana Junková

ÚVODEM	3
VÝHODY POČÍTAČOVÉHO TESTOVÁNÍ A JEHO ORGANIZACE	3
VYUČOVACÍ PROCES A ZPĚTNÁ VAZBA	4
VYMEZENÍ POJMU „ŠKOLNÍ HODNOCENÍ“	5
PRINCIP OBJEKTIVITY A SUBJEKTIVITY VE ŠKOLNÍM HODNOCENÍ.....	5
ÚSTNÍ ZKOUŠKA A TESTOVÁNÍ.....	6
DIDAKTICKÝ TEST	6
DRUHY DIDAKTICKÝCH TESTŮ.....	6
VLASTNOSTI DOBRÉHO DIDAKTICKÉHO TESTU	9
TVORBA DIDAKTICKÉHO TESTU.....	10
PLÁNOVÁNÍ DIDAKTICKÉHO TESTU	11
TECHNIKA SPECIFIKAČNÍ TABULKY	11
TECHNIKA SEZNAMU VÝUKOVÝCH CÍLŮ	13
KONSTRUKCE DIDAKTICKÉHO TESTU	14
NÁVRH TESTOVÝCH ÚLOH.....	14
NÁVRH PROTOTYPU DIDAKTICKÉHO TESTU	23
OVĚŘOVÁNÍ A OPTIMALIZACE DIDAKTICKÉHO TESTU.....	23
ANALÝZA VLASTNOSTÍ TESTOVÝCH ÚLOH	24
ÚPRAVA VYTVOŘENÉHO DIDAKTICKÉHO TESTU	27
TVORBA STANDARDIZOVANÉHO DIDAKTICKÉHO TESTU	28
URČENÍ RELIABILITY DIDAKTICKÉHO TESTU	28
STANDARDIZACE DIDAKTICKÉHO TESTU	30
VYTVOŘENÍ KLASIFIKAČNÍHO STANDARDU PRO DIDAKTICKÝ TEST.....	31
POUŽÍVÁNÍ DIDAKTICKÝCH TESTŮ VE ŠKOLNÍ PRAXI	32
DIAGNOSTICKÝ ROZBOR VÝSLEDKŮ TESTU	32
POSOUZENÍ CELKOVÝCH VÝSLEDKŮ TŘÍDY	33
KLASIFIKACE VÝSLEDKU TESTU.....	33
ZÁVĚR.....	35

ÚVODEM

Dnešní prudký rozvoj v oblasti informačních technologií se musí také nutně odrazit i ve vzdělávání. Zatím vývoj směřoval k získání základních teoretických informací a praktických dovedností v oblasti práce s aplikacemi a programování. Proto logicky dalším možným krokem je využití počítačů k testování znalostí žáků.

Tato práce by měla nastínit správný postup při vytváření a ověřování testů, doporučí vhodný nástroj na počítačové testování, případně využití písemného testování v jakémkoliv vyučovacím předmětu. Dále práce poukáže na výhody testování a jeho organizační možnosti na škole.

Výstupem pro ovládání počítačové aplikace pro tvorbu a spouštění testů je Manuál II, výstupem pro tvorbu a ověřování testů je příručka Didaktické testování, jejíž součástí je i návrh organizační struktury pro samotné počítačové testování na školách.

VÝHODY POČÍTAČOVÉHO TESTOVÁNÍ A JEHO ORGANIZACE

Počítačové technologie zasahují téměř do všech oblastí vzdělávání žáků. Žáci se je učí ovládat, využívat k získávání informací, s jejich pomocí si ověřují své znalosti a dovednosti. Na řadě vyšších odborných a vysokých škol již probíhá přijímací řízení a následně pak výuka pomocí výpočetní techniky, totéž se v brzké době bude týkat přijímacích testů na střední školy. Proto skórované testování se brzy stane běžnou součástí vyučování na školách.

Jako vhodné řešení se jeví například program doSystem-EduBase(www.dosli.cz). Tato aplikace je určena pro přípravu výukových materiálů a jejich využití ve výuce. Výukové materiály mohou obsahovat výukové texty (učebnice), příklady včetně jejich řešení a testy. Tyto testy lze vytvářet pomocí otevřených a uzavřených otázek typu klasické (dichotomické a s výběrem odpovědí), obrázkové, přiřazovací, uspořádací a doplňovací. Ve srovnání s testy v prostředí Moodle má program doSystem-EduBase možnost tisku, poskytuje jednoduchou statistiku dosažených testových výsledků, umožňuje archivaci výsledků.

Při samotném testování je výhodou pro žáky vždy dobře čitelné zadání testů, možnost neměnné časové volby délky testu, objektivní, nestranné a rychlé vyhodnocení testu strojem s možností zpětné kontroly. Žáci si tedy mohou prolistovat seznam svých odpovědí včetně správného řešení. Po ukončení testu znají svůj dosažený celkový počet bodů.

Výhodou testovacího programu pro učitele je jeho snadné ovládání při vytváření výukových materiálů. Učitel již nebude potřebovat žádný jiný program, vše, co potřebuje, je v něm již zahrnuto (tzn. editor textů, obrázků, tabulek, vzorců a možnost vkládání různých symbolů a odkazů). S ukončením testu učitel okamžitě získává vyhodnocené výsledky žáků ve formě celkově dosaženého počtu bodů, případně procentuálního vyhodnocení. Může si výsledky řadit, filtrovat, archivovat a vytvářet tiskové výstupy. Učitel má možnost procházet testy žáků a tak zjistit příčiny chyb. Následně pak lze test upravit, některé otázky odstranit, jiné přidat.

DoSystem-EduBase vyhoví různým nárokům na provoz sítě. Je schopen pracovat i jako jednouživatelská verze.

K testování žáků mohou posloužit libovolné učebny vybavené počítači s testovacím programem. Většinou jsou používány k výuce informačních technologií a jejich kapacita je maximálně sedmáct počítačů, tedy pro stejný počet žáků. Jak zajistit testování celé třídy i v jiných předmětech během jedné vyučovací hodiny?

Teorie testů stanoví jejich optimální délku do dvaceti minut, což se dá ideálně zužitkovat. Rozdělíme žáky na poloviny. Jedna část bude testována, druhá část dostane například samostatnou práci. Dozor lze zajistit prostřednictvím kolegy, který právě nevyučuje. K zajištění učebny lze využít rezervační systém (realizovaný formou vývěsky nebo v elektronické podobě), přesuny tříd v učebnách, případně období, kdy učebny již nejsou plně vytíženy, tedy v době maturit a následně po nich.

Pokud není možno provádět počítačové testování, lze je nahradit testováním písemným. Program DoSystem-EduBase dokáže vytisknout testovací listy v různých variantách, kdy test obsahuje stále tytéž otázky, ale jejich pořadí mění a mění také pořadí odpovědí na otázku. Takže každá varianta má jiné pořadí otázek i odpovědí a čím více variant, tím získáme objektivnější výsledky. Dále program umožňuje vytisknout pro každou variantu karty pro odpovědi žáka a karty správné volby odpovědi pro zkoušejícího.

Odborně sestavený a správně používaný didaktický test může být velmi užitečným prostředkem k získávání objektivních informací o znalostech a dovednostech žáka. Pro potřeby škol se zatím nevydávají vždy ověřené (standardizované) didaktické testy a pokud ano, pak za nemalý finanční obnos.

Tvorba kvalitního didaktického testu je činnost velmi náročná na čas i odbornost autora. Někteří učitelé dají vždy přednost přípravě vlastního testu. A proto každý učitel by měl být také obeznámen alespoň s hlavními principy používání, hodnocení a interpretace výsledků didaktických testů.

Neocenitelnou předlohou této práce je publikace Miroslava Chrásky *Didaktické testy*, z níž pochází odborná teorie testů.

VYUČOVACÍ PROCES A ZPĚTNÁ VAZBA

Vyučovací proces je možno při určitém zjednodušení chápat jako řízený proces, ve kterém lze rozlišit dvě základní funkce:

- sdělování nových poznatků
 - kontrola množství a kvality osvojených vědomostí a dovedností.
- Tyto dvě fáze vyučovacího procesu tvoří jednotný celek, který nelze násilně rozdělovat. První fázi se v dnešní době vyčleňuje velký prostor, ať již při přípravě budoucích učitelů, nebo v literatuře. Druhá fáze sice samotnými učiteli opomíjena není a také nemůže být, protože by tu nebyla zpětná vazba, ale chybí dostatek publikací a odborné literatury jako podpůrný prostředek.

Zpětná vazba se realizuje pohovorem s žákem, zkouškou, testem, řešením problémových úkolů spojováním s praxí, ... Toto vše pak slouží jako podklad pro hodnocení žáka.

Hodnocení je nedílnou součástí každé lidské činnosti, spolu s procesem rozhodování ovlivňuje každé jednání člověka. Je přirozenou součástí každé výchovně-vzdělávací činnosti. a tudíž je přítomno i v činnosti učitele i v činnostech žáků.

Pro řízení, ovlivňování a usměrňování učebních činností žáků používá učitel celou řadu prostředků. Hodnocení žáků, jejich výkonů, činností, chování, je jeden ze základních, velmi účinných prostředků učitele, jak řídit a usměrňovat složité a pro žáky značně náročné učební činnosti.

VYMEZENÍ POJMU „ŠKOLNÍ HODNOCENÍ“

V odborné pedagogické literatuře můžeme nalézt různá vysvětlení pojmu školní hodnocení. Např.:

- Slavík (1999, s. 23-24) rozumí školním hodnocením „všechny hodnotící procesy a jejich projevy, které bezprostředně ovlivňují školní výuku nebo o ní vypovídají“.
- J. Velikanič (1973, s. 156-157) chápe hodnocení jako „proces stálého poznávání a posuzování žáka, jeho vědomostní úrovně, pracovní a učební činnosti, jeho výsledků“. Hodnocení má vyjádřit ocenění žákovy práce, nebo naznačit cestu, jak nedostatky napravit.
- J. Skalková (1971, s. 95) hodnocení chápe jako „zaujímání a vyjadřování kladného nebo záporného stanoviska k různým činnostem a výkonům žáků při vyučování, které může mít v praxi nejrůznější formy: od souhlasného nebo nesouhlasného pokývnutí hlavou, přísného pohledu, tónu hlasu, kladné či negativní poznámky, zájmu o osobnost žáka, pochvaly či napomenutí, odměny či trestu až po známku, případně podrobnější analýzu výkonu včetně závěrečného hodnotícího soudu aj.“
- M. Pasch a kol. (1998, s. 104) hovoří o hodnocení jako „systematickém procesu, který vede k určení kvalit a výkonů vykazovaných žákem nebo skupinou žáků, je to činnost systematická, tj. činnost připravená, organizovaná a opakovaně prováděná, jejíž výsledky jsou podrobovány revizím či opravám“.

Hodnocení žáků ve vyučování je také jediné hodnocení v životě člověka, které je ve své podstatě systematické (Pasch, 1998, s. 104). Všechna ostatní hodnocení, se kterými se člověk ve svém životě zákonitě setkává mají charakter až příliš všeobecný a značně náhodný. Také hodnocení, která sám jedinec provádí v každodenních běžných situacích, nemusí mít systematický charakter.

Systematičnost hodnocení žáků ve výuce lze spatřovat v tom, že tuto činnost učitel připravuje, organizuje, provádí ji pravidelně, v určitých časových intervalech a výsledky jsou porovnávány se zvolenými normami. A právě touto kvalitou (systematičností, permanentností) má školní hodnocení nesmírně silný vliv na povahu vyučování vůbec, na intenzitu učebních činností žáků i na kvalitu a schopnost sebehodnocení žáků.

Školní hodnocení tedy poskytuje informace o tom, jak úspěšně probíhá výuka a jaké jsou její výsledky. Je v podstatě zpětnou vazbou, která vypovídá, zda práce ve škole dosahuje předpokládaných cílů.

Základní pravidlo školního hodnocení

Žáci by měli vždy vědět, co se bude hodnotit, jaká je tedy k dané konkrétní učební látce formulována cílová norma. zda stačí si zapamatovat, nebo je nutné pochopit souvislosti, nebo je nezbytné zvládnout aplikaci na nové situace atd. Tedy to, co je standard, co je případně nadstandard a co je bráno ve vztahu k dané cílové normě jako nedostatečný výkon.

PRINCIP OBJEKTIVITY A SUBJEKTIVITY VE ŠKOLNÍM HODNOCENÍ

Otázky subjektivity a objektivit ve školním hodnocení vždy vyvolávaly zájem žáků, učitelů i rodičů. S hodnocením žáků byly a jsou stále spojeny pochybnosti, zda je toto hodnocení pravdivé, zda vyjadřuje to, co vyjadřovat má.

Hodnocení je vždy významným projevem úrovně racionality hodnotitele a zároveň je i projevem jeho subjektivní účasti na procesech, které jsou předmětem jeho hodnotících aktivit. Zvláště to platí ve školních podmínkách, kdy učitel je ve většině případů tím, kdo

žákům předává vědomosti, rozvíjí dovednosti, návyky a formuje jejich zájmy a postoje. Zároveň je však hlavním hodnotitelem výkonů u žáků.

Do školního hodnocení se promítají osobnostní faktory hodnotitele i hodnoceného žáka. V hodnocení se projevuje nejen kvalita hodnocené činnosti, ale i kvalita hodnotitele, tzn. že jakékoli hodnocení bude nutně poznamenáno subjektivními prvky.

Proces kvalitního hodnocení je značně náročnou činností kladoucí vysoké nároky na osobnost učitele. Jednou ze součástí hodnocení jsou také výsledky školního měření, což lze chápat jako kvantitativní popis výkonu žáka. Měření je objektivní, neosobní a přesně definuje pravidla hodnocení. Prostředkem takového měření jsou didaktické testy. Objektivnost testu by měla být zaručena omezením, resp. úplným vyloučením vlivu osobnosti učitele ve funkci zadavatele a interpreta testu. Hodnocení testu je zaměřeno především na přísné změření výkonu a tak dává každému rovné šance. Testy se staly potřebným a využívaným nástrojem zejména v oblasti pedagogicko-psychologické diagnostiky žáka nebo jako prostředek porovnávání školních výkonů v rámci různých evaluačních programů.

Ovšem pozor, učitel *nikdy* nesmí založit hodnocení jen na výsledcích školního měření, protože pak by neplnilo všechny své funkce, které plnit má, zejména funkci výchovnou, regulační a motivační. ale může je použít jako objektivně měřitelnou součást hodnocení.

ÚSTNÍ ZKOUŠKA A TESTOVÁNÍ

K hodnocení žáků může učitel použít různých metod, přesto je současný systém hodnocení a klasifikace žáků založen převážně na individuálním ústním zkoušení. Dobře připravená a provedená ústní zkouška má velký význam zvláště pro rozvoj vyjadřování a myšlení žáka. Přes tyto klady má ústní zkouška i řadu nedostatků. Vede spíše k memorování než k řešení problémů a k rozvoji myšlení, je časově náročná, neexistují pevné normy pro hodnocení (každý učitel klasifikuje jinak)

Tradiční ústní zkouška však dnes již nemůže sama o sobě objektivní a spravedlivé hodnocení žáka stačit. Vhodným doplňkem ústní zkoušky může být i kvalitní *didaktický test*.

DIDAKTICKÝ TEST

Didaktický test se orientuje na objektivní zjišťování úrovně zvládnutí učiva u určité skupiny osob. Od běžné zkoušky se didaktický test ovšem liší zejména tím, že je navrhován, ověřován, hodnocen a interpretován podle určitých, předem stanovených pravidel. Stručná a výstižná je definice didaktického testu podle P. Byčkovského(1982): „didaktický test je nástroj systematického zjišťování (měření) výsledků výuky“.

DRUHY DIDAKTICKÝCH TESTŮ

V pedagogické praxi se lze setkat s didaktickými testy různé kvality a různého druhu. Jednotlivé druhy didaktických testů mají své specifické vlastnosti a liší se tím, jaké informace pomocí nich získáváme. Při popisu jednotlivých typů didaktických testů budeme postupovat podle klasifikace, kterou navrhl P. Byčkovský (1982).

Tabulka 1: Druhy didaktických testů

KLASIFIKAČNÍ HLEDISKO	DRUHY TESTŮ		
měřená charakteristika výkonu	rychlosti	úrovně	
dokonalost přípravy testu a jeho příslušenství	standardizované	nestandardizované	
povaha činnosti testovaného	kognitivní	psychomotorické	
míra specifičnosti učení zjišťovaného testem	výsledků výuky	studijních předpokladů	
interpretace výkonu	rozlišující (relativního výkonu)	ověřující (absolutního výkonu)	
časové zařazení do výuky	vstupní	průběžné (formativní)	výstupní (sumativní)
tematický rozsah	monotematické	polytematické (souhrnné)	
míra objektivity skórování	objektivně skórovatelné	kvaziobj. skórovatelné	subjektivně skórovatelné

Testy rychlosti

U těchto testů se zjišťuje, jakou rychlostí je žák schopen řešit určitý typ testových úloh. Testy rychlosti mají pevně stanovený časový limit pro řešení a obsahují velmi snadné úlohy.

Testy úrovně

Čisté testy úrovně nepoužívají žádné časové omezení (časový limit) a výkon v nich je dán pouze úrovní vědomostí nebo dovedností zkoušeného. Úlohy jsou v testu zpravidla řazeny se vzrůstající obtížností. Z praktických důvodů však bývá nutné s určitým volným limitem pracovat.

Testy standardizované

Standardizovaný didaktický test je připravován profesionálně, je důkladně ověřen, takže jsou známy jeho základní vlastnosti. Součástí příslušenství standardizovaného didaktického testu je testová příručka (manuál), ze které se uživatel dozví o vlastnostech testu, o jeho správném použití atd. Většinou je také k dispozici standard (testová norma) pro hodnocení dosažených výkonů. Tyto testy vydávají většinou specializované instituce.

Nestandardizované didaktické testy

Tyto testy si připravují učitelé sami pro svoji vlastní potřebu. Nebyly u nich realizovány všechny kroky obvyklé při přípravě a ověřování testů standardizovaných. Neproběhlo u nich ověřování na větším vzorku žáků, a nejsou tudíž známy všechny jejich vlastnosti.

Testy kognitivní a testy psychomotorické

Dělení didaktických testů na kognitivní a psychomotorické vychází z dělení lidského učení do tří oblastí podle B. S. Blooma (učení kognitivní, afektivní a psychomotorické). Výsledky učení *afektivního* se didaktickými testy nezjišťují (k tomuto účelu se používají např. dotazníky, různé škály apod.) Pokud didaktický test měří úroveň (kvalitu) poznání u žáků, jde o test *kognitivní*, pokud testem zjišťujeme výsledky psychomotorického učení, hovoříme o testu *psychomotorickém*. Příkladem kognitivních testů jsou např. testy, ve kterých má žák

řešit úlohy z matematiky, překládat text do cizího jazyka atd., příkladem testu psychomotorického je např. test psaní na stroji..

Testy výsledků výuky a testy studijních předpokladů

V běžné pedagogické praxi se doposud téměř výlučně používají didaktické *testy výsledků výuky*, které měří to, co se žáci v dané oblasti naučili.

Testy studijních předpokladů měří úroveň obecnějších charakteristik jedince, které jsou potřebné k dalšímu studiu. Měly by se používat zejména při přijímání žáků ke studiu na vyšší typ školy. Konstrukce testů studijních předpokladů je náročná a vyžaduje vedle pedagogické kvalifikace autora také dobrou kvalifikaci psychologickou.

Testy rozlišující (testy relativního výkonu)

Podle toho, jakým způsobem vysvětlujeme a hodnotíme výkon žáka v testu, můžeme rozlišit tzv. didaktické testy rozlišující (testy relativního výkonu) a didaktické testy ověřující (testy absolutního výkonu).

Hlavní rozdíl mezi těmito dvěma druhy testů spočívá v tom, že u rozlišujících testů se výkon žáka určuje vzhledem k populaci testovaných, zatímco u testů ověřujících se výkon určuje vzhledem ke všem možným úlohám, které určité učivo reprezentují.

V naší pedagogické praxi se zatím používají téměř výlučně testy rozlišující. Základní ideou, o kterou se opírá koncepce rozlišujících didaktických testů, je snaha dosáhnout maximální možné objektivitu a diferencovanosti hodnocení testových výkonů.

Výkon žáka v testu se srovnává s výkony ostatních žáků, v případě standardizovaných rozlišujících testů s výkony celé žákovské populace. Rozlišující didaktické testy jsou tedy konstruovány tak, že umožňují rozhodnout, jaký výkon v testu žák dosáhl vzhledem k celé populaci, k níž patří. Umožňují posoudit, zda určitý konkrétní žák je ve srovnání s ostatními žáky např. „velmi slabý“, „podprůměrný“, „průměrný“ atd.

Testy ověřující (testy absolutního výkonu)

Didaktické testy ověřující jsou často v literatuře označovány také jako *kriteriální* testy nebo CR testy. Úkolem ověřujících testů je prověřit úroveň vědomostí a dovedností žáka v přesně vymezené oblasti (části učiva). Výkon testovaného se vyjadřuje vůči všem úlohám, které reprezentují dané učivo.

U ověřujících testů je kritériem úspěchu předem stanovený stupeň zvládnutí učiva. Ověřující testy požadují u vybraných základních poznatků téměř úplné zvládnutí. Cílem je rozhodnout, zda žák zvládl učivo nebo nikoli. Při konstrukci ověřujících testů je základním problémem výběr učiva, které musí žák bezpečně zvládnout. Toto učivo se potom transformuje do testových úloh. Aby se bezpečně ověřilo zvládnutí určitého učiva, požaduje se, aby každý testovaný jev byl pokryt větším počtem testových úloh.

Testy vstupní, průběžné a výstupní

Vstupní didaktické testy se zadávají na začátku výuky určitého celku učební látky a jejich cílem je postihnout úroveň vědomostí a dovedností, které jsou pro úspěšné zvládnutí daného celku učiva důležité.

Průběžné didaktické testy se zadávají v průběhu výuky a jejich úlohou je poskytovat učitelům zpětnovazební informace potřebné k optimálnímu řízení výuky. Obvykle zkouší jen poměrně

Didaktické testování

malou část učiva a jejich posláním je sledovat, jak žáci probírané učivo přijímají, chápou a jak si je osvojují. Tyto testy neslouží většinou k hodnocení žáků, nýbrž k hodnocení výuky.

Výstupní didaktické testy se zadávají buď na konci výukového období nebo na konci určitého celku a většinou poskytují informace potřebné pro hodnocení žáků. Bývají také označovány jako testy sumativní.

Testy monotematické a polytematické

Monotematické testy zkouší jediné téma učební látky, testy polytematické zkouší učivo několika tematických celků. Testy polytematické jsou proto náročnější z hlediska přípravy i konstrukce.

Testy objektivně skórovatelné

Testy objektivně skórovatelné obsahují úlohy, u nichž lze objektivně rozhodnout, zda byly řešeny správně či nikoli. Výhodou těchto testů je, že skórování může provádět jakákoli osoba (někdy i stroj).

Vzhledem k tomu, že velká většina používaných didaktických testů se vyznačuje možností objektivního skórování, vznikla u části pedagogické veřejnosti nesprávná představa, že test je zkouška, která vždy obsahuje pouze objektivně hodnotitelné úlohy (např. úlohy, kde žák vybírá správnou odpověď, nebo úlohy, kde žák formuluje vlastní, ale velmi stručnou, a tudíž objektivně hodnotitelnou odpověď).

Testy subjektivně skórovatelné

Subjektivně skórovatelné testy (označované často jako esej testy) obsahují úlohy, u nichž není možno stanovit jednoznačná pravidla pro skórování. Mezi subjektivně skórovatelné testové úlohy patří např. tzv. otevřené široké úlohy, ve kterých žák volně odpovídá na položenou otázku uvedením rozsáhlejší odpovědi. Ukazuje se, že není rozumné vyhýbat se používání takových úloh jen proto, že neumožňují objektivní skórování. Otevřené široké úlohy totiž mohou zkoušet daleko komplexnější vědomosti a dovednosti

VLASTNOSTI DOBRÉHO DIDAKTICKÉHO TESTU

Základními vlastnostmi dobrého didaktického testu jsou validita, reliabilita a praktičnost.

Validita (platnost) didaktického testu

Test je validní tehdy, pokud se jím zkouší skutečně to, co má být zkoušeno. Tak například:

- testy studijních výsledků: zde zkoumáme, jak dalece se shoduje obsah testu s cílem a obsahem vyučování; v těchto případech jde především o tzv. obsahovou validitu testu. Obsah úloh testu by měl být reprezentativním vzorkem zkoušeného učiva.
- testy výsledků výuky: tady jsou kritériem validity příslušná kurikula vyučovacích předmětů
- testy studijních předpokladů: by měly mít schopnost předpovídat budoucí úspěšnost v učení (je nutno zvážit, ze kterých vědomostí nebo dovedností je možno usuzovat na budoucí úspěšnost ve studiu).

Reliabilita (spolehlivost a přesnost) didaktického testu

Aby byl didaktický test reliabilní, je třeba, aby byl spolehlivý a přesný. *Spolehlivost* spočívá v tom, že za týchž podmínek by měl poskytovat stejné (velmi podobné) výsledky. Didaktický

test je *přesný* tehdy, jestliže při jeho použití nedochází k velkým chybám měření. Oba požadavky, tj. spolehlivost a přesnost, zahrnujeme pod společný pojem *reliabilita*.

Výsledek didaktického testu u určitého žáka je tvořen dvěma složkami: pevnou složkou (skutečné vědomosti nebo dovednosti) a náhodnou složkou (okamžitá kondice, vnější podmínky atd.). Náhodná složka způsobuje, že při zdánlivě stejných podmínkách se výsledky testování mohou podstatně lišit. U dobrého didaktického testu by se vliv náhodné složky měl uplatňovat co nejméně. O testu poskytujícím výsledky, které jsou jen minimálně dotčeny náhodnými vlivy, můžeme říci, že má vysokou *reliabilitu*.

K přesnému posouzení míry *reliability* didaktického testu slouží *koeficient reliability*. Tento koeficient v praxi nabývá hodnot od 0 (pro případ naprosté nespolehlivosti a nepřesnosti) až po hodnoty blízké 1 (pro případ dokonalé spolehlivosti a přesnosti didaktického testu). Pro individuální pedagogickou diagnostiku se většinou požaduje koeficient *reliability* minimálně 0,80.

Reliabilita testu je závislá na:

- kvalitě testových úloh, z nichž je vytvořen
- počtu úloh; obecně platí, že čím více úloh test obsahuje, tím má větší *reliabilitu*. U testů s počtem úloh 10 nebo méně koeficient *reliability* zpravidla dosahuje maximálně hodnoty kolem 0,60. Tzn. dobrý didaktický test by měl vždy obsahovat dostatečný počet úloh (za spodní hranici lze považovat 10 úloh).

Reliabilita didaktického testu je velmi důležitým ukazatelem jeho technické kvality. Čím nižší je *reliabilita* testu, tím skeptičtěji je nutno výsledky testování posuzovat. Koeficient *reliability* je možné vypočítat několika způsoby. Často se např. užívá Kuderova – Richardsonova vzorce nebo tzv. metody půlení (viz. kap. Tvorba standardizovaného didaktického testu). Stanovení koeficientu *reliability* je nezbytné při ověřování všech standardizovaných didaktických testů. Je však možné je doporučit i v případě testů nestandardizovaných, má-li autor zájem vytvořit test skutečně kvalitní.

Vedle *validity* a *reliability* testu zvažuje učitel při jeho hodnocení i *praktické* výhody, které mu z jeho používání plynou. Dobrý test se vyznačuje tím, že jeho použití je jednoduché, oprava výsledků snadná a rychlá, a proto představuje úsporu času ve srovnání s jinými způsoby zkoušení žáků.

TVORBA DIDAKTICKÉHO TESTU

Vytváření didaktického testu bychom neměli začínat tím, že začneme navrhovat testové úlohy. Tento postup vede k tvorbě nevyváženého didaktického testu, který nepokrývá rovnoměrně celé učivo a zaměřuje se většinou na reprodukci zapamatovaných poznatků.

Následně uváděné činnosti odpovídají svým obsahem i rozsahem konstrukci testu standardizovaného. U testů nestandardizovaných (učitelských) je možno některé kroky vypustit a nebo je uskutečnit ve zjednodušené podobě. Tvorbu didaktického testu lze rozdělit do tří základních etap:

1. plánování testu,
2. konstrukce testu
3. ověřování testu.

PLÁNOVÁNÍ DIDAKTICKÉHO TESTU

Nejdříve se musí určit *účel testu* (např. zjištění výsledků výuky na konci tematického celku nebo na konci pololetí či roku, zjištění, jak žáci probírané učivo přijímají a chápou).

Po ujasnění účelu testování se rámcově vymezuje *obsah testu*.

Příklad 1: Rámcové vymezení obsahu testu

Funkce a logaritmy.

(učivo 20 hodin matematiky, 2. ročník střední odborné školy)

Kompresa a dekomprese dat.

(učivo 3 hodin ICT, 1. ročník střední odborné školy)

Rámcově vymezený obsah testu je třeba upřesnit tak, aby bylo zřejmé, jaký obsah mají jednotlivé úlohy zkoušet, na jakou úroveň osvojování vědomostí se při tom mají zaměřovat, kolik úloh je nutno navrhnout atd. Je známo několik technik, kterými se toto upřesnění může uskutečnit (Byčkovský, 1982). V učitelské praxi přicházejí nejvíce v úvahu dvě z nich - *technika specifikační tabulky a technika seznamu výukových cílů*.

TECHNIKA SPECIFIKAČNÍ TABULKY

Specifikační tabulka (kromě jiného) upřesňuje, jaká úroveň osvojení znalostí má být jednotlivými úlohami testu zkoušena: Dobrý didaktický test by měl ověřovat nejen pamětní osvojování učiva, nýbrž by měl zkoušet i vyšší cílové úrovně, jako je porozumění poznatkům, aplikace poznatků, analýza a syntéza poznatků atd.

Existuje několik tabulek taxonomie výukových cílů (např. Tollingerové taxonomie učebních úloh, Bloomova taxonomie výukových cílů (tab. 2), Niemiarkova taxonomie (tab. 3).

Tabulka 2: Bloomova taxonomie cílů

	Hladina	Popis	Aktivní slovesa
1.	Znalost	žák si dokáže vybavit, reprodukovat nebo rozeznat dříve naučené informace, znovupoznání informace	definovat, doplnit, napsat, reprodukovat, vybavit si, uvést seznam, identifikovat, nazvat, označit, vybrat, seřadit
2.	Porozumění	žák dokáže vlastními slovy vyjádřit dříve naučenou látku, pochopení a schopnost užití znalosti	Dokázat jinak formulovat, ilustrovat, uvést příklad, vyjádřit vlastními slovy, vypočítat, popsat, shrnout
3.	Aplikace	žák dokáže použít dříve naučenou látku, např. pojmy, pravidla nebo generalizace při zpracování nové látky	Aplikovat, demonstrovat, interpretovat, navrhnout, nalézt vybrat, roztrždit, uvést vztah, vyčíslit, vyzkoušet
4.	Analýza	žák dokáže rozčlenit složitou věc, proces na komponenty a vysvětlit, proč je daná sestava vztahů uspořádána daným způsobem nebo jaké příčiny k danému uspořádání vedly	Porovnat, rozhodnout, rozlišit, analyzovat, rozdělit, vysvětlit proč, ukázat jak, nakreslit schéma, načrtnout
5.	Syntéza	žák dokáže z několika jednodušších komponentů vytvořit původní a složitý výrobek aj., vytváří strukturu	Klasifikovat, kombinovat, tvořit, stavět, organizovat, vytvořit originál, komponovat, psát, řešit, předvést, stanovit, předpovědět
6.	Hodnocení	žák umí na základě dříve naučených norem a kritérií stanovit hodnotu nebo cenu složitěho produktu, posouzení s využitím kritérií	Argumentovat, obhájit nebo vyvrátit, rozvíjet a kritizovat, posoudit, zaujmout nebo podpořit stanovisko, ospravedlnit, diskutovat, rozhodnout

Tabulka 3: Niemierkova taxonomie výukových cílů (pro vzdělávací oblast)

<p>A Zapamatování poznatků Této kategorie je dosaženo, jestliže je žák schopen vybavit si určitá fakta (např. termíny, zákony), přičemž je nesmí mezi sebou zaměňovat. <i>Typická aktivní slovesa:</i> definovat, napsat, opakovat, pojmenovat, reprodukovat</p>
<p>B Porozumění poznatkům V tomto případě je již žák schopen zapamatované poznatky předložit v jiné formě než v té, ve které si je zapamatoval, dovede poznatky uspořádat nebo zestručnit. <i>Typická aktivní slovesa:</i> jinak formulovat, ilustrovat, objasnit, odhadnout, přeložit, převést, vyjádřit vlastními slovy</p>
<p>C Používání vědomostí v typových situacích U této kategorie dovede žák použít vědomostí k řešení situací, které ve výuce již byly řešeny. <i>Typická aktivní slovesa:</i> aplikovat, použít, prokázat, řešit, diskutovat, načrtnout, vyzkoušet, registrovat, demonstrovat</p>
<p>D Používání vědomostí v problémových situacích Žák dovede použít vědomostí k řešení problémových situací, které nebyly ve výuce doposud řešeny. <i>Typická aktivní slovesa:</i> rozhodnout, provést rozbor, kombinovat, vyvrátit, obhájit, prověřit, zhodnotit, posoudit</p>

Jestliže formulujeme výukový cíl, jehož dosažení má testová úloha zkoušet, potom k tomu používáme sloves (nebo slovesných vazeb), která označujeme jako *aktivní slovesa*. Některá z nich jsou v tabulkách uvedena.

a) Určení struktury učiva, které má být testováno

Při sestavování specifikační tabulky pro didaktický test se nejdříve téma (témata), které má být testem zkoušeno, rozdělí na dílčí části (např. podle struktury výkladu v učebnici, podle názvů kapitol v programu výuky atd.). Každé této dílčí části učiva se potom přiřadí určitá váha, např. podle toho, kolik hodin bylo výuce dané části učiva věnováno, podle toho, jaký je rozsah dané části učiva v učebnici atd.

b) Určení počtu úloh v testu

Dalším krokem při sestavování specifikační tabulky je rozhodnutí, kolik úloh celkem má test obsahovat. O počtu úloh (min. 10) v testu rozhoduje řada okolností. Na prvním místě je to požadavek dostatečně vysoké spolehlivosti a přesnosti, tj. *reliability testu* (viz v předchozí text). Horní hranice délky testu je dána *časovými možnostmi* ve výuce. Nejdelší testy mívají čistý testovací čas 35 - 40 minut, monotematické didaktické testy 15 - 20 minut čistého času, didaktické testy, které ověřující pochopení výkladu nepřesahují 10 minut.

Počet úloh v testu závisí také na *druhu používaných testových úloh* a na jejich složitosti. U jednodušších otevřených úzkých úloh a u jednodušších úloh s výběrem odpovědí lze orientačně počítat s časem od 0,5 minuty do 1,5 minuty na jednu úlohu. U složitějších úloh jsou časové nároky přiměřeně větší.

c) určení úrovně osvojení poznatků, kterou mají úlohy ověřovat

U každé části učiva je třeba uvážit, jakou úroveň osvojení poznatků mají úlohy zkoušet. Z hlediska rozvíjení myšlení žáků a z hlediska kvality osvojovaných vědomostí je žádoucí preferovat vyšší úroveň osvojení. Není však normativně stanoveno, kolik úloh musí určitou úroveň osvojení zkoušet. Bude vždy záležet na úvaze autora testu

Specifikační tabulka představuje pro autora didaktického testu základní vodítko pro jeho konstrukci. Na základě jejího zpracování získá představu o tom, na které části učiva se mají jednotlivé testové úlohy zaměřovat, i o tom, které cílové kategorie osvojení mají postihovat. Specifikační tabulka tak vytváří předpoklady k vytvoření vyváženého didaktického testu, zkoušejícího reprezentativní výběr učiva.

Příklad 2: Specifikační tabulka pro test z fyziky POHYB A SÍLA

(7. roč. základní školy, 31 vyuč. hodin)

Při sestavování této specifikační tabulky bylo rozhodnuto, že test bude vytvořen ze 25 testových úloh výběrem odpovědí (čas na práci žáků byl orientačně stanoven na 20 minut). Plánovaných 25 testových úloh bylo rozděleno na jednotlivé části učiva tak, aby to přibližně odpovídalo procentovým podílům počtu hodin, které byly výuce dané části věnovány.

OBSAH	POČET HODIN		POČET ÚLOH		ÚROVEŇ OSVOJENÍ (DLE NIEMIERSKÉ TAX.)			
					A	B	C	D
Vzájemný pohyb těles	2	6 %	2	8 %	-	1	1	-
Posuvný pohyb tělesa	8	26 %	6	24 %	1	2	2	1
Pohybové a deformační účinky síly. Siloměr	8	26 %	6	24 %	2	2	1	1
Skládání sil. Rovnováha sil	8	26 %	6	24 %	2	1	2	1
Moment síly vzhledem k ose	3	10 %	3	12 %	-	1	1	1
Tření	2	6 %	2	8 %	-	1	1	-
<i>Celkem</i>	<i>31</i>	<i>100 %</i>	<i>25</i>	<i>100 %</i>	<i>5</i>	<i>8</i>	<i>8</i>	<i>4</i>

Specifikační tabulka dále upřesňuje, které úrovně osvojení (dle Niemiervy taxonomie) mají jednotlivé testové úlohy ověřovat. Z tabulky vyplývá, že 5 úloh v testu má zkoušet pouze zapamatování poznatků (kategorie A), 8 úloh, tj. asi 1/3, bude zkoušet porozumění poznatkům (kategorie B), dalších 8 úloh Používání vědomostí v typových situacích (kategorie C) a 4 úlohy se zaměří na zjišťování schopností žáků používat vědomosti v problémových situacích (kategorie D).

TECHNIKA SEZNAMU VÝUKOVÝCH CÍLŮ

Pracnější, zato však přesnější technikou upřesňování obsahu testu je *technika seznamu výukových cílů*. U této techniky se učební látka, která má být testována, převádí na seznam výukových cílů, kterých chceme ve výuce dosáhnout.

Při upřesňování obsahu testu touto technikou postupujeme tak, že pro dané učivo formulujeme co možná největší počet výukových cílů. V této fázi ještě nepřihlížíme k významu jednotlivých cílů, snažíme se však, aby žádný důležitý cíl nezůstal opomenut. Na druhé straně dbáme, aby v seznamu byly jen takové cíle, na něž byla výuka skutečně zaměřena. Výukové cíle formulujeme tak, aby byly zcela konkrétní, vyjadřovaly určitý výkon žáka a byly jednoznačně kontrolovatelné.

Každý výukový cíl musí být v testu zkoušen tolika úlohami, kolik odpovídá jeho výukovému významu. Výukový význam určitého cíle lze posoudit podobně jako u specifikační tabulky podle počtu hodin, které jsou nutné k jeho dosažení, případně podle rozsahu učebního textu, který bylo nutno ke splnění výukového cíle prostudovat atd. Stejně jako u specifikační

tabulky lze potom jednotlivým výukovým cílům přiřadit určitý počet testových úloh. Výhodou v tomto případě je, že úrovně osvojení poznatků přímo vyplývají z formulovaných výukových cílů a netřeba je zvlášť stanovovat.

Po skončeném plánování didaktického testu by autorovi mělo být jasné, co a na jaké úrovni a kolika testovými úlohami, má být zkoušeno.

KONSTRUKCE DIDAKTICKÉHO TESTU

Ve fázi konstrukce didaktického testu se jedná o vytvoření jednotlivých testových úloh a o vytvoření prvního návrhu didaktického testu.

Autor testu rozhoduje o použití typu úloh v didaktickém testu (zde rozhoduje cíl, který testování má plnit, obsah učiva, který má být předmětem testování, materiální a technické podmínky, ale v i obliba určitého druhu testových úloh u autora testu).

NÁVRH TESTOVÝCH ÚLOH

Didaktický test je vytvořen z jednotlivých testových úloh. Na kvalitě testových úloh závisí v podstatné míře kvalita celého testování. V didaktických testech se používají různé typy testových úloh. Nejdůležitější z nich uvedeme ve členění podle P. Byčkovského (1982):

1. otevřené (s tvořenou odpovědí)
 - 1.1. se širokou odpovědí
 - 1.1.1. nestrukturované
 - 1.1.2. se strukturou
 - 1.1.2.1. vymezenou
 - 1.1.2.2. danou konvencí
 - 1.2. se stručnou odpovědí
 - 1.2.1. produkční
 - 1.2.2. doplňovací
2. uzavřené (s nabízenou odpovědí)
 - 2.1. dichotomické
 - 2.2. s výběrem odpovědí
 - 2.3. přiřazovací
 - 2.4. uspořádací

OTEVŘENÉ ŠIROKÉ ÚLOHY

V otevřených širokých úlohách se požaduje od žáka rozsáhlejší odpověď (např. ½ strany nebo i delší). Může se např. požadovat pojednání na určité téma (např. *Význam díla K. H. Borovského pro českou poezii*; nebo *Jaké byly hlavní příčiny vzniku druhé světové války?*), vyřešení zadaného problému (např. *Navrhněte postup, kterým je možno předejít infikování počítače virem.*), popis konkrétního procesu (např. *Popište jak se zapisují data na pevný disk.*) apod. Požadovaný rozsah odpovědi se žákovi naznačuje velikostí vynechaného místa v testovém zadání.

Někdy je u otevřených širokých úloh vhodné vymezit strukturu požadované odpovědi, např.:

Příklad 3: Výpočet kvadratické rovnice

(uveďte definiční obor, početní a grafické metody používané pro výpočet kořenů rovnice).

Někdy není nutno strukturu požadované odpovědi vymezovat, protože vyplývá z konvence, kterou by zkoušený měl znát, např.:

Příklad 4: Popište princip tisku na barevné laserové tiskárně.

Otevřené široké úlohy lze doporučit zejména při zkoušení komplexních vědomostí nebo dovedností, osvojovaných v delším časovém období. Jsou vhodné pro zkoušení vyšších úrovní osvojení učiva (např. řešení problémových situací apod.), zatímco pro zkoušení nižších úrovní osvojení učiva (např. zapamatování atd.) jsou výhodnější testy s úlohami objektivně skórovatelnými. Skutečnost, že široké testové úlohy nelze objektivně skórovat, by neměla vést k jejich nepoužívání. Komplexní vědomosti a dovednosti lze sice rozložit na dovednosti dílčí (které se dají zkoušet mnohem snadněji), avšak zvládnutí dílčích dovedností ještě nemusí znamenat zvládnutí dovedností komplexních.

Široké testové úlohy se poměrně snadno navrhnou, ale jejich hlavní nevýhodou je nemožnost objektivního skórování. Při skórování širokých úloh se často postupuje tak, že za správné a úplné zodpovězení úlohy se přisuzuje určitý počet bodů (např. 10). Za každou chybnou nebo chybějící část odpovědi se potom strhává určitý počet bodů. V některých případech lze pro skórování otevřených testových úloh vypracovat detailní předpis, který umožňuje téměř objektivní skórování.

Testy vytvořené z otevřených širokých úloh se často označují jako *esej testy*. Tyto testy obsahují zpravidla jen několik úloh. Od běžných písemných zkoušek se esej testy liší tím, že při jejich konstrukci, hodnocení i interpretaci výsledků se využívají všechna základní pravidla a postupy obvyklé u ostatních didaktických testů.

ÚLOHY SE STRUČNOU ODPOVĚDÍ

Úlohy se stručnou odpovědí požadují od žáka, aby vytvořil a uvedl vlastní krátké odpovědi. Může to být např. uvedení čísla, značky, symbolu, vzorce, jednoduchého grafu, určitého slova, příp. několika slov či krátké věty. Úlohy se stručnou odpovědí mohou být dvojího typu: produkční a doplňovací.

Příklad 5: Produkční testové úlohy

Napište tři krajská moravská města.

1.....

2.....

3.....

Definuj jednotku jeden bit (b).

Napiš Pythagorovu větu.

Příklad 6: Doplňovací testové úlohy

Městem Kroměříž protéká řeka

Starší chlapci a děvčata hráli volejbal a menší děti je povzbuzovali.....

Rokubyl v Kostnici upálen Jan Hus.

K exponenciální funkci $x = a^y$ je inverzní funkce... , která má vyjádření

Výhodami úloh se stručnou odpovědí je zejména to, že se snadno navrhnou. Další jejich výhodou je, že neumožňují žákům tak lehce uhadnout správnou odpověď bez příslušných vědomostí. Většinou se předpokládá, že vytvoření odpovědi je pro žáka náročnější. Nevýhodou úloh se stručnou odpovědí je to, že žák mnohdy odpovídá správně, ale jinak, než si představoval autor testu.

Příklad 7: Uved' jednotku délky?

V tomto příkladě může být správná odpověď např. palec, *km*, *m*, *míle* apod. Odpověď může být zapsána i slovy, gramaticky nesprávně nebo nepřesně. V podobných případech nutně vznikají pochybnosti, zda úloha byla či nebyla zodpovězena správně.

Pochybnostem při skórování odpovědí je možno se částečně vyhnout tím, že budeme úlohu formulovat co nejpřesněji (v uvedeném příkladě např. můžeme požadovat uvedení „základní jednotky délky“). V zásadě však u tohoto typu testových úloh mohou vždy nastat situace, že úloha je zodpovězena jen částečně správně. Skórování testu, ve kterém jsou použity úlohy se stručnou odpovědí, proto nemůže provádět laik, nýbrž jen odborník, který zkoušenému učivu dokonale rozumí.

Doporučení pro návrh úloh se stručnou odpovědí (upraveno podle P. Byčkovského, 1982)

1. Úlohu užívejte jen tehdy, lze-li odpovědět velmi stručně (nejlépe jen jedním údajem).
2. Úlohu formulujte zcela jasně a jednoznačně.
3. Nevyžadujte doslovné opakování textu z učebnice.
4. Uvažte předem všechny možné odpovědi, a je-li jich mnoho, raději úlohu nepoužívejte.
5. Ponechte v úlohách vždy dostatek místa pro uvedení odpovědi.
6. Dávejte přednost produkčním úlohám před doplňovacími. Chcete-li přece jen použít doplňovací úlohy, dodržujte následující doporučení:
7. Vynechávejte jen důležité údaje.
8. Z neúplné věty musí být patrné, co se má doplnit.
9. Údaj, který se má doplnit, umíst'ujte pokud možno na konec věty.
10. Pokud se má doplnit několik údajů, vynechtejete pro doplnění zhruba stejné místo.

ÚLOHY DICHOTOMICKÉ

U dichotomických testových úloh jsou žákovi předkládány dvě alternativy odpovědi s tím, že jedna je správná a tu má označit (např. podtržením, zakroužkováním apod.). Tyto úlohy bývají také často označovány jako úlohy s dvoučlennou volbou. Nejčastější formy dichotomických úloh uvádějí následující příklady.

Příklad 8: Dichotomické testové úlohy

V počítači je nejmenší jednotkou nesoucí informaci.

bit - byte

Jan Masaryk byl první československý prezident.

ano - ne

Pohoří Apeniny prochází:

Řeckem - Itálií

Výhodou dichotomických úloh je, že se velmi snadno navrhují. Tato konstrukční jednoduchost však může svádět k testování jednotlivých detailů, pouhých faktů.

Nedostatkem dichotomických úloh je velká pravděpodobnost uhodnutí správné odpovědi i bez příslušných vědomostí. Aby se věrohodnost výsledků získaných testem s dichotomickými úlohami zvýšila, je třeba, aby test obsahoval těchto úloh dostatečný počet.

Doporučení pro návrh dichotomických úloh (upraveno podle P. Byčkovského, 1982)

1. Tvrzení uváděné v úloze musí být jednoznačně správné a nebo nesprávné.
2. Nepoužívejte příliš dlouhých tvrzení.
3. V tvrzeních nepoužívejte dvojího záporu.
4. V tvrzeních neužívejte výrazů typu často, téměř, vždy, nikdy, zřídka apod.
5. Navrhujte zhruba stejný počet správných a nesprávných tvrzení.
6. Nepoužívejte vět vytržených z učebnice, ani je neobměňujte zařazením záporu.

ÚLOHY S VÝBĚREM ODPOVĚDÍ

Úlohy s výběrem odpovědí (úlohy s vícečlennou či vícenásobnou odpovědí) vděčí za svoji teoretickou rozpracovanost především rozvoji programovaného učení (hlavně tzv. větvených programů). Úloha s výběrem odpovědí se skládá ze dvou částí: problému nebo otázky (tzv. kmenu úlohy) a nabídnutých odpovědí.

Úlohy s výběrem odpovědí se v didaktických testech vyskytují v několika formách.

a) Úlohy typu „jedna správná odpověď“

Základní formou je úloha, ve které žák vybírá (např. zakroužkováním) jednu správnou odpověď z několika nabídnutých alternativ.

Příklad 9: Úlohy s výběrem odpovědí (jedna odpověď správná)

a) *Povídku Pohorská vesnice napsal(a) spisovatel(ka)*

Karolína Světlá

Božena Němcová

Jindřich Šimon Baar

František Palacký

b) *Pythagorova věta lze vyjádřit vztahem:*

$$c^2 = a^2 - b^2$$

$$a^2 = c^2 - b^2$$

$$b^2 = a^2 + c^2$$

$$a^2 = b^2 + c^2$$

c) *Evropu od Afriky odděluje :*

Suezský průplav

Gibraltarský průplav

Panamský průplav

Maltský průplav

b) Úlohy typu *Jedna nejpřesnější odpověď*

Jinou formou úloh s výběrem odpovědí jsou úlohy, kde se požaduje nejlepší nebo nejsprávnější odpověď. Takové úlohy mohou být pro žáky velmi obtížné, obtížnější než odpovídající úlohy otevřené.

Příklad 10: Které z následujících tvrzení správně vyjadřuje rovnost mezi vztahy:

$$(a - b)^3 =$$

$$A : a^3 - 3a^2b + 3ab^2 - b^3$$

$$B : a^3 - 3a^2b - 3ab^2 + b^3$$

$$C : a^3 + 3a^2b - 3ab^2 - b^3$$

$$D : a^3 + 3a^2b - 3ab^2 + b^3$$

c) Úlohy typu „jedna nesprávná odpověď“

V některých případech lze také požadovat uvedení nesprávné odpovědi. V tomto případě je ovšem nutné zápor ve kmenu úlohy patřičně *zdůraznit*, protože jinak může snadno dojít k přehlédnutí a nesprávné odpovědi přesto, že žák má příslušné vědomosti.

Příklad 11: Mezi vnitřní paměti nepatří:

RAM
EPROM
CD-ROM
ROM

d) Úlohy s vícenásobnou odpovědí

Jestliže má žák v testové úloze vybrat několik správných odpovědí, hovoříme o tzv. *vícenásobné odpovědi*. Jestliže se rozhodneme pro použití této úlohy, je nutné na to žáky předem upozornit. V úlohách, kde se požaduje výběr jen jedné odpovědi, je totiž výběr většího počtu odpovědí považován za chybu, a žáci by proto mohli váhat, zda více odpovědí uvést.

Příklad 12: Mezi savce řadíme:?

člověka *
netopýra
delfína *
hada
psa*

Mohou nastat problémy u úloh s vícenásobnou odpovědí při jejich skórování. Neexistuje zde totiž jen jedna naprosto správná a jedna naprosto nesprávná odpověď, ale také několik částečně správných odpovědí. Lze doporučit dva přístupy, z nichž první lze stručně vyjádřit slovy „*všechno a nebo nic*“. Podle tohoto přístupu přidělíme 1 bod v případě, kdy žák označí všechny správné odpovědi, a 0 bodů tehdy, jestliže bude (třeba jen jedna) odpověď nesprávná. Druhý přístup (diferencovanější) spočívá v tom, že přidělíme 1 bod za každou označenou správnou část odpovědi a -1 bod za každou nesprávnou část odpovědi.

e) Situační úlohy

Zvláštní modifikací testových úloh s výběrem odpovědí jsou úlohy označované někdy jako úlohy situační či interpretační. Jsou to úlohy, u nichž žák vybírá z podstatně většího počtu nabídek, než je obvyklé, přičemž nabídky nejsou předkládány ve formě dlouhého a nepřehledného seznamu, nýbrž vyplynou přímo z dané situace. Pravděpodobnost uhodnutí správné odpovědi bez příslušných vědomostí je u tohoto typu úloh zpravidla velmi malá.

Příklad 131: Situační úloha

$823 * 43$
Na místo označené hvězdičkou napište takovou číslici, aby výsledné šesticiferné číslo bylo dělitelné třemi.

V uvedené testové úloze na první pohled nejsou nabízeny žádné možnosti, ve skutečnosti však žák vybírá z deseti číslic. Pravděpodobnost náhodné volby správné odpovědi je v tomto případě 10 %.

Problém hádání správných odpovědí u úloh s výběrem odpovědí

Častou námitkou proti užívání testových úloh s výběrem odpovědí bývá, že žák může mnohdy uhodnout správnou odpověď, aniž by učivo náležitě zvládl. Není to až tak pravda. Kromě toho lze vhodně volenými opatřeními vliv „*hádání*“ omezit na minimální míru.

Je pravda, že u uzavřených úloh existuje vždy určitá pravděpodobnost, že žák zvolí správnou odpověď zcela náhodně. Toto nebezpečí se zmenšuje s rostoucím počtem nabízených odpovědí. Jako optimální počet předkládaných odpovědí se uvádí číslo *čtyři*. Menší počet

odpovědí než čtyři se pro velkou pravděpodobnost uhodnutí správné odpovědi nedoporučuje, více než pět odpovědí činí zase úlohu nepřehlednou a také sestavování většího počtu přijatelných odpovědí je značně obtížné.

Pro dvě a tři nabízené odpovědi se někdy doporučuje tzv. *korekce na hádání*. Při používání korekce na hádání se přisoudí žákovi počet bodů podle toho, kolika chyb se v testu dopustil. Vychází se z toho, že žák, který odpověď hádá, se dopouští častěji chyb než ten, který úlohy skutečně řeší a odpovídá jedině tehdy, když odpověď zná. Korekci dosažených bodových výsledků lze provést podle vzorce

$$s_o = s_n - \frac{n}{y-1} \quad (1)$$

s_o je tzv. opravené skóre (opravený počet bodů), s_n je neopravené skóre, n počet nesprávných odpovědí daného žáka v testu a y je počet nabízených odpovědí v jedné úloze.

Příklad 14: Korekce „na hádání“

V didaktickém testu, který byl sestaven z 30 dichotomických úloh, žák K uvedl 18 správných odpovědí a 10 nesprávných odpovědí (v 2 úlohách neodpověděl) v testu byly tři alternativy nabízených odpovědí. Dosadíme-li tyto hodnoty do vztahu (1) dostáváme

$$s_o = s_n - \frac{n}{y-1} = 18 - \frac{10}{3-1} = 13$$

Žákovi K tedy přisoudíme, přestože odpověděl v 18 úlohách správně, pouze 13 bodů.

Jestliže se provádí korekce na hádání, je nutno na tuto skutečnost žáky v úvodní informaci upozornit. Je třeba jim vysvětlit, že ve sporných případech je pro ně výhodnější neodpovídat vůbec. Naopak v testech, kde se korekce na hádání neprovádí, je pro zkušeného výhodnější zodpovědět všechny úlohy. Zkušenosti ukazují, že použití korekce na hádání je sice „*statisticky spravedlivé*“, ale nemusí být spravedlivé vzhledem k jednotlivým žákům. Při tomto postupu dochází např. k poškozování žáků velmi kritických ke své práci. Užití korekce na hádání má své opodstatnění zejména tehdy, jestliže chceme objektivně posoudit úroveň vědomostí celé skupiny žáků (třída, ročník).

U některých testových úloh s výběrem odpovědí lze omezit hádání správné odpovědi tím, že požadujeme (jako podmínku pro uznání odpovědi) výpočet nebo zdůvodnění odpovědi.

Pravděpodobnost uhodnutí správné odpovědi lze také v některých případech snížit užitím tzv. *neurčité odpovědi*.

Příklad 15: Úloha s neurčitou odpovědí

Určete kolik minut je 60 % ze 2 hodin:

A 36 min

B 72 min

C 18 min

D správná hodnota není uvedena

V tomto případě musí žák počet minut skutečně vypočítat, nestačí jen odhadovat. Nemůže si totiž být jist tím, že správná hodnota je nabídnuta (v uvedeném příkladě je správná odpověď pod písmenem B).

Jestliže jsou v testu používány neurčité odpovědi, je třeba, aby občas byly správné. Jinak žáci velmi brzy pochopí, že jsou zařazováni jen formálně, při úvodní instrukci je také třeba žáky

na možnost neurčité odpovědi upozornit. Neobvyklost těchto odpovědí totiž často odrazuje od jejich volby.

Konstrukce testových úloh s výběrem odpovědí

Sestavování úloh s výběrem odpovědí je podstatně obtížnější než příprava úloh produkčních. Produkční úlohy vyžadují pouze správnou formulaci problému nebo otázky, úlohy s výběrem odpovědí vyžadují navíc vytvoření vhodných nabídek odpovědí, které by byly pro žáky stejně přijatelné.

Nesprávné odpovědi, které se žákům předkládají k výběru, označujeme jako *distraktory*. Návrh vhodných distraktorů je největším problémem při návrhu úloh s výběrem odpovědí. Při navrhování distraktorů většinou vycházíme z logické úvahy anebo ze zkušeností s nejčastěji se vyskytujícími chybami. U pečlivě připravovaných testů se někdy postupuje tak, že se úloha nejdříve zadá žákům jako otevřená a potom se nejčastěji vyskytující chyb použije jako distraktorů. U testové úlohy, kde distraktory plní svoji funkci, by mělo platit, že žák (který správnou odpověď nezná) vybírá náhodně ze všech předložených nabídek. Přitom je třeba si uvědomit, že přijatelnost (atraktivita) distraktorů je vlastnost značně relativní. Stejný distraktor se může jevit žákovi s nižší úrovní vědomostí dostatečně atraktivní, zatímco žákovi s vyšší úrovní vědomostí jako nepřijatelný.

Důležitou zásadou při navrhování úloh s výběrem odpovědí je, že informace, *otázka i navrhované odpovědi musí být co nejstručnější*. Dlouhý a nepřehledný text odvádí žáky od podstaty problému a může je dezorientovat. Je třeba si uvědomit, že někteří žáci (zejména na základní škole nebo na některých středních odborných učilištích) mají značné potíže se čtením a většina jejich intelektuální energie se může spotřebovávat na překonávání dlouhého a nepřehledného textu. Často bývá možné nahradit dlouhý text vhodným obrázkem, náčrtem nebo grafem. Obrázek (i poměrně složitý) dokáží žáci číst daleko rychleji.

Příklad 16: V školním roce 2005/2006 bylo třídě 2.1 10 hochů a 20 dívek. Najdi graf popisující danou situaci:

A)



B)



C)



D)



Testové úlohy s výběrem odpovědí bývají často ze strany odpůrců testování kritizovány. Bývá jim např. vytýkáno, že jsou pro žáky daleko snadnější než úlohy otevřené, že pro žáka je daleko snazší vybrat správnou odpověď mezi předloženými odpověďmi nesprávnými než vytvořit odpověď vlastní. Tyto námitky jsou plně opodstatněné zejména u úloh, které zkoušejí pouze zapamatování konkrétních poznatků. Jestliže však testová úloha zkouší některou z vyšších úrovní osvojení, potom u kvalifikovaně sestavených úloh bývá obtížnost stejná, někdy dokonce vyšší než u úloh otevřených.

Doporučení pro návrh úloh s výběrem odpovědí (upraveno podle P. Byčkovského, 1982)

1. Úlohami s výběrem odpovědí nezkoušíme pokud možno zapamatování konkrétních poznatků.
2. Ve formulaci úlohy se vyhýbáme slovům nebo údajům, které by mohly sloužit jako nápověda.
3. Pokud se ve formulaci úlohy vyskytuje zápor, zvýrazníme jej např. podtržením.
4. Soubor nabízených odpovědí k jedné úloze by měl být homogenní, tj. podobný obsahovým zaměřením i formou.
5. Distraktory se nesmějí navzájem překrývat nebo jinou formou vyjadřovat totéž.
6. Umístění správné odpovědi mezi distraktory se má volit zcela náhodně.
7. Navrhujeme jen takové distraktory, u nichž je předpoklad, že budou využívány.
8. Při používání úloh s vícenásobnou volbou odpovědi a při používání neurčitých odpovědí na tuto skutečnost žáky upozorníme.
9. Při formulaci úloh s výběrem odpovědí dáváme přednost otázkám před neúplnými tvrzeními.
10. V úlohách s výběrem odpovědí se vyhýbáme příliš dlouhým slovním formulacím.

PŘÍRAZOVACÍ ÚLOHY

Přířazovací úlohy obsahují dvě množiny pojmů a instrukci. Úkolem žáka je správně přiřadit pojmy jedné množiny k pojmům množiny druhé.

Příklad 17: Přířazovací úlohy

a) K názvům paměťových médií v levém sloupci správně přiřaďte hodnoty jejich kapacit z pravého sloupce:

<i>HDD</i>	()	<i>A</i>	<i>128 TB</i>
<i>FDD</i>	()	<i>B</i>	<i>512 MB</i>
<i>RAM</i>	()	<i>C</i>	<i>650 MB</i>
<i>CD-ROM</i>	()	<i>D</i>	<i>3 GB</i>
	()	<i>E</i>	<i>1,44 MB</i>

b) K jednotlivým světadílům přiřaďte pohoří:

<i>Austrálie</i>	()	<i>A</i>	<i>Velký Atlas</i>
<i>Evropa</i>	()	<i>S</i>	<i>Kordillery</i>
<i>Amerika</i>	()	<i>C</i>	<i>Velké předělové pohoří</i>
<i>Asie</i>		<i>D</i>	<i>Karákoram</i>
		<i>E</i>	<i>Pyreneje</i>

V uvedených příkladech je vždy počet pojmů v pravém sloupci úmyslně větší než počet pojmů v levém sloupci. Tuto zásadu bychom měli u tohoto typu úloh dodržovat vždy. Pokud by totiž obě množiny pojmů byly stejně početné, měl by žák, který některá přiřazení zná, situaci usnadněnu tím, že by se mu počet možných přiřazení zmenšoval.

Výhodou přířazovacích testových úloh je, že omezují možnost uhodnutí správné odpovědi na minimální míru. Jejich použití je však možné jen v poměrně omezeném okruhu učiva.

USPOŘÁDACÍ ÚLOHY

V uspořádacích testových úlohách se od žáka požaduje, aby uspořádal prvky dané množiny pojmů jedné třídy do řady. Úloha tohoto typu se skládá z dané množiny prvků a z instrukce, která uvádí, podle kterého kritéria a jakým způsobem se mají prvky uspořádat. Prvky je možno řadit např. podle velikosti, významu, stupně obecnosti, chronologicky atd.

Příklad 18: Uspořádací úlohy

a) Seřadte významné české krále podle období ve kterém vládli a to od období nejstaršího směrem k mladšímu období:

Josef II.
Přemysl I. Otakar
Marie Terezie
Václav III.
Jan Lucemburský
Rudolf II.

b) Seřadte uvedené živočichy podle délky života a začněte s nejvyšším věkem až po nejnižší:

liška obecná
sloní želva
moucha domácí
medvěd hnědý
buk lesní

c) Poskládejte jednotlivé státy od nejlidnatějšího po méně lidnaté:

Německo
Japonsko
Čína
USA
Indie

Určité problémy může působit skórování uspořádacích úloh. Nesprávné seřazení prvků ve skupině může totiž být provedeno mnoha způsoby, přičemž se jedná o různě závažné chyby.

Nejjednodušší způsob skórování se provádí tak, že za zcela správné vyřešení úlohy, tj. za uvedení naprosto správného pořadí, se přiděluje 1 bod, za všechna ostatní řešení 0 bodů. Tento postup se doporučuje v případech, kdy počet seřazovaných pojmů je menší nebo roven pěti..

ZÁSADY PLATNÉ PRO NAVRHOVÁNÍ VŠECH DRUHŮ TESTOVÝCH ÚLOH

Při navrhování všech typů testových úloh by se měly dodržovat některé obecně platné zásady.

1. Vyhýbáme se úlohám *kvizového charakteru*. Tyto úlohy mohou být sice pro žáky zábavné, ale k serióznímu měření výsledků vzdělávací činnosti se nehodí.

Příklad 19: Počítačová ikona je

A grafická značka zastupující program, dokument, soubor, ...
B židovský náboženský obraz
C fotografie vlastníka počítače
D nástroj na spouštění her

2. Snažíme se navrhovat testové úlohy, které jsou *navzájem nezávislé*, tj. takové, u nichž správné vyřešení jedné úlohy není vázáno na správné vyřešení jiné úlohy v testu.

3. Při formulaci testových úloh musíme dbát na to, aby již tyto formulace neobsahovaly *nápovědu správné odpovědi* (tzv. nezamýšlená nápověda). Zdrojem nápovědy u úloh s výběrem odpovědi bývá často např. obsahová, formální nebo jiná odlišnost správné odpovědi od distraktorů.

4. V didaktických testech zásadně *nepoužíváme* tzv. „*chytáky*“, u nichž nezkoušíme stupeň zvládnutí učiva, ale zcela jiné charakteristiky žáka (např. postřeh, vtip atd.).

5. Při hodnocení odpovědí v testu je nevhodnější užívat tzv. *jednoduchého skórování* úloh, kdy za správnou odpověď v kterékoli úloze připisujeme vždy jen jeden bod. Složitější a pracnější - tzv. *vážené skórování* (u něhož přiřazujeme různé počty bodů úlohám podle jejich náročnosti) se většinou doporučuje u úloh, které se výrazně liší časem, který je třeba k jejich vyřešení (zodpovězení).

6. Testových úloh bychom měli *navrhovat* vždy o něco *více*, než kolik má obsahovat konečná podoba testu. Při ověřování testu se totiž zpravidla ukáže, že řada úloh z různých příčin nevyhovuje.

7. Náležitou pozornost je třeba také věnovat *grafické úpravě* úloh. Text musí být dokonale čitelný a přehledný, písmo dostatečně velké a výrazné.

NÁVRH PROTOTYPU DIDAKTICKÉHO TESTU

Pokud autor testu respektoval při návrhu úloh doporučení uváděná v předchozí části, lze předpokládat, že se vyvaroval základních nedostatků v obsahové i formální stavbě úloh. Málodky se však podaří navrhnout napoprvé úlohy zcela dokonalé. Zkušenosti ukazují, že bývá prospěšné vytvořené testové úlohy na nějaký čas (alespoň několik dnů) po návrhu odložit a teprve potom se k nim znovu kriticky vrátit.

Další metodou, jak přispět k vytvoření co nejkvalitnějších úloh, je nechat navržené úlohy posoudit další kompetentní osobou (ještě lépe několika osobami). Je dobré ptát se na následující otázky:

- Zkouší navržené úlohy skutečně to, co je v daném učivu nejdůležitější? Není něco podstatného vynecháno?
- Jsou očekávané správné odpovědi jednoznačné?
- Jak jsou navržené úlohy obtížné?

Z úloh, které obstály při opakovaném posuzování autorem, (příp. při posuzování jinými odborníky) sestavíme první návrh didaktického testu. Při konstrukci rozlišujícího testu, je třeba, úlohy v testu seřadit podle odhadované vzrůstající obtížnosti.

Součástí přípravy testu k jeho použití je také přibližné určení času, který budou žáci k vypracování potřebovat (u jednodušších otevřených úloh a u jednodušších úloh s výběrem odpovědí je dán čas od 0,5 minuty do 1,5 minuty na jednu úlohu).

Konkrétnější představu o časových nárocích na zpracování testu je možno si učinit až po prvním použití testu na vzorku žáků. Většinou se doporučuje, aby časový limit pro testy úrovně byl zvolen tak, aby 80 - 90 % žáků stačilo testem projít. Abychom předešli opisování žáků při práci s testem, bývá zpravidla nutné vypracovat test ve dvou ekvivalentních formách. Nejjednodušší je vytvořit druhou formu testu náhodným přeskupením úloh v původní formě.

OVĚŘOVÁNÍ A OPTIMALIZACE DIDAKTICKÉHO TESTU

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

Důkladnost ověřování testu závisí zejména na počtu žáků, kteří se ověřování zúčastní. Pokud je na škole několik paralelních tříd, je výhodné využít všech tříd. Učitelé na menších školách mohou např. také využít spolupráce s kolegy z jiných škol. U standardizovaných didaktických testů se při ověřování pracuje se vzorky 300 - 500 žáků.

Při výkladu jednotlivých analýz budeme postupovat tak, že popíšeme všechny důležité kroky (které přichází v úvahu) s tím, že vždy upozorníme na to, který krok je zcela nezbytný a který je možno (např. v případě nestandardizovaného testu) provést zjednodušeně nebo případně vůbec vynechat.

ANALÝZA VLASTNOSTÍ TESTOVÝCH ÚLOH

Na kvalitě vlastností úloh je závislá kvalita testu jako celku. Analýza vlastností testových úloh se zaměřuje na obtížnost úloh, na citlivost úloh a na tzv. nenormované odpovědi.

OBTÍŽNOST ÚLOHY

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

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

$$Q = 100 \frac{n_n}{n} \quad (2)$$

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

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

$$P = 100 \frac{n_n}{n} \quad (3)$$

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

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

$$Q = 100 - P \quad (4)$$

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

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

CITLIVOST TESTOVÝCH ÚLOH

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

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

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

Velice nízkých (případně záporných) hodnot nabývá koeficient citlivosti např. v úlohách, které jsou příliš komplikovaně formulovány a kdy obě skupiny žáků (skupina *L* a skupina *H*) mohou volit rozdílné strategie řešení. Zatímco žáci „lepší“ se snaží k řešení komplikované úlohy dojít složitými úvahami, při kterých se snadno dopouštějí chyb, žáci „horší“ se složitými úvahami příliš neznepokojují a odpověď s větším či menším štěstím hádají. Nízké hodnoty koeficientu citlivosti se objevují také u některých velmi obtížných úloh s výběrovými odpověďmi a nebo u úloh, kde se zkouší formálně zvládnuté učivo.

METODY VÝPOČTU KOEFICIENTU CITLIVOSTI

a) Koeficient ULI

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

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

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

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

b) Tetrachorický koeficient citlivosti

Poněkud pracnější, ale většinou spolehlivější metodou výpočtu koeficientu citlivosti je výpočet tzv. tetrachorického koeficientu citlivosti r_{tet} . Pro výpočet tohoto koeficientu je třeba pro každou testovou úlohu sestavit tzv. čtyřpolní (tetrachorickou) tabulku, která uvádí počty žáků ze skupin L a H, kteří v úloze odpověděli správně (+) a nebo špatně, případně neodpověděli (-).

Jednotlivé počty žáků jsou ve schématu čtyřpolní tabulky označeny písmeny a, b, c, d.

		Odpověď	
		+	-
Skupina	L	a	b
	H	c	d

Tabulka 4: Schéma čtyřpolní tabulky

Z uvedeného schématu vyplývá, že a je počet žáků ze skupiny L, kteří odpověděli v úloze správně, b je počet žáků ze skupiny L, kteří odpověděli v úloze špatně (nebo neodpověděli), c je počet žáků ze skupiny H, kteří odpověděli v úloze správně a d je počet žáků ze skupiny H, kteří odpověděli v úloze špatně a nebo neodpověděli.

Tetrachorický koeficient citlivosti úlohy se vypočítává ze vztahu

$$t_{tet} = \cos\left(180 \frac{\sqrt{bc}}{\sqrt{bc} + \sqrt{ad}}\right) \quad (6)$$

kde \cos je goniometrická funkce kosinus a čísla a , b , c , d mají význam, který vyplývá ze schématu čtyřpolní tabulky.

Většinou se uvádí, že tetrachorický koeficient citlivosti by neměl být u vyhovujících testových úloh nižší než 0,15. Tato hodnota platí v případě, že skupiny L a H byly vytvořeny z 50 % všech testovaných žáků. Jestliže k vytvoření skupin L a H použijeme menší počet žáků (např. 33 % nebo 25 %), pak při posuzování vypočítaných hodnot r_{tet} je nutno uplatňovat poněkud přísnější měřítko.

c) Bodově biseriální koeficient citlivosti

Někdy se citlivost testových úloh posuzuje také pomocí bodově biseriálního koeficientu. Příslušný vzorec pro výpočet je

$${}_b r_{bis} = \frac{\bar{x}_s - \bar{x}_n}{s_x} \sqrt{pq} \quad (7)$$

kde ${}_b r_{bis}$ je bodově biseriální koeficient citlivosti testové úlohy, \bar{x}_s je průměrný počet bodů v testu u žáků, kteří danou úlohu řešili správně, \bar{x}_n průměrný počet bodů u žáků, kteří danou úlohu řešili nesprávně, s_x směrodatná odchylka, vypočítaná ze všech testových výsledků, $p = 0,01P$, kde P je index obtížnosti testové úlohy, $q = 1 - p$.

Vyhovující testová úloha by měla vykazovat bodově biseriální koeficient citlivosti minimálně 0,20.

ANALÝZA NENORMOVANÝCH ODPOVĚDÍ

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

Rozbor vynechaných odpovědí

Jestliže zjistíme, že některé odpovědi jsou vynechány, může to znamenat neznalost učiva, nepochopení formulace úlohy, nedostatek času k vypracování odpovědí atd. Je třeba věnovat zvýšenou pozornost zejména těm otevřeným úlohám, ve kterých odpověď vynechalo více než 30 - 40 % žáků. U uzavřených úloh je však třeba věnovat zvýšenou pozornost i úlohám, kde neodpovědělo více než 20 % žáků.

Rozbor nesprávných odpovědí

Rozbor nesprávných odpovědí je velmi jednoduchý u úloh s výběrem odpovědí. V tomto případě postačí překontrolovat, zda všechny nabídnuté distraktory (nesprávné nabídky) jsou pro žáky dostatečně atraktivní. To, zda všechny distraktory jsou dostatečně atraktivní, se pozná podle toho, že žáci vybírají skutečně ze všech nabídnutých. Ten distraktor, který nikdo (nebo téměř nikdo) ze žáků nevolí, neplní svoji funkci a měl by být, nahrazen jiným, atraktivnějším distraktorem anebo z úlohy odstraněn. Nefungující distraktor zbytečně test zatěžuje, protože žáci musí část své energie neúčelně vynaložit na jeho čtení a úvahy o něm. Při správně fungujících distraktorech by žák, který správnou odpověď nezná, měl více méně náhodně volit jednu z nabídek. Nalezení dostatečného počtu vhodných distraktorů je jedním z nejobtížnějších problémů při konstrukci úloh s výběrem odpovědí.

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

ÚPRAVA VYTVOŘENÉHO DIDAKTICKÉHO TESTU

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

- úloha je příliš obtížná anebo příliš snadná (hodnota obtížnosti Q je buď větší než 80 anebo menší než 20).
- úloha málo rozlišuje mezi žáky s dobrými a špatnými vědomostmi (např. koeficient citlivosti d je u středně obtížných úloh menší než 0,25 apod.).
- v testové úloze je příliš mnoho vynechaných odpovědí (u otevřených úloh např. více než 30 – 40 %, u uzavřených úloh více než 20 %).
- počet vedlejších chyb v úloze převažuje nad počtem hlavních chyb (u úloh otevřených).
- žáci nevybírají ze všech nabídnutých distraktorů v úloze (u úloh s výběrem odpovědí).

Vytvoření definitivní podoby didaktického testu

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

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

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

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

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

U některých testových úloh lze ekvivalentní formu vytvořit obměnou údajů v zadání úlohy (např. změnou číselných hodnot v příkladech). Ekvivalentní formu testu lze vytvořit i kombinací způsobů shora uvedených.

TVORBA STANDARDIZOVANÉHO DIDAKTICKÉHO TESTU

Kvalitně zpracovaný nestandardizovaný test může být pro učitele zdrojem velmi cenných informací o průběhu a výsledcích výuky. U nestandardizovaného didaktického testu však nemáme možnost dosažené výsledky žáků srovnávat s ostatními žáky, nemáme k dispozici klasifikační nebo jiný standard, na jehož základě bychom mohli žáky objektivně hodnotit nebo klasifikovat. Konečně, u nestandardizovaného didaktického testu také zpravidla nevíme, jak dalece jsou výsledky testování dotčeny náhodou, jak dalece jsou spolehlivé a přesné, tj. *reliabilní*.

URČENÍ RELIABILITY DIDAKTICKÉHO TESTU

Nejdůležitější vlastností didaktického testu je jeho *validita*, tj. záruka, že test skutečně zkouší (měří) to, co zkoušet má. Nutným předpokladem dobré validity (i když ne jediným) je, že výsledky testování nejsou v příliš ovlivněny náhodnými vlivy. Takový didaktický test je spolehlivý a přesný, tj. má dobrou *reliabilitu*. Zatímco posuzování validity je značně komplexní a je často značně ovlivňováno postoji a názory autora testu, je *reliabilita* do jisté míry záležitostí technickou a lze ji poměrně snadno exaktně odhadovat.

U standardizovaných didaktických testů je posouzení *reliability* běžnou a obecně vyžadovanou záležitostí. U testů nestandardizovaných se *reliabilita* většinou nestanovuje, přestože nejde o příliš složitou záležitost.

Výpočet *reliability* umožňuje (a to i v případě nestandardizovaného didaktického testu) posoudit, jak dalece jsou výsledky testování ovlivněny náhodou a jak jsou tudíž věrohodné.

Vypočítaná reliabilita u nestandardizovaných didaktických testů má, vzhledem k malým počtům žáků, z nichž byla počítána, jen omezenou platnost a výsledky budou spíše nižší než ty, které by byly získány na velkých vzorcích. Ale i tak představuje pro autora testu zdroj cenných informací.

V dalším textu budou popsány dva často používané postupy, kterými lze reliabilitu vypočítat.

RELIABILITA TESTU POMOCÍ KUDEROVA-RICHARDSONOVA VZORCE

Tento model výpočtu koeficientu reliability je vhodný pro didaktické testy úrovně, které jsou složeny z obsahově homogenních úloh. U testů vytvořených z obsahově nehomogenních úloh vychází touto metodou koeficient reliability příliš nízký. Výpočet koeficientu reliability se provádí pomocí Kuderova-Richardsonova vzorce

$$r_{kr} = \frac{k}{k-1} \left(1 - \frac{\sum pq}{s^2} \right) \quad (8)$$

kde k je počet úloh v testu, p je podíl žáků ve vzorku, kteří řešili určitou úlohu v testu správně. Platí tedy, že

$$p = \frac{n_s}{n} \quad (9)$$

kde n_s je počet žáků, kteří určitou úlohu řešili správně, a n je celkový počet žáků; $q = 1 - p$ a s je směrodatná odchylka pro celkové výsledky žáků v testu.

K výpočtu koeficientu reliability bylo dále třeba určit pro dosažené testové výsledky aritmetický průměr a směrodatnou odchylku pomocí vzorců:

$$\bar{x} = \frac{1}{n} \sum n_i * x_i \quad (10)$$

$$s^2 = \frac{1}{n-1} \sum n_i (x_i - \bar{x})^2 \quad (11)$$

kde \bar{x} je vážený aritmetický průměr výsledků žáků v testu, s směrodatná odchylka pro výsledky v testu (výraz s^2 se nazývá rozptyl), n celkový počet testovaných žáků, x_i jednotlivé dosažené počty bodů, n_i počty žáků, kteří dosáhli výsledků x_i

RELIABILITA TESTU METODOU PŮLENÍ

Výhodou této metody je, že se dá použít jak pro testy úrovně, tak pro testy rychlosti. Podmínkou pro použití tohoto modelu výpočtu je, že didaktický test obsahuje *sudý* počet úloh a jednotlivé úlohy jsou řazeny podle vzrůstající obtížnosti.

Při výpočtu se postupuje většinou tak, že celý test se rozdělí na dvě poloviny tím způsobem, že jednu polovinu tvoří úlohy s lichým pořadovým číslem a druhou polovinu úlohy se sudým pořadovým číslem. Výsledky dosažené jednotlivými žáky v obou polovinách testu se potom navzájem korelují. Z hodnoty vypočítaného korelačního koeficientu se vychází při stanovení koeficientu reliability. Samotný výpočet koeficientu reliability metodou půlení se provádí pomocí Spearmanova-Brownova vzorce.

Didaktické testování

Spearmanův-Brownův vzorec:

$$r_{sb} = \frac{2 * r_p}{1 + r_p} \quad (12)$$

kde r_{sb} je koeficient reliability a r_p je koeficient korelace mezi výsledky žáků v obou polovinách didaktického testu.

Výsledky žáků v polovině L jsou označeny x_L a výsledky žáků v polovině S jsou označeny x_S .

Koeficient korelace r_p pro hodnoty x_L a x_S lze vypočítat podle vzorce

$$r_p = \frac{n \sum x_L * x_S - \sum x_L * \sum x_S}{\sqrt{[n \sum x_L^2 - (\sum x_L)^2] * [n \sum x_S^2 - (\sum x_S)^2]}} \quad (13)$$

kde n je celkový počet testovaných žáků .

STANDARDIZACE DIDAKTICKÉHO TESTU

Standardizace je významnou etapou tvorby didaktického testu. Smyslem standardizace je vytvoření testového standardu (testové normy), který umožní zařadit žáka podle dosaženého počtu bodů do určitého žebříčku (stupnice, škály).

Dosažený počet bodů v testu (hrubé skóre) sám o sobě nic neříká o tom, zda výkon žáka je průměrný, dobrý či slabý. Získá-li např. žák ve dvou testech stejný počet bodů, může to znamenat v jednom testu vynikající výkon, zatímco ve druhém pouze výkon průměrný. Teprve na základě srovnání dosaženého výkonu žáka s výkony ostatních žáků můžeme daného jedince adekvátně posoudit.

U standardizovaných didaktických testů se výkon jednotlivých žáků porovnává s reprezentativním vzorkem žáků (zpravidla se jedná o stovky žáků). Postup, kterým se toto srovnávání realizuje, se nazývá *standardizace testu*. Standardizovat výsledky určitého testu znamená vyjádřit je vzhledem k výsledkům standardizačního vzorku žáků. Jednodušší metody standardizace jsou založeny na zjišťování procent žáků, kteří v reprezentativním vzorku dosáhli určitého výsledku. Složitější metody standardizace většinou předpokládají, že výsledky testování odpovídají předpokladu tzv. normálního rozdělení a vycházejí z určování vzdálenosti jednotlivého výsledku od aritmetického průměru, přičemž jednotkou této vzdálenosti je směrodatná odchylka.

Mezi metody standardizace didaktických testů patří percentilová škála, C – škála, škála Stanin, z – škála, Z – škála, T – škála. Blíže bude vysvětlena percentilová škála.

Nejjednodušší metodou standardizace je standardizace pomocí percentilů. U této metody se ke každému dosaženému počtu bodů (hrubému skóre) přiřadí tzv. *percentilové pořadí*, které udává, kolik procent žáků ve vzorku dosáhlo horšího výkonu. To umožňuje posoudit, jaké je relativní postavení žáka ve skupině (např. v populaci žáků).

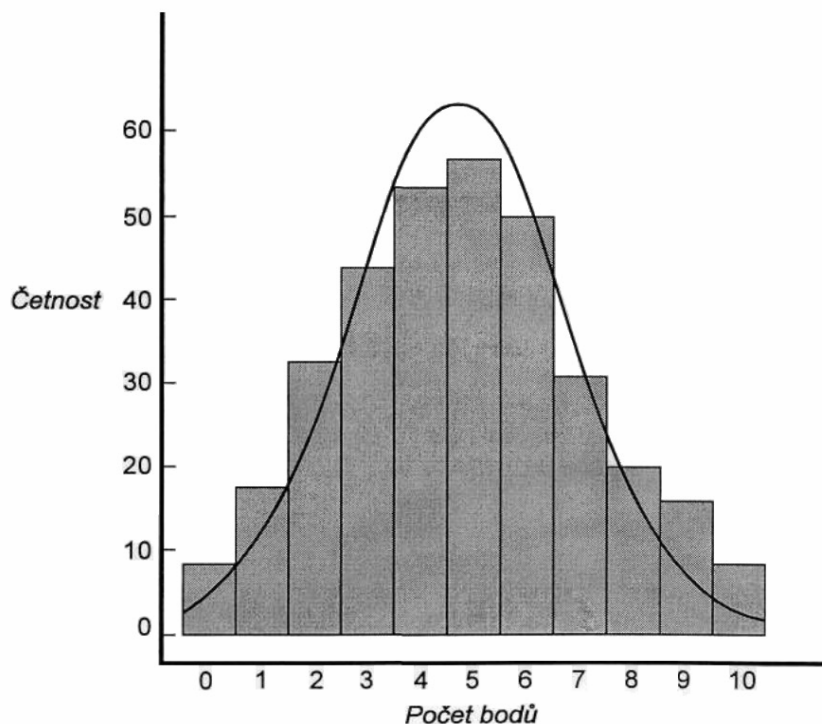
Percentilové pořadí pro určitý výsledek v testu lze vypočítat podle vzorce

$$PR = 100 * \frac{n_k - \frac{n_i}{2}}{n} \quad (14)$$

kde PR je percentilové pořadí žáka pro daný výsledek v testu, n_k je kumulativní četnost u daného výsledku, n_i je četnost daného výsledku a n je počet testovaných žáků.

Výsledky testu lze také graficky znázornit jako závislost četnosti (počtu žáků) na dosaženém počtu bodů. Pak dostáváme histogram četností. Obrázek ukazuje ideální rozdělení četností pro test s celkovým počtem deseti úloh, kde každá úloha byla ohodnocena jedním bodem. Z výsledků je patrné, že výkon 5 bodů v testu znamená téměř přesně průměrný výkon.

Obr. 1: Histogram četností pro výsledky testu



VYTVOŘENÍ KLASIFIKAČNÍHO STANDARDU PRO DIDAKTICKÝ TEST

Některé standardizované didaktické testy umožňují také vyjádřit výkon žáka v podobě klasifikačních stupňů. Při stanovení tohoto klasifikačního standardu (normy) se většinou vychází z předpokladu normálního rozdělení četností výsledků v populaci testovaných žáků. Splnění tohoto předpokladu by se mělo v každém případě opět ověřovat.

Nejdříve se rozhodneme, jaké procentové ekvivalenty zvolíme pro jednotlivé klasifikační stupně. Často se doporučuje např. rozdělení:

- 7 % nejlepších žáků – výborně
- 24 % žáků – chvalitebně
- 38 % žáků dobře
- 24 % žáků dostatečně
- 7 % nejhorších žáků nedostatečně.

Volbou těchto procentových ekvivalentů je vlastně do klasifikace vkládán jediný subjektivní moment.

Pro jednotlivé procentové ekvivalenty se potom vypočítávají odpovídající skóre (počty bodů). Pro výše uvedené rozdělení se tato skóre vypočítají ze vztahů:

$$\begin{aligned} x_{(1)} &= \bar{x} + 1,48 * s \\ x_{(2)} &= \bar{x} + 1,50 * s \\ x_{(3)} &= \bar{x} - 0,50 * s \\ x_{(4)} &= \bar{x} - 1,48 * s \end{aligned} \quad (15)$$

kde $x_{(1)}$ je minimální počet bodů, které nutno dosáhnout pro klasifikaci stupněm 1, $x_{(2)}$ je minimální počet bodů pro klasifikaci stupněm 2, $x_{(3)}$ minimální počet bodů pro klasifikaci stupněm 3 a $x_{(4)}$ je minimální počet bodů potřebných k dosažení klasifikace stupněm 4, \bar{x} je aritmetický průměr všech výsledků v testu a s je směrodatná odchylka výsledků.

Pokud bychom chtěli použít jiné procentové ekvivalenty pro jednotlivé klasifikační stupně, potom by ve vztazích (15) byly jiné tzv. *kvantity normálního rozdělení*. Např. pro rozdělení 15 % - 20 % - 30 % - 20 % - 15 % by se příslušné skóre pro dosažení určitého klasifikačního stupně určovaly ze vztahů:

$$\begin{aligned} x_{(1)} &= \bar{x} + 1,04 * s \\ x_{(2)} &= \bar{x} + 0,39 * s \\ x_{(3)} &= \bar{x} - 0,39 * s \\ x_{(4)} &= \bar{x} - 1,48 * s \end{aligned} \quad (16)$$

Takže pro test s deseti testovými úlohami, kde každá úloha je ohodnocena jedním bodem můžeme dostat následující klasifikační standard:

<i>Klasifikace</i>	<i>Počet bodů</i>
<i>výborně.....(1)</i>	8 - 10
<i>chvalitebně.....(2)</i>	6 - 7
<i>dobře.....(3)</i>	4 - 5
<i>dostatečně.....(4)</i>	1 - 3
<i>nedostatečně.....(5)</i>	0

POUŽÍVÁNÍ DIDAKTICKÝCH TESTŮ VE ŠKOLNÍ PRAXI

Z výsledků didaktického testu by měl učitel získat co nejvíce informací k hodnocení žáků a také k optimalizaci svého dalšího pedagogického působení.

DIAGNOSTICKÝ ROZBOR VÝSLEDKŮ TESTU

Diagnostický rozbor výsledků by měl následovat prakticky po každém použití didaktického testu. Při tomto rozboru si učitel všímá především chyb, kterých se žáci dopustili, a hledá jejich pravděpodobné příčiny. Tento rozbor učitel naznačuje, kterým směrem by se měla jeho činnost ubírat po provedení testu.

POSOUZENÍ CELKOVÝCH VÝSLEDKŮ TŘÍDY

Dosažené výsledky třídy, případně školy, se obvykle posuzují podle průměrného počtu dosažených bodů. Aritmetický průměr se u výsledků didaktického testu nejvýhodněji počítá podle vzorce (10).

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

KLASIFIKACE VÝSLEDKU TESTU

Mnoho nejasností panuje v otázce převodu bodového hodnocení na klasifikační stupně. Je to pochopitelné, neboť ani v teorii nebyl tento problém doposud spolehlivě a jednoznačně rozřešen. Uvedeme některé přístupy, které se v této oblasti nejčastěji uplatňují.

a) Intuitivní přístup ke klasifikaci

Někteří učitelé přistupují k převodu bodových výsledků na klasifikační stupně zcela subjektivně a sami víceméně intuitivně určují, kolik bodů je potřeba na dosažení určité známky.

Někteří odborníci doporučují jako optimální řešení otázky převodu bodového hodnocení na klasifikační stupně normativní přiřazování klasifikačních stupňů, které se provádí na základě posudku skupiny odborníků. Tato technika je ve školní praxi dobře použitelná. Vytvořený didaktický test necháme posoudit co možná největšímu počtu učitelů v daném předmětu s tím, že je požádáme o návrh klasifikační stupnice. Z jednotlivých posudků je možno určit průměr, čímž se eliminují extrémní názory.

b) Klasifikace na základě procenta správných odpovědí

Někdy se při převodu bodových výsledků na klasifikační stupně vychází z procenta správných odpovědí, kterého žák v testu dosáhl. Jeden z návrhů (Sedláčková, 1993), jak převádět dosažené počty bodů na klasifikační stupně touto metodou, uvádí tab. 5.

Tabulka 5: Klasifikace podle procenta správně vyřešených úloh

<i>Procento správně vyřešených úloh v testu</i>			<i>Klasifikační stupeň</i>
<i>Klasifikace běžná</i>	<i>Klasifikace přísná</i>	<i>Klasifikace velmi přísná</i>	
91 - 100	96-100	95-100	1
81 - 90	88-95	90-94	2
71 - 80	82-87	85-89	3
61 - 70	70-81	80-84	4
0 - 60	0-69	0-79	5

Učitelé touto metodou často postupují např. při hodnocení písemných prací. Zadávají např. celkem 4 rozsáhlejší úlohy a hodnotí takto: všechny vyřešené úlohy - výborně, tři vyřešené úlohy - chvalitebně, dvě vyřešené úlohy-dobře, jedna vyřešená úloha - dostatečně, žádná vyřešená úloha - nedostatečně. Užívání této metody převodu bodových výsledků na

klasifikační stupně je v některých případech *diskutabilní*, protože nepřihlíží ke skutečnosti, že úlohy v testu mohou mít rozdílnou obtížnost. Odpoví-li např. žák 50 % úloh správně v testu s velmi obtížnými úlohami, může to znamenat dobrý výkon, zatímco stejný výsledek v testu velmi snadném může být výkonem velmi slabým. Tento způsob klasifikace může vcelku vyhovovat u testů, které neobsahují extrémně snadné ani extrémně obtížné úlohy.

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

c) Klasifikace na základě normálního rozdělení

Bodové výsledky žáků v testu můžeme klasifikovat také na základě *normálního rozdělení četností*. U tohoto postupu se vychází z předpokladu, že výkony dosažené v testu u dostatečně velké skupiny žáků odpovídají tzv. Gaussově křivce. To znamená, že nejvíce je vždy výkonů průměrných a na obě strany od průměru potom četností symetricky ubývá (srov. obr. 1). Jestliže klasifikujeme ve shodě s touto zákonitostí, potom nejvíce žáků klasifikujeme stupněm 3, poněkud méně žáků stupněm 2 a 4 a nejméně žáků stupněm 1 a 5. Mezi jednotlivými autory není shoda v tom, kolik procent žáků by mělo být jednotlivými stupni klasifikováno. Všechny návrhy se však shodují v tom, že rozdělení jsou symetrická kolem stupně 3 a tento klasifikační stupeň má také největší četnost. Tab. 21 uvádí některá nejčastěji doporučovaná rozdělení.

Tabulka 5: Doporučovaná rozdělení klasifikačních stupňů

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

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

Klasifikace testu odvozená z relativních kumulativních četností pro třídu má pochopitelně jen relativní platnost. Slouží spíše k vyjádření příslušnosti žáka v určité oblasti učení k určité prospěchové skupině ve třídě. Jestliže klasifikaci odvozujeme z výsledků několika paralelních tříd na škole, je její platnost pochopitelně větší.

Jestliže stejný didaktický test používáme v nezměněné podobě a za stejných podmínek několik po sobě následujících let, můžeme srovnávat a hodnotit dosahované výsledky z hlediska delšího časového období a provádět jakousi standardizaci v rámci školy a v delším čase. V tomto případě lze vytvořit ze všech výsledků, kterých žáci za určité období dosáhli, určitý „*školní klasifikační standard*“, který umožňuje posoudit určitý žakovský výkon vzhledem ke škole.

Některé standardizované testy vyjadřují výkon žáků přímo pomocí klasifikačních stupňů, jiné ho vyjadřují pomocí různých druhů standardních škál, např. percentilové škály.

V souvislosti s klasifikací výsledků testu je třeba upozornit na důležitou skutečnost, že žák, který má určité vědomosti, může dosáhnout bodových výsledků v určitém pásmu. Pásma výkonů odpovídající jednotlivým klasifikačním stupňům se překrývají, a proto žák, který dosáhl výkonu na hranici klasifikačních intervalů, může mít docela dobře vědomosti na známku lepší i na známku horší. Z toho, co bylo uvedeno vyplývá, že dosáhne-li žák výkonu na hranici intervalů, nelze z toho činit jednoznačné závěry. V těchto případech bychom měli při rozhodování o klasifikaci raději volit přezkoušení, jiný test apod.

S používáním didaktických testů je spojena řada problémů a nedořešených otázek.

Používání a zejména tvorba didaktických testů klade na učitele vysoké nároky. Autor testu by měl být dobrým odborníkem i pedagogem, měl by mít určitou kvalifikaci psychologickou, musí být orientován v oblasti statistických metod apod. Vynaložená námaha se však zanícenému učiteli vyplatí, protože testy mohou určitým způsobem zhmotnit výsledky jeho jinak jen velmi obtížně postižitelného úsilí a práce.

ZÁVĚR

Počítačové testování si stále více upevňuje svou pozici v oblasti testování žáků. A není divu. Je přitažlivé pro frekventanty testu, zlepšuje dovednosti při práci s informačními technologiemi, dává možnost zapojení studentů se změněnou pracovní schopností (například dysgrafie).

Pevně věřím, že se dobře navržený a ověřený test prováděný na počítači stane běžnou součástí vyučovacího procesu. Také věřím, že většina pedagogických pracovníků na základě svých znalostí a zkušeností je schopna vytvořit dobré didaktické testy, a tak pro školy vlastními silami zajistit kvalitní testovací materiál.

LITERATURA

BYČKOVSKÝ, P.: Základy měření výsledků výuky. Tvorba didaktického testu. Praha, ČVUT 1982.

CHRÁSKA, M.: Základy výzkumu v pedagogice. Olomouc, Pedagogická fakulta UP 1993.

CHRÁSKA, M.: Didaktické testy, Brno, Paido 1999.

SEDLÁČKOVÁ, J.: Diagnostické metody ve vyučování matematice. Olomouc, Přírodovědecká fakulta UP 1993.

SKALKOVÁ, J.: Pedagogika a výzvy nové doby. Brno, Paido 2004.

SLAVÍK, J.: Hodnocení v současné škole. Východiska a nové metody pro praxi. Praha, Portál 1999.

PASCH, M. a kol.: Od vzdělávacího programu k vyučovací hodině. Jak pracovat s kurikulem. Praha, Portál 1998.

VELIKANOVIČ, J.: Skúšanie, hodnotenie a klasifikácia žiakov. Bratislava, Slovenské pedagogické nakladateľství 1973.