

Univerzita Karlova v Praze
Přírodovědecká fakulta
Katedra učitelství a didaktiky biologie



**Ověřování biologických poznatků –
didaktické testy**

Bakalářská práce

Blanka Vošmerová

Praha 2007

Vedoucí bakalářské práce: doc. RNDr. Věra Čížková, CSc.

Prohlašuji, že jsem předloženou bakalářskou práci vypracovala samostatně, na základě uvedené literatury.

V Praze dne 18. 8. 2007

.....

Blanka Vošmerová

Na tomto místě bych chtěla poděkovat doc. RNDr. Věře Čížkové, CSc. za cenné odborné rady a připomínky, trpělivost a vstřícnost, kterou mi při tvorbě této práce věnovala.

Seznam zkratk

ACER	Australian Council for Educational Research (Australská rada pro pedagogický výzkum)
ACSPRI	Australian Consortium for Social and Political Research Incorporated (Australské společenství pro sociální a politický integrovaný výzkum)
APEID	Affiliated Centre of the Asia - Pacific Programme of Educational Innovation for Development (Program společného prostoru v Asii a Pacifiku pro zlepšení ve vývoji vzdělávání)
AusAID	Australian Agency for International Development (Australská agentura pro mezinárodní rozvoj)
CERMAT	Centrum pro reformu maturitní zkoušky
CITO	Netherlands National Institute for Educational Measurement (Nizozemský institut pro měření výsledků ve vzdělávání)
CIVED	Civic education study (Mezinárodní výzkum občanské výchovy)
CKE	Centralna Komisja Egzaminacyjna (Centrální vzdělávací komise)
COMPED	Computers in Education (Program počítače ve vzdělávání)
CZVV	Centrum pro zjišťování výsledků ve vzdělávání
ETS	Education Testing Service (Testovací servis ve vzdělávání)
ESF	Evropský sociální fond
ICCS	International Civic and Citizenship Education Study (Mezinárodní občanská a státní vzdělávací studie)
IDEA	Skupina pro vzdělávací alternativy
IEA	International Association for Evaluation of Educational Achievement (Mezinárodní organizace pro hodnocení výsledků ve vzdělávání)
LSAY	Longitudinal Survey of Australian Youth (Dlouhodobý výzkum Australské mládeže)
MŠMT	Ministerstvo školství, mládeže a tělovýchovy
NAEP	The National Assessment of Educational Progress (Národní analýza pokroku ve vzdělávání)
NIES	National Indian Education Study (Národní studie vzdělávání Indiánů)
NÚOV	Národní ústav odborného vzdělávání
OECD	Organization for Economic Cooperation Development (Organizace pro hospodářskou spolupráci a rozvoj)

OKE	Okręgowa Komisja Egzaminacyjna (Okresní vzdělávací komise)
PIRLS	Progress in International Reading Literacy Study (Mezinárodní výzkum čtenářské gramotnosti)
PISA	Programme for International Student Assessment (Program pro mezinárodní hodnocení žáků)
PTDE	Polskie Towarzystwo Diagnostyki Edukacyjnej (Polská společnost diagnostického vzdělávání)
SITES	Second Information Technology in Education Study (Druhý mezinárodní výzkum informačních technologií ve vzdělávání)
TIMSS	Third International Mathematics and Science Study (Třetí mezinárodní výzkum matematického a přírodovědného vzdělávání)
TUDA	NAEP Trial Urban District Assessment (Výzkum na úrovni městských částí v rámci projektu NAEP)
ÚIV	Ústav pro informace ve vzdělávání
UNESCO	United Nations Educational, Scientific and Cultural Organization (Organizace OSN pro výchovu, vědu a kulturu)
VÚP	Výzkumný ústav pedagogický

Abstrakt

V bakalářské práci jsou charakterizovány významné světové instituce a projekty, které se zabývají hodnocením vzdělávacího procesu v zahraničí a hodnocením výsledků vzdělávání u nás. Dále jsou popsány typy didaktických testů, testových položek a doporučení pro jejich konstrukci. Práce také seznamuje s historií testování a pohledem zahraničních autorů na tuto problematiku.

Klíčová slova

testování, didaktický test, testová položka, konstrukce testů, NAEP, CITO, IEA, ACER, ETS, CKE, PTDE, CERMAT

Abstract

This bachelor work discusses important world institutions and projects that deal with evaluating educational processes in foreign countries and evaluating educational results in our country. Further, it describes types of achievement tests, testing items and introductions for test construction. This work also introduces history of testing and informs us about perspectives of foreign authors at these problems.

Keywords

testing, achievement test, testing item, tests construction, NAEP, CITO, IEA, ACER, ETS, CKE, PTDE, CERMAT

Obsah

1. Úvod	7
2. Historie testování	8
3. Didaktické testy a evaluace vzdělávacího procesu	10
3.1 Instituce a projekty zabývající se hodnocením vyučovacího procesu v zahraničí.....	10
3.1.1 NAEP (The National Assessment of Educational Progress – Národní analýza pokroku ve vzdělávání)	10
3.1.2 CITO (Netherlands National Institute for Educational Measurement - Nizozemský institut pro měření výsledků ve vzdělávání)	13
3.1.3 IEA (International Association for Evaluation of Educational Achievement - Mezinárodní organizace pro hodnocení výsledků ve vzdělávání)	14
3.1.4 ACER (The Australian Council for Educational Research - Australská rada pro pedagogický výzkum)	16
3.1.5 ETS (Educational Testing Service - Testovací servis ve vzdělávání)	17
3.1.6 CKE (Centralna Komisja Egzaminacyjna - Centrální vzdělávací komise)	18
3.1.7 PTDE (Polskie Towarzystwo Diagnostyki Edukacyjnej - Polská společnost diagnostického vzdělávání)	18
3.2 Instituce a projekty zabývající se hodnocením vyučovacího procesu u nás	19
3.2.1 CERMAT	19
3.2.2 Scio	25
3.2.3 Kalibro	27
4. Didaktické testy	29
4.1 Didaktické testy a testování v zahraniční literatuře	29
4.2 Základní typy didaktických testů	30
4.3 Základní typy testových položek	34
4.3.1 Uzavřené položky	34
4.3.2 Otevřené položky	36
4.4 Porovnání uzavřených a otevřených položek	37
4.5 Konstrukce didaktických testů	39
5. Závěr	42
6. Použitá literatura a internetové zdroje	44

1. Úvod

Evaluace vzdělávacích výsledků je nejstarší a nejrozsáhlejší součástí pedagogické evaluace. Jejím cílem je zjišťování, měření a vyhodnocování vzdělávacích výsledků žáků, které jsou dosahovány působením vzdělávacích procesů. Thorndike a Hagen (1967) uvádějí, že do poloviny 19. století se veškeré hodnocení vzdělávacích výsledků provádělo na základě ústních zkoušek. Jejich hodnocení bylo subjektivní a neexistovala možnost srovnávání škol. Ve druhé polovině 19. století se prosadilo písemné hodnocení, postupně se začaly prosazovat standardy hodnocení písemných odpovědí a vyvinulo se testování vzdělávacích výsledků. Evaluační techniky vyvinuté v Americe se začaly používat i v evropských zemích. Rozvíjelo se konstruování testů pro různé vyučovací předměty, druhy škol a věkové skupiny žáků.

V posledních dvaceti letech tento proces pokročil od pouhého popisu jevů k cílevědomému a systematickému shromažďování, analýze, interpretaci a hodnocení informací o průběhu, podmínkách a výsledcích vzdělávání s hlavním cílem zkvalitnit vzdělávání a zvýšit odpovědnost za vynaložené finanční prostředky. Evaluace má přispět ke zdokonalení managementu a vést ke zkvalitnění kurikula. Systémy evaluace v jednotlivých zemích jsou ovlivněny podmínkami země a jejími tradicemi, což se projevuje v odlišných prioritách evaluace, konkrétních postupech a charakteru používaných evaluačních nástrojů. Společným základem evaluace je monitorování výsledků vzdělávání s využitím standardizovaných testů. Volba postupů a metod monitorování závisí na tom, jaké vzdělávací hodnoty vzdělávací politika státu podporuje a jaké stanoví priority (Průcha 1996).

Také v České republice se zájem o problematiku testování neustále zvyšuje a to na úrovni národní i mezinárodní. Vznikly u nás nové organizace na podporu vzdělávání, k nimž patří např. CERMAT, Scio a Kalibro. Tyto organizace se podílejí také na tvorbě didaktických testů a následném testování. Postupně testování překračovalo hranice škol, regionů a nakonec i státu. Naše země se zapojila do řady mezinárodních projektů, z nichž k nejvýznamnějším patří TIMSS a PISA.

K problematice hodnocení výsledků vzdělávání a didaktickým testům je také zaměřena tato práce, jejímž cílem je:

- podat přehled o historii testování
- charakterizovat vybrané instituce a projekty, zabývající se hodnocením vyučovacího procesu v zahraničí a u nás
- charakterizovat základní typy didaktických testů a typy testových položek

2. Historie testování

Slovo test je anglického původu a představuje zkoušku, jejíž podmínky jsou pro všechny stejné a jejíž výsledky mají číselný charakter. Hniličková aj. (1972) definují test jako zkoušku založenou na vědeckých poznatcích a postupech, v rámci možností objektivní, spolehlivou, přesnou a výstižnou, se srovnatelnými výsledky. Testovat lze kvalitu výrobků, přítomnost určitých chemických látek v prostředí, lidí a zvířata – je možné měřit a poměřovat jejich výkony, testovat jejich vlastnosti, dovednosti, znalosti apod. Tradiční doménou testování je psychologie, která zkoumá lidské schopnosti, testuje určité znaky osobnosti nebo psychický výkon. Testy dělíme na psychologické (např. IQ testy), pedagogické (např. výkonová motivace) a didaktické - NR testy, CR testy, E- testy (Hniličková aj., 1972). Výraz test se velmi rozšířil a dnes je běžně používán.

Klasickou doménou testů je školství v Severní Americe. Počátek testování se klade do 20. let 20. století v USA. Snahy o racionální řízení výchovně vzdělávacího procesu vedly k potřebě získat pravdivý obraz o úrovni a kvalitě osvojovaných vědomostí.

Pedagogické možnosti školního měření a testování potvrzuje historie i současný stav ve vyspělých zemích. U nás se testuje o poznání méně než např. v USA. Už ve druhé polovině dvacátých let minulého století bylo vydáno v USA ročně 30 – 40 miliónů didaktických a psychologických testů. V šedesátých letech minulého století bylo testováno třikrát ročně kolem pěti miliónů žáků. V 90. letech vydávalo asi 150 nakladatelství a organizací 250 miliónů testů ročně (Hrabal aj., 1994). V USA se provádí celostátní testování a od konce šedesátých let minulého století se pravidelně zjišťují výkony žáků v celých Spojených státech. Didaktické testy se používají i v japonském a západoevropském školství. Trvalé rozšíření didaktických testů neznamena ustrnulé přetrvávání původních forem a pojetí testování. Postupem času se začaly vedle standardizovaných didaktických testů prosazovat i méně formálně propracované, ale flexibilní testy konstruované učitelem, tzv. teacher – made testy, včetně písemných zkoušek testového typu (essay test). Od šedesátých let minulého století se v mnoha školách užívají tzv. kriteriální testy, které sice zachovávají kvantitativní charakter zpracování společný všem testům, ale soustřeďují pozornost především na obsah, zvláště na základní učivo a jeho zvládnutí žákem.

Vedle forem testování se rozvíjely také oblasti, ve kterých se testy používaly. Testování se stalo běžným způsobem zkoušení na středních a vysokých školách, při přijímacích zkouškách na školy, ale také při výběrových řízeních do zaměstnání, a tím spolurozhoduje o budoucím studijním i pracovním uplatnění. Velmi významnou roli mají

didaktické testy při zjišťování stavu vzdělání a úrovně výsledků školy v regionálním, celostátním i mezinárodním měřítku a při ověřování efektivity školských soustav.

Srovnání vývoje a současného stavu testování pro nás nevyznívá příliš příznivě. A to i navzdory relativně brzkému vstupu testování na naši půdu díky profesoru Příhodovi. Používání testů nebylo sice před 2. světovou válkou velmi rozšířené, ale zato bylo intenzivní a tvořivé, především v tzv. reformních školách. Na konci meziválečného období se rozšířilo i na školách využívání výsledků standardizovaných psychologických testů. Po 2. světové válce až do šedesátých a sedmdesátých let minulého století bylo psychologické testování nepřetržitě kritizováno a mlčky trpěno jen v oblasti zdravotnictví s určitým přesahem do škol. Didaktické testy byly zpočátku z ideologických důvodů odmítány, poté s výhradami připuštěny a teprve v sedmdesátých letech minulého století začala vydávat Psychodiagnostika Bratislava standardizované psychologické a didaktické testy, především pro vyšší stupeň základních škol. Pomalu se začaly objevovat na školách, často jen z iniciativy školské správy a nebyly vždy plně využity (Hrabal aj., 1994).

Po roce 1989 se probouzí stále větší zájem o testy a testování v praxi, zvláště v souvislosti s diferenciací, např. s výběrem do soukromých gymnázií a středních škol. Také školská správa a školní inspekce pocítuje potřebu objektivnějšího přístupu k hodnocení škol v souvislosti s nastupujícími organizačními a obsahovými změnami ve školské soustavě. Na cestě potřebnému rozšíření kvalitního testování však dosud stojí mnoho překážek. K hlavním patří nedostatek kvalitních didaktických testů a nedostatečná připravenost učitelů na práci s testy vlastní konstrukce a na adekvátní používání standardizovaných testů. Rozvoj kvalitních testů také brzdí nedostatek finančních prostředků ve školství na jejich nákup. V současné době se tvorbou didaktických testů a testováním zabývá u nás několik organizací, mezi něž patří např. CERMAT, Scio nebo Kalibro.

3. Didaktické testy a evaluace vzdělávacího procesu

Didaktické testy (Achievement tests) nabízejí řadu možností využití. Dají se využít klasicky pro klasifikaci samotných studentů, ale v současnosti jsou v zemích s rozvinutým školským systémem stále častěji používány k celkové evaluaci vzdělávací soustavy a vyučovacího procesu. Pomocí měření dosažené úrovně vědomostí a dovedností žáků se dají odhalit například problémové oblasti vyučované látky a následně navrhnout potřebné změny ve způsobu nebo rozsahu výuky jednotlivých předmětů.

3.1. Instituce a projekty zabývající se hodnocením vyučovacího procesu v zahraničí

Testováním se zabývá celá řada organizací po celém světě. Tyto organizace si uvědomují důležitost testování a jejich následné porovnávání v mezinárodním měřítku. Testování žáků napomáhá k mezinárodní spolupráci jednotlivých zemí. Mezi nejznámější a významné projekty a organizace bezesporu patří NAEP, CITO, IEA, TIMSS, SITES, ACER, PISA, ETS, CKE a PTDE.

3.1.1 NAEP (The National Assessment of Educational Progress – Národní analýza pokroku ve vzdělávání)

The National Assessment of Educational Progress (NAEP), také znám jako "the Nation's Report Card" je již od roku 1969 celonárodní projekt, který se zabývá soustavným sledováním úrovně vědomostí a dovedností studentů. Testování je prováděno pravidelně z čtení, matematiky, přírodních věd, psaní, americké historie, občanské výchovy, geografie a umění. Projekt sleduje výsledky týkající se vědomostí a dovedností populace studentů s ohledem např. na jejich věk, pohlaví, národnost, vzdělání rodičů a bydliště.

NAEP je spravován sekci Ministerstva školství USA, zaměřenou na šetření v oblasti vzdělávacího procesu. Ministerstvo školství také jmenovalo správní radu, která řídí celý projekt.

Projekt je možné rozdělit do několika částí:

- *National NAEP - Celonárodní výzkum* – vyhodnocuje a podává informace o vědomostech a dovednostech studentů na celonárodní úrovni a ve zvláštních geografických regionech země. Výzkum zahrnuje žáky 4., 8. a 12. ročníku ze státních

i soukromých škol. V roce 2007 proběhlo testování ze čtení a matematiky v 4. a 8. ročnicích státních škol a z psaní v 8. a 12. ročnicích státních i soukromých škol.

- State NAEP - Výzkum na úrovni států - vyhodnocuje a podává informace o vědomostech a dovednostech studentů na úrovni jednotlivých států USA. To umožňuje zjištění úrovně všech studentů ve státě a následně jejich porovnávání s ostatními státy, popřípadě s národními výsledky. V roce 2003 proběhl v rámci projektu State NAEP výzkum zaměřený na vědomosti a dovednosti v matematice a čtení ve 4. a 8. ročníku celkem v 53 státech. Na jaře 2006 se testovaly přírodní vědy. Poslední výzkum byl v roce 2007 a testovalo se z matematiky a čtení v 4. a 8. ročnicích. v 8. ročnicích se testovalo také psaní.
- NAEP Trial Urban District Assessment (TUDA) – Výzkum na úrovni městských částí - vyhodnocuje a podává informace o vědomostech a dovednostech studentů na úrovni městských částí. Poprvé se uskutečnil v roce 2002 v oblasti psaní a čtení v pěti městských částech v Atlantě, Chicagu, Houstonu, Los Angeles a New Yorku + Distrikt of Columbia. V roce 2003 se účastnilo 9 městských částí v testování čtení a matematiky. V roce 2005 se testovalo v 10 distriktech čtení, matematika a přírodní vědy v 4. a 8. ročníku. K tomu se ještě zapojily distrikty z roku 2003 + Austin Independent School District, pro ně se ale testovalo pouze ve veřejných školách. Výsledky pro městské části jsou také srovnatelné s výsledky na veřejných školách ve velkých městech a s národními výsledky. TUDA se v roce 2007 prováděl společně se State NAEP ve vybraných oblastech včetně District of Columbia.
- Long - term trend – Sledování dlouhodobých trendů - vyhodnocuje a podává informace o vědomostech a dovednostech studentů v dlouhodobém časovém horizontu. Výzkum se provádí celonárodně jednou za čtyři roky (ne na státní a distriktní úrovni) u studentů ve věku 9, 13 a 17 let v matematice a čtení. Aby se daly zachytit změny ve výkonech studentů, nepodléhají soubory testových úloh této části projektu změnám v důsledku změn v kurikulu nebo pedagogické praxe.
- Special Studies – Speciální studie – zahrnují další aktivity pod záštitou projektu NAEP. Mezi tyto aktivity patří např. High School Transcript Study a Technology - Based Assessment project, které se zaměřují na využití techniky, zejména počítačů ve vyučovacím procesu a sledují jejich vliv na kvalitu vyučování. V roce 2002 začal NAEP novou studii The oral reading study a v roce 2005 zajišťoval National Indian

Education Study (NIES). NIES se v roce 2007 uskutečnil v 4. a 8. ročníku ve školách, které navštěvuje alespoň jeden americký Indián nebo Alaska Native student.

Výzkum programu NAEP v oblasti přírodních věd:

Výzkum NAEP Science prezentuje výsledky, které ukazují, co američtí studenti znají a dovedou využít v oblasti přírodních věd. V této oblasti lze vymezit tři pole zájmu: vědy o Zemi (pevná složka Země, voda, vzduch, Země v prostoru), fyzikální vědy (fyzika a chemie - hmota a její transformace, energie a její transformace, pohyb) a vědy o životě (příroda a funkce živých objektů).

Tab. č.1: Časová dotace na jednotlivé části v NAEP Science v roce 2005 z celkového času testu

	Vědy o Zemi		Fyzikální vědy		Vědy o životě	
	předpoklad [%]	skutečnost [%]	předpoklad [%]	skutečnost [%]	předpoklad [%]	skutečnost [%]
4. ročník	33	33	33	34	33	33
8. ročník	30	30	30	33	40	37
12. ročník	33	34	33	35	33	31

Zdroj: www.nces.ed.gov (28.7.2007)

Analýza se provádí pomocí uzavřených testových položek s vícenásobnou volbou odpovědí (multiple - choice), otevřených položek s krátkou nebo delší odpovědí (short or extended constructed - response questions) a z praktických úloh (hands - on tasks). Zhruba polovina studentů na každé škole vykoná jeden ze tří praktických úkolů a zodpoví otázky, vztahující se k těmto úlohám. Tyto úlohy prověřují, jak dokážou studenti používat získané znalosti v praxi, jak si dokážou věci pospojovat, domýšlet důsledky, využít dodaný materiál a zaznamenat výsledky pozorování a celkové závěry.

Studenti jsou podle výsledků ve výzkumu NAEP charakterizováni jedním ze tří typů hodnotící škály. Tyto stupně jsou označovány jako basic level (základní úroveň), proficient level (středně pokročilý) a advanced level (pokročilý). Znalosti a vědomosti studentů na úrovni basic jsou nezbytné pro pochopení některých základních přírodních zákonitostí. Studenti zvládnou provést jednoduchý pokus a následně zhodnotit jeho průběh. Na úrovni proficient studenti rozumí podstatě průběhu základních životních, fyzikálních a chemických dějů. Své znalosti aplikují na situace z běžného života. Na úrovni advanced student nejen rozumí, ale i dokáže předvídat průběh základních životních, fyzikálních a chemických dějů.

V roce 2005 bylo z přírodních věd testováno více než 300 000 studentů v 4., 8. a 12. ročníku. Procento studentů 4. ročníku, kteří byli na úrovni vyšší než basic, se oproti roku 1996 zvedlo z 63% na 68% v roce 2005. V 8. a 12. ročníku výsledky zůstaly podobné jako v roce 1996 a pohybovaly se okolo 60%.

3.1.2 CITO (Netherlands National Institute for Educational Measurement - Nizozemský institut pro měření výsledků ve vzdělávání)

CITO je institut založený v Nizozemí. Je to jedna ze světových organizací v oboru testování a hodnocení společnosti. Měření a monitorování lidského potenciálu je její hlavní náplní od roku 1968. Pracuje pro školské instituce, vlády a společnosti v oblasti výkonnosti učení, výkonu na pracovišti a vzdělávání. Práce je založena na nejnovějších pokrocích v ICT (informačních a komunikačních technologiích) a psychometrických výzkumech.

CITO byla založena v r.1968 vládou, jako jediná národní instituce pro hodnocení ve vzdělávání v Nizozemí. V roce 1999 byla celkově privatizována. Má dvě pomocné společnosti. Jednu v Německu (Cito Deutschland GmbH) a druhou v USA (Cito USA Inc.). CITO konzultanti pracují dohromady s výzkumníky a analytickými centry ministerstev z třiceti zemí. V těchto zemích pomáhají více zpřístupnit vzdělávání.

Produkty tohoto institutu ukazují cestu vzdělávání na mezinárodním trhu s testy. Vedlejší pobočky v Německu a USA prodávají pyramidové metody a bi-lingvální jazykové testy pro 5 - 7leté žáky. V oblasti měření kompetencí a výkonů souvisejících s prací čerpá ze zkušeností CITO stále větší počet společností a státních institucí. CITO také připravuje testy pro National final examinations v Nizozemí, dále testy pro testování cizích jazyků (audiovizuální testy, testy pro ústní zkoušku, zkoušku čtenářských dovedností...). Na přípravě testů CITO spolupracuje také s mnoha učiteli z praxe. Jednou za 5 let také provádí testování v posledním ročníku základních škol. Testuje se i angličtina, hudba, fyzické vzdělávání, pravidla silničního provozu a komunikativní schopnosti. Testování pomáhá odhadnout skutečné znalosti a schopnosti žáků (www.cito.com, 10.7.2007).

3.1.3 IEA (International Association for Evaluation of Educational Achievement - Mezinárodní organizace pro hodnocení výsledků ve vzdělávání)

IEA je nezávislá mezinárodní organizace složená z národních a vládních výzkumných institucí.

Cílem organizace je:

- poskytovat mezinárodní srovnávací testy, které by poukazyvaly i na slabé stránky vzdělávacích systémů
- zajistit vysoce kvalitní data, která by ukazovala i na mimoškolní faktory ovlivňující vzdělávací proces
- poskytovat vysoce kvalitní data, která mají sloužit pro přípravu a hodnocení reformy vzdělávání
- vyvinout a zlepšit vzdělávací systémy a národní strategie pro hodnocení vzdělávání
- přispět ke vzniku světoznámého společenství výzkumníků v oblasti vzdělávání

IEA vznikla v roce 1956 a od té doby organizovala více než 25 výzkumů v oblasti vzdělávání, které přesahovaly hranice jednotlivých národů. Poskytuje důležitá data z oblasti vzdělávání studentů na lokální, národní a mezinárodní úrovni. Periodicita výzkumů přináší možnost porovnání výsledků jednotlivých zemí, jejich pokroky ve vzdělávání, ale také změny ve vzdělávací politice a výsledky reforem školství. Prováděné studie mají silný empirický základ a kvantitativní orientaci a jsou zaměřeny zejména na průřezové a dlouhodobé projekty se sběrem dat pomocí výběrových metod.

Ve více než 80 státech se provádělo testování z matematiky a přírodních věd (TIMSS 1995, TIMSS 1999, TIMSS 2003, TIMSS 2007), čtenářské gramotnosti (PIRLS 2001, PIRLS 2006), občanské nauky (CIVED, ICCS), informačních technologií ve vzdělávání (SITES-M1, SITES-M2, SITES 2006), vzdělávání učitelů apod.(www.iea.nl, 17.7.2007).

➤ TIMSS (Third International Mathematics and Science study - Třetí mezinárodní výzkum matematického a přírodovědného vzdělávání)

TIMSS se orientuje na komplexní výzkum vzdělávacích výsledků a podmínek v oblasti výuky matematiky a přírodních věd.

Česká republika se zapojila do testování celkem třikrát a to v roce 1995, 1999 a 2007. Studie byla v roce 1995 realizována pro věkovou kategorii devítiletých žáků - 3. a 4. ročník

ZŠ - populace 1, věkovou kategorii třináctiletých žáků - 7. a 8. ročník ZŠ- populace 2 a pro studenty závěrečného ročníku všech středních škol - populace 3 (Straková, Tomášek, Palečková, 1996). V roce 1999 se u nás testovali pouze žáci 8. ročníků ZŠ a odpovídající ročník víceletých gymnázií. Testování se účastnilo 148 škol, u více než 3600 žáků. V roce 2007 se testovalo v populaci 1 a 2. Pilotáže projektu se v roce 2006 zúčastnilo 1350 žáků 4. ročníku z 38 základních škol a 1465 žáků 8. ročníku z 34 základních škol a víceletých gymnázií. Šetření probíhalo v první polovině roku.

Kromě zjišťování výsledků vzdělávání prostřednictvím testů se shromažďovaly i informace o vzdělávacích systémech jednotlivých zemí, analýzy učebních osnov a učebnic. Prostřednictvím dotazníků byly zjišťovány kromě žakovských, učitelských a školních charakteristik také informace o výukových metodách, o podmínkách výuky ve školách a třídách apod.

➤ **SITES (Second Information Technology in Education Study - Druhý mezinárodní výzkum informačních technologií ve vzdělávání)**

Program SITES navazuje na výzkum Počítače ve vzdělávání (COMPED - Computers in Education), který proběhl v roce 1989 - 1992 v 19 školských systémech celého světa. Jeho cílem bylo porovnat situaci v jednotlivých zemích z hlediska vybavení počítačovou technikou, způsobu jejího využití ve školách a problémů, které jsou spojeny s jejím zaváděním do škol a následným využíváním. Sledoval se také přístup ředitelů, učitelů a žáků k výpočetní technice. Česká republika se výzkumu COMPED bohužel nezúčastnila.

Cílem programu SITES je postihnout změny, které nastaly od doby realizace projektu COMPED v oblasti dostupnosti a využívání informačních technologií ve výuce a v připravenosti žáků v jednotlivých zemích. Snaží se také zmapovat propojení výpočetní techniky s ostatními vyučovanými předměty ve školách.

Výzkum má tři fáze:

První fáze spočívala v administraci dotazníků pro ředitele škol a koordinátory výpočetní techniky na vybraných školách v jednotlivých zemích. V dotaznících se zjišťovalo např. vybavení škol výpočetní technikou, rozmístění počítačů, dostupnost elektronické pošty a externích sítí, rozsah a obsah výuky výpočetní techniky, míru integrace informačních technologií (ICT) do vyučovacích předmětů a cíle školy z hlediska výuky výpočetní techniky. První fáze probíhala od roku 1997 - 1999. Do její realizace se zapojilo 25 zemí, u nás se do projektu zapojilo 700 ZŠ a SŠ.

Druhá fáze vyhledávala v jednotlivých zemích zajímavé příklady využívání informačních technologií ve výuce a dokumentuje je s cílem vytvořit jakousi databázi takových příkladů, která by byla dobře využitelná učiteli výpočetní techniky. Tato fáze byla zahájena v roce 1999. V České republice ji koordinuje ing. Bořivoj Brdička z Oddělení informační výchovy Pedagogické fakulty UK Praha.

Třetí fáze bude zaměřena na žáky a jejich učitele výpočetní techniky. Bude zkoumat zejména postoje, znalosti a kompetence žáků a učitelů a dále se bude snažit podrobně popsat rozdíly, které existují ve způsobu používání informačních technologií ve výuce v jednotlivých zemích (www.uiv.cz, 14.7.2007).

3.1.4 ACER (The Australian Council for Educational Research - Australská rada pro pedagogický výzkum)

ACER poskytuje nejmodernější výzkum, produkty a služby ve vzdělávání. Byl založen v r. 1930. Má dlouhou historii a dobrou pověst jako poskytovatel spolehlivé podpory pro vzdělávací politiku a profesionály. Dnes je ACER jedním z hlavních světových vzdělávacích center. Je orientován na vytváření a distribuci znalostí zaměřených na výzkumu, produkty a služby zlepšující učení po formální i neformální stránce.

ACER poskytuje velké množství služeb pro zahraniční klienty. V posledních letech byly programy zaměřené na hodnocení ve vzdělávání zavedeny na Fidži, Malaysii, Indonésii, Filipínách, Hong Kongu, Kambodži, Bhutanu, Srí Lance, Singapuru a v jižním Pacifiku.

ACER je zakládajícím členem Asia Pacific Educational Research Association, která vznikla v Bangkoku v roce 2001, členem International Association for the Evaluation of Educational Achievement (IEA), Australian Consortium for Social and Political Research Incorporated (ACSPRI) a Affiliated Centre of the Asia - Pacific Programme of Educational Innovation for Development (APEID). Na mezinárodních projektech spolupracuje s OECD, AusAID, UNESCO a Světovou Bankou. Mezi projekty, které ACER organizuje patří např. PISA nebo LSAY - Longitudinal Survey of Australian Youth (www.acer.edu.au, 27.4.2007).

➤ PISA (Programme for International Student Assessment - Program pro mezinárodní hodnocení žáků)

Projekt PISA je realizován mezinárodním konsorciem, které je řízeno Australskou radou pro pedagogický výzkum (ACER). PISA provádí výzkumy v oblasti výsledků patnáctiletých žáků v matematice, přírodovědných předmětech a čtenářské gramotnosti.

Testování probíhá jednou za tři roky. Při každém sběru dat je věnován jedné oblasti větší prostor, tak aby o ní bylo možné získat podrobnější informace. První sběr dat v roce 2000 kladl větší důraz na oblast čtenářské gramotnosti, hlavní oblastí v roce 2003 byla matematická gramotnost a v roce 2006 gramotnost přírodovědná. V roce 2000 se v ČR projektu účastnilo zhruba 10 000 žáků z 250 škol, v roce 2003 to bylo 9 919 žáků z 260 škol a v roce 2006 se testování účastnilo 9 016 žáků z 245 škol.

Výsledky se zjišťují pomocí testů, v nichž jsou zařazeny jak úlohy s výběrem ze 4 nebo 5 nabízených možností, tak úlohy s tvorbou odpovědi žákem. V testech se klade důraz zejména na zjištění vědomostí a dovedností, které jsou důležité pro uplatnění žáků v jejich budoucím životě v moderní společnosti. Cílem projektu je také zmapovat mezipředmětové dovednosti prostřednictvím problémových úloh. Tyto úlohy jsou zakotveny v situacích z reálného života, které přesahují situace, se kterými se žáci běžně setkávají ve škole (Straková, 2001).

3.1.5 ETS (Educational Testing Service - Testovací servis ve vzdělávání)

ETS je nezisková organizace v USA, jejíž cílem je zlepšit kvalitu a poctivost ve vzdělávání pro všechny lidi na celém světě. Organizace pomáhá učitelům učit, studentům se učit a rodičům poznávat a měřit vzdělávací a myšlenkové pokroky u svých dětí. ETS za své hodnoty považuje odpovědnost, poctivost a kvalitu. Těchto cílů dosahuje pomocí:

- naslouchání pedagogům, rodičům a kritikům
- zjišťování, co studenti a jejich učitelé potřebují
- vedení ve vývoji nových a inovačních produktů a služeb

Organizace usiluje o zlepšení kvality a poctivosti ve vzdělávání za použití poctivých a platných ustanovení, výzkumu a souvisejících služeb. Její produkty a služby měří znalosti a schopnosti, podporují učení a vzdělávací výkon a podporují vzdělávání a profesionální vývoj pro všechny lidi na celém světě. ETS provádí testování uchazečů o studium na univerzitách v USA již řadu let. Podle výsledků v těchto testech se pak studenti ucházejí o přijetí na vybrané fakulty (www.scio.cz, 27.4.2007).

ETS chce být známá (za předpokladu poctivého a platného ustanovení, výzkumu a souvisejících produktů a služeb) jako světový odborník a výzkumník v pomoci jednotlivcům, rodičům, učitelům, školským institucím, podnikatelům, vládě, zemím, státům a celé oblasti školství (www.ets.org, 21.4.2007).

3.1.6 CKE (Centralna Komisja Egzaminacyjna - Centrální vzdělávací komise)

CKE byla založena ve Varšavě v roce 1999 Ministerstvem Národního Vzdělávání a Sportu, aby připravila a organizovala „Systém externího hodnocení“ (external examination system) ve spolupráci s osmi Okresními vzdělávacími komisemi OKE (Okręgowa Komisja Egzaminacyjna). OKE sídlí např. v Krakově, Gdaňsku, Poznani nebo Lodži a podílí se na přípravě testů, standardů, následných analýz a hodnocení testů.

Systém vznikl postupně od roku 2002 a umožňuje hodnotit úspěchy, ale i neúspěchy vzdělávání žáků a efektivitu vyučovacího procesu. Za cíl si také klade srovnatelnost výstupů z jednotlivých stupňů škol (vysvědčení, diplomy).

External examination system má několik částí: The Competence Test v 6. třídě základních škol, The Lower Secondary School (Gymnasium) Examination ve třetím ročníku nižšího gymnázia, The Matura Exam pro absolventy gymnázií, speciálních středních škol, průmyslových SŠ či nástavbových škol a The Examination confirming Vocational Qualifications (vocational examination) pro absolventy odborných SŠ, průmyslových SŠ a nástavbové školy (upraveno podle www.oke.krakow.pl, 17.7.2007).

CKE organizuje celou řadu mezinárodních konferencí, např. v roce 2001 to byla v Krakově „Tvorba a hodnocení testů a testových položek“ a „Generální zkouška společné části maturit“, v roce 2004 v Krakově „Standards požadavků a testové normy ve vzdělávací diagnostice“ a v roce 2005 v Gdaňsku „Různé metody vzdělávací diagnostiky“.

OKE pracuje za finanční podpory Evropského sociálního fondu EFS na projektu Příprava učitelů na ověřování znalostí. Projekt je zaměřen pro učitele z Lublinu, z Malopolska a Podkarpatské provincie. Je založen na rozvíjení společnosti na základě vědomostí a přizpůsobení vzdělávání potřebám trhu práce. Dále CKE realizovala za podpory EFS např. projekt „Publikace materiálů o Systému externího hodnocení pro žáky“ a „Bádání, vztahující se k výsledkům vnějšího hodnocení“. V roce 2005 byl ukončen projekt „Modernizace počítačových softwarů a hardwarů v Centrálních a Okresních vzdělávacích komisích“.

3.1.7 PTDE (Polskie Towarzystwo Diagnostyki Edukacyjnej - Polská společnost diagnostického vzdělávání)

Sídlem této organizace je Krakov. Jejím cílem je pomáhat v rozvoji systému hodnocení v rámci škol, ale i vnějšího testování; propagovat trendy vzdělávání ve školských institucích; provádět výzkum a diskuze o nových vzdělávacích trendech; spolupracovat s organizacemi a

institucemi podílejícími se na realizaci vzdělávacích trendů a podporovat další vzdělávání členů společnosti.

PTDE pořádá celou řadu konferencí s mnoha referáty a podnětnými diskuzemi. Takové konference se uskutečnily např. v Krakově, Gdaňsku, Lodži, Katovicích a Vratislavi. Tématy mohou být např. etické otázky vzdělávání, vzdělávání v očích žáků nebo přístup učitelů k žákům.

3.2. Projekty zabývající se hodnocením vyučovacího procesu u nás

V posledních více než deseti letech byla i v České republice realizována řada projektů, které se zabývaly konstrukcí didaktických testů, jejich ověřováním a analýzou výsledků získaných jejich praktickým využitím. Mezi ně patří především projekty CERMATu, Scia a Kalibra.

3.2.1 CERMAT

Centrum pro zjišťování výsledků vzdělávání (CZVV, používá se i CERMAT - dnes již neexistující Centrum pro reformu maturitní zkoušky) je organizační složkou státu přímo zřízenou v roce 2006 Ministerstvem školství, mládeže a tělovýchovy (MŠMT). Mezi nejvýznamnější aktivity CZVV patří příprava a realizace nové maturitní zkoušky a projekt hodnocení výsledků vzdělávání žáků 9. tříd a odpovídajících ročníků víceletých gymnázií (řešen v rámci projektu Kvalita 1). CZVV spolupracuje také s partnerskými organizacemi v rámci resortu školství, především s Ústavem pro informace ve vzdělávání (ÚIV), Národním ústavem odborného vzdělávání (NÚOV) a Výzkumným ústavem pedagogickým (VÚP).

➤ Maturitní zkouška

V roce 1999 byl CERMAT pověřen Ministerstvem školství, mládeže a tělovýchovy přípravou reformované maturitní zkoušky. CERMAT do tohoto projektu zahrnul dlouhodobý cyklus programů pod názvem *Krok za krokem k nové maturitě*. Cílem tohoto projektu je pomoc školám v přípravě na změny, které vzniknou zavedením nové maturitní zkoušky. Tato pomoc je školám nabízena každoročně a je zcela dobrovolná a bezplatná. V rámci tohoto cyklu se v roce 2001 uskutečnil projekt *Seznamte se: Nová maturita*, v roce 2002 program *Maturita po internetu* a v roce 2003 program *Maturita nanečisto*. Označení Maturita

nanečisto (MANA, Mana) s označením příslušného roku se ustálilo pro následné programy: Maturita nanečisto 2004, 2005, 2006 a 2007 (www.ceremat.cz, 14.4.2007).

Do programu Maturita nanečisto 2007 (Mana 2007) se zapojilo 1025 škol a 66 029 žáků, přičemž jde o největší účast od roku 2001. Např. přírodovědně technického základu se zúčastnilo 5388 žáků. Původně měla být Mana 2007 generálkou na nové maturity (zkoušky společné i profilové části včetně strukturované písemné části - SPP), ale z finančních důvodů byla realizována minimální varianta s omezenou nabídkou zkoušek. Nicméně velkým pozitivem bylo odzkoušení nového evaluačního nástroje v podobě strukturované písemné části z českého jazyka, cizích jazyků 1 a 2 a polského jazyka pro žáky polské národnostní menšiny. Pro zkoušku cizí jazyk jsou nabízeny testy z anglického, německého, francouzského, ruského, španělského a italského jazyka.

Součástí cyklu Krok za krokem k nové maturitě jsou také od školního roku 2001/2002 přípravné programy na všech speciálních středních školách s maturitními ročníky a na školách běžného typu, v nichž jsou integrováni žáci se speciálními vzdělávacími potřebami. Programy nesou název *Maturita bez handicapu*: Maturita bez handicapu 2002, 2003, 2004, 2005, 2006. Pro rok 2007 je tento speciální přípravný program organicky včleněn do programu Maturita nanečisto 2007 (www.ceremat.cz, 14.4.2007).

Nová maturita má mít dvě části - společnou a profilovou.

Společná část maturitní zkoušky se skládá z českého jazyka, cizího jazyka a z volitelné zkoušky. Má odrážet základní okruhy všeobecných požadavků státu na znalosti a dovednosti maturantů. Volitelnou zkoušku žák skládá podle své volby buď z matematiky, občanského základu, přírodovědně technického základu nebo informačně technického základu. Škola je povinna zajistit podle svého rozhodnutí přípravu žáka na nejméně dvě tyto oblasti. Zkouška z českého a cizího jazyka je formou písemnou i ústní, volitelná zkouška pouze písemnou.

Tab. č.2: Struktura zkoušek zadávaných MŠMT ve společné části maturitní zkoušky

Zkouška	Didaktický test		Poslechový test	Strukturovaná písemná práce	Ústní zkouška
	60 minut	90 minut	35 - 40 minut	60 min.	15 minut
Český jazyk	U			E	ano
Cizí jazyk	U		U+O	E	ano
Volitelná zkouška		U+O			

Zdroj: www.ceremat.cz (14.4.2007)

Pozn. U - uzavřené úlohy; O - otevřené úlohy se stručnou odpovědí; E - úlohy se širokou odpovědí

Jednou ze čtyř volitelných zkoušek ve společné části maturity je přírodovědně technický základ. Obsah této zkoušky zasahuje do biologie, chemie, fyziky, geografie a zároveň obsahuje i technické prvky. Cílem zkoušky není ověřovat znalosti a dovednosti z jednotlivých oborů do hloubky, ale spíše ověřit, nakolik student dokáže pracovat s dostupnými informacemi z této oblasti a využít je při řešení problémů. Tato zkouška je koncipována na integrovaném pohledu na svět přírody a techniky a zatím nemá oporu v komplexně pojatém vyučovacím předmětu.

MŠMT bude každý rok vymezovat obsah a podmínky konání jednotlivých zkoušek, které zadává v rámci maturity, formou *katalogů požadavků*. Katalog také obsahuje popis zkoušky a příklady úloh, které mohou být do zkoušky zařazeny.

Požadavky, zařazené do katalogu pro biologii, vycházejí z platných pedagogických dokumentů: Učební dokumenty pro gymnázia (učební osnovy a plány). Schválilo MŠMT s platností od 1.9.1999. Praha, Fortuna, 1999 a Standard vzdělávání ve čtyřletém gymnáziu. Schválilo MŠMT s platností od 12.2.1996. Praha, Fortuna, 1999. Podkladem pro formulaci zkoušky byly také současně používané učebnice pro gymnázia, neboť poskytují komplexní pojetí středoškolského učiva biologie. Biologie je zde brána z celkového pohledu a neklade důraz na vybraná témata, jako to bývá u odborných škol. Zároveň je zde zahrnuta i skutečnost, že na některých středních školách jsou již ověřovány rámcové vzdělávací programy (RVP). Katalog vymezuje znalosti a dovednosti obsažené v maturitních požadavcích tak, aby si je mohli osvojit studenti bez ohledu na typ navštěvované školy i na druh programového dokumentu, z něhož vychází studijní plán dané školy. Při vymezování maturitních požadavků bylo bráno v potaz, že zkouška z biologie zadávaná MŠMT může sloužit také jako podklad při přijímacím řízení na vysoké školy. Předpokládá se, že této maturitní zkoušky se budou účastnit studenti, kteří mají o biologii zájem a hodlají se jí ve svém studiu na vysoké škole dále zabývat (www.ceremat.cz, 14.4.2007).

V rámci profilové části maturity jsou v katalogu pro biologii uvedeny i cílové kompetence, které jsou rozděleny do tří kategorií (převzato z www.ceremat.cz, 14.4.2007):

Znalost s porozuměním

Žák dovede:

- používat správnou terminologii, symboly a značení
- definovat základní biologické pojmy
- popsat jev, objekt, systém podle skutečnosti, modelu, nákresu
- vysvětlit základní biologické zákony, definice a teorie

- orientovat se v přirozených systémech organismů a posuzovat fylogenetické vztahy
- poznat a pojmenovat charakteristické představitele taxonů podle uvedených charakteristik

Aplikace znalostí a řešení problémů

Žák dovede:

- aplikovat teoretické biologické poznatky při řešení konkrétních životních situací nebo při řešení problémových úloh
- posoudit důsledky určitého jevu nebo lidské činnosti z ekologického, ekonomického nebo zdravotního hlediska
- využívat při řešení biologických problémů poznatků z dalších oborů, především z chemie, fyziky, zeměpisu a matematiky
- využívat biologické znalosti pro pochopení moderních technologií
- zdůvodnit význam nových biologických poznatků pro společnost - zdravotnictví, šlechtitelství, různá průmyslová odvětví a další aplikované obory

Práce s informacemi

Žák dovede:

- vyhledat, vybrat, uspořádat a prezentovat informace z různých zdrojů (text, graf, tabulka)
- klasifikovat a kategorizovat biologické objekty a jevy dle rozlišovacích znaků
- vybírat z nadbytku biologických informací podstatné
- převést informace z jedné formy do druhé a naopak (text, graf, tabulka)
- pracovat s určovacími klíči přírodnin
- určit na základě obrázků, mikrofotografií či schémat daný organismus či jeho část
- vyhodnotit a vyvodit závěry z údajů v tabulkách a grafech
- zapsat, tabelárně a graficky zpracovat a synteticky vyhodnotit empirické údaje získané při řešení biologických úloh

Profilová část se koná ze tří povinných zkoušek, které zohledňují hlavně odborný profil škol a zároveň profesní a vzdělávací záměry žáků. Žák může konat ještě nejvýše 4 nepovinné zkoušky. Ty si může volit z důvodu větších šancí na trhu práce nebo z důvodu možného zohlednění jejich výsledků v přijímacím řízení na vysoké školy. Zkoušky se konají formou:

- vypracování maturitní práce a její obhajoby před zkušební maturitní komisí
- ústní zkoušky před zkušební maturitní komisí

- c) písemné zkoušky
- d) praktické zkoušky
- e) kombinací dvou nebo více forem

Profilovou i společnou část koná žák ve škole, jejíž je žákem, nepovinné zkoušky zadávané ministerstvem koná žák ve škole stanovené CZVV.

Tab. č.3: Příklad možné struktury maturitní zkoušky jednotlivého žáka pro gymnázium se všeobecným zaměřením (maturant/ka podává přihlášky na různé typy VŠ)

Společná část	český jazyk
	anglický jazyk
	přírodovědně technický základ
Profilová část - výběr ředitele	anglický jazyk/ německý jazyk/ ruský jazyk na rozšířené úrovni (kombinace písemné a ústní zkoušky)
	dějepis/ základy společenských věd/ zeměpis/ dějiny umění (obhajoba maturitní práce)
	biologie/ fyzika/ chemie/ matematika (obhajoba maturitní práce)
Profilová část - výběr žáka	matematika na rozšířené úrovni (z nabídky ministerstva)
	občanský a společenskovední základ (z nabídky ministerstva)
	filozofie (písemná zkouška)

Zdroj: www.cermat.cz (14.4.2007)

Realizace Nové maturity již byla několikrát odsunuta (z roku 2002 na 2004) a poprvé se měla uskutečnit ve školním roce 2007/2008. Poslanecká sněmovna Parlamentu ČR zavedení Nové maturity opět odložila na rok 2010, kdy by se měla poprvé konat maturita z českého jazyka a cizího jazyka. Volitelnou zkoušku by měli studenti skládat až od roku 2012. Novela byla již schválena senátem a podepsaná prezidentem republiky. CZVV by mělo zajišťovat společnou část maturitní zkoušky a zkoušky zadávané ministerstvem v rámci profilové části maturitní zkoušky.

➤ **Kvalita 1**

Systémový projekt **Kvalita 1** je projekt Evropského sociálního fondu (ESF), jehož cílem je postupné vytvoření nového komplexního systému monitorování a hodnocení výsledků základního a středního stupně vzdělávání. Významnou měrou se na něm od května 2005 podílí i CERMAT. Celý projekt připravilo Ministerstvo školství, mládeže a tělovýchovy (MŠMT) ve spolupráci s NÚOV a ÚIV. Nutnost vytvoření tohoto systému hodnocení se uvádí ve školských dokumentech a doporučuje se i v mezinárodních expertních hodnoceních české vzdělávací soustavy. Tento projekt by měl přinést do českého školství větší objektivitu, školám poskytnout možnost vzájemného srovnání, žákům možnost srovnání se stejně starými

děťmi a ředitelům informace o výsledcích jejich škol. Cílem je také zkvalitnění vzdělávání a podpora kurikulární reformy.

Oblasti hodnocení vzdělávání jsou ve třech úrovních, v nichž se jednotlivé organizace angažují různě:

CERMAT:

- na úrovni žáka je vytvářen systém evaluace výsledků žáků v uzlových bodech jejich vzdělávání

Zdrojem informací jsou:

- výsledky *národní hodnotící zkoušky*, která je realizovaná formou didaktických testů na konci základního vzdělávání

- *osobní portfolio žáka*, což je souhrn informací vzdělávacích procesů a výsledky žáka, přičemž se využívá různorodých prací žáka a jejich hodnocení

Projekt také zjišťuje, vytváří a ověřuje *podmínky ovlivňující zavedení nového modelu maturitní zkoušky* a to zavedení nových forem konání maturitní zkoušky, genderové aspekty a ověřování úrovně znalostí a dovedností žáka v cizích jazycích.

- na úrovni školy je hlavní aktivita zaměřena na tvorbu nástrojů pro evaluaci a autoevaluaci výsledků vzdělávání. Pomocí ověřovacích didaktických testů, vycházejících z rámcových vzdělávacích programů, budou základní a střední školy získávat informace o úrovni znalostí a dovedností žáků, což povede ke zkvalitnění výuky.
- podpůrný systém projektu zahrnuje
 - podporu vzdělávání pedagogických pracovníků ve využívání systému evaluace a autoevaluace, v souvislosti s reformou maturitní zkoušky a s hodnocením v oblasti základního vzdělávání
 - distribuci výsledků CERMATu mezi odbornou i širší veřejnost

ÚIV:

- na úrovni vzdělávacího systému jde především o:
 - využití výsledků výzkumu PISA pro podporu práce učitelů základních škol, především českého jazyka a matematiky (hodnocení úrovně vzdělávání v ČR s ohledem na krajevou diferenciaci - krajská PISA)
 - zjišťování a hodnocení úrovně vzdělávání ve vzdělávacím systému a jeho částech. Zejména o hodnocení úrovně vzdělání v jednotlivých krajích ČR (profily krajů) a o

české školství ve světle dlouhodobě zjišťovaných výsledků vzdělávání v mezinárodních šetřeních.

- vytvoření systému statistických ukazatelů pro hodnocení vzdělávacího systému

NÚOV:

- na úrovni žáka jde o nové pojetí *závěrečné zkoušky*, která má přinést větší srovnatelnost výsledků, zvýšení prestiže absolventů učebních oborů a jejich lepší uplatnitelnost na trhu práce. Zadání musí odpovídat standardům vytvořeným v Národní soustavě kvalifikací, kterou spoluvytváří zaměstnavatelé. Závěrečná zkouška se koná z písemné a ústní části a z praktické zkoušky. Každé téma při ústní části obsahuje také podotázku ze světa práce u nás i možností absolventů uplatnit se v zahraničí (www.esf-kvalita1.cz, www.kvalita1.cz, www.uiv.cz, 14.7.2007).

V rámci projektu se testují 5. a 9.třídy a odpovídající třídy víceletých gymnázií. Žáci jsou testováni z dovedností v českém jazyce, matematice a obecných dovedností. Součástí testování je i vyplnění dotazníku. V únoru 2007 testování proběhlo ve všech krajích ČR (kromě Prahy). Testuje se i ve školách s polským vyučovacím jazykem, mohou se ho zúčastnit i žáci se speciálními vzdělávacími potřebami (SCP), kteří jsou integrováni v běžných nebo speciálních školách. Projekt ale není určen pro žáky s mentálním postižením, kteří se vzdělávají ve třídách nebo školách s upraveným vzdělávacím programem. Testuje se testy s výběrem právě jedny správné odpovědi.

3.2.2 Scio

Společnost Scio působí na poli českého školství a vzdělávání již od roku 1996. Od té doby se rozšířily její aktivity a s nimi se rozrostl také odborný záběr v této oblasti. Mezi projekty Scia patří především Srovnávací zkoušky, Národní srovnávací zkoušky (NSZ), Porovnávací testy pro žáky 5. a 9. tříd, Srovnávací testy 2. ročníků SŠ a Vektor.

➤ Srovnávací zkoušky

Společnost Scio organizovala v roce 1996 *Srovnávací zkoušky* pro gymnázia. Studenti byli testováni z českého jazyka, matematiky, biologie a obecných studijních předpokladů. Výsledky šetření, které zaslalo 43 gymnázií, byly použity jako pilotní verze pro sestavování *Srovnávacích testů maturantů*. V roce 1997 společnost Scio nabízela již všem typům

středních škol soubor 23 testů ze všeobecně vzdělávacích předmětů, a to variantu pro žáky z daného předmětu maturující i nematurující (www.scio.cz, 27.4.2007).

➤ **Národní srovnávací zkoušky (NSZ)**

Scio již od roku 1996 spolupracuje se středními i vysokými školami na testech pro přijímací řízení. Před jedenácti lety myšlenka Národních srovnávacích zkoušek jako přijímacích zkoušek na vysoké školy vzhledem k negativnímu postoji fakult neuspěla. NSZ tedy po tuto dobu fungovaly jako „přijímačky nanečisto“ a jejich výsledky byly řadou fakult zohledňovány vedle výsledků přijímacích zkoušek. V roce 2005 zkoušky s testy Scio absolvovalo přes 30 000 uchazečů o studium na středních školách. Scio organizovalo také průběh nebo dodávalo testy pro přijímací řízení na některé vysoké školy (Právnická fakulta UK Praha, Biologická fakulta Jihočeské univerzity v Českých Budějovicích,...). V akademickém roce 2006/2007 celkem 28 fakult českých vysokých škol a univerzit přijímalo uchazeče na základě výsledků *Národních srovnávacích zkoušek*, které jsou inspirovány modelem praktikovaným již desítky let v USA.

Na některých vysokých školách se na výsledky Národních srovnávacích zkoušek (NSZ) přihlíží, na některých zcela nahrazují přijímací zkoušky. Ve školním roce 2006/2007 mohli účastníci skládat NSZ z obecných studijních předpokladů, základů společenských věd, anglického/německého jazyka a matematiky. V minulých letech se konaly NSZ i z českého jazyka a přírodních věd. Formou zkoušky je test s výběrem odpovědí, kde správná je vždy právě jedna. Velký počet úloh a skutečnost, že testy jsou organizovány na všech místech ve stejnou chvíli se stejným zadáním výrazně ovlivňuje vliv náhody na výsledek.

NSZ se konají i pro žáky základních škol a to pro 5., 7. a 9. třídu. Některé střední školy zohledňují výsledky NSZ u přijímacích zkoušek. Tyto testy se skládají z českého jazyka, matematiky a obecných studijních předpokladů. Žáci vypracovávají test s s výběrem odpovědí, kdy právě jedna je správná.

➤ **Porovnávací testy pro žáky 5. a 9. tříd**

Scio vytváří také *Porovnávací testy pro žáky 5. a 9. tříd*. Tyto testy se skládají z českého jazyka, matematiky a obecných studijních předpokladů. Žáci 9. tříd dále mohou vypracovávat testy také z anglického jazyka a biologie.

➤ **Srovnávací testování 2. ročníků SŠ**

Scio se také zabývá *Srovnávacím testováním 2. ročníků SŠ*. Srovnává se český jazyk, matematika a obecné studijní předpoklady. Rozšířená verze zahrnuje i testy z anglického a německého jazyka. Tyto testy umožňují zhodnotit v poločase čtyřletého studijního cyklu, jaké jsou vědomosti studentů ve vztahu k jejich vstupní úrovni, a zároveň u každého studenta posoudit předpoklady k úspěšnému zvládnutí státní maturitní zkoušky. Všechny testy jsou sestaveny z úloh s výběrem jediné správné odpovědi.

➤ **Vektor**

Dalším projektem je projekt *Vektor*, který zajišťuje středním školám základ dlouhodobé (auto)evaluace. Žáci jsou v něm testováni na začátku a na konci studia na SŠ. To umožňuje stanovit, co škola žákům za dobu studia dala. Projekt má 3 moduly (modul 1, 3 a M). Modul 1 je určen pro 1. ročník středních škol a skládá se ze dvou sad. Základní sada obsahuje dotazník a testy z obecných studijních předpokladů (OSP), českého jazyka, anglického a německého jazyka a matematiky. Rozšířená sada obsahuje testy základní sady + testy z humanitního základu (dějepis, základy společenských věd, zeměpis), přírodovědného základu (fyzika, chemie, biologie) a informatiky. Modul 3 je určen pro žáky 3. ročníků středních škol. Testování je stejné jako v Modulu 1. Modul M je maturita. Tento model však není momentálně realizovatelný, z důvodu odložení státních maturit.

3.2.3 Kalibro

Projekt Kalibro je organizovaný Skupinou pro vzdělávací alternativy IDEA. Nabízí srovnávací testy z českého jazyka, matematiky a dvou cizích jazyků pro střední i základní školy. Pro základní školy a třídy nižšího gymnázia dále ještě test širšího humanitního základu a širšího přírodovědného základu. Pomocí srovnávacích testů tohoto projektu jsou získávána reálná měřítka výsledků vzdělávání a mohou být poskytována dále školám.

Testy jsou určeny pro jednoho žáka na jednu vyučovací hodinu. Nejsou zaměřeny na vědomosti, ale na praktické použití základních znalostí v nových situacích. V testech převažují uzavřené otázky s výběrem odpovědí, kde správných odpovědí může být nula až všechny. Tím se eliminuje efekt hádání a používání tzv. vylučovací metody. Žák je tak donucen se nad otázkou opravdu zamyslet.

Dále Kalibro využívá Dovednostních testů, kde žáci nevybírají odpovědi na otázky z nabídky, ale formulují je vlastními slovy. Hledají příklady či protipříklady, zapisují postup

úvah nebo výpočtu, uvádějí argumenty ap. Při řešení tak uplatňují klíčové dovednosti potřebné v reálném životě. K těmto testům náleží ještě podrobná pravidla jejich objektivního hodnocení, řada ukázek žákovských odpovědí a souhrnné výsledky za ČR.

4. Didaktické testy

Didaktický test je nástroj systematického zjišťování (měření) výsledků výuky (Byčkovský, 1982). Je zaměřen na objektivní zjišťování úrovně zvládnutí obsahu učiva ve školních předmětech, na vysokých školách v jednotlivých vědních oborech a jejich strukturálních komponentách. Měl by být sestaven tak, aby nemohlo dojít k subjektivnímu zkreslení výsledků ze strany učitele a výsledky byly jednoznačné, komplexní a objektivní.

Kvalitní test musí být:

- objektivní - výsledek testování je plně závislý na testovaném studentovi a na testu; na testujícím je nezávislý
- validní – platný - plní ty požadavky, pro které byl vytvořen a použit
- reliabilní - přesný a spolehlivý; výsledky testu jsou stabilní i při vícenásobném zadání testu; hodnota reliability testu by měla být uváděna v každém manuálu testu a měla by být vyšší než 0,85
- ekonomický - z hlediska času testovaného a testujícího; i pokud jde o vynaložené finanční prostředky v porovnání s kvalitou a množstvím získaných informací (Hrabal aj., 1994)

4.1 Didaktické testy a testování v zahraniční literatuře

Ve světové literatuře se problematikou testování a tvorby didaktických testů zabývá velké množství autorů, k nimž patří např. S. Morrison, P. Smith, R. Britt, D. Boyda a A. Murray (1996), S. J. Osterlind, T. M. Haladyna, S. H. Irvine, P.C. Kyllonen.

Největší pozornost se soustředí především na tvorbu testových položek pro standardizované testy. Autoři těchto publikací se většinou shodují, že tvorba otázek pro standardizované testy, které měří výkon, schopnosti a nadání studentů, je úkol velmi významný, a že možnost interpretace výsledků testů vyplývá z kvality testových položek. Primárním předpokladem pro snížení výskytu nechtěných odchylek či chyb měření je pečlivé vytvoření testových položek. Díky tomu roste spolehlivost výsledků testů.

Haladyna (1997) a Osterlind (1998) seznamují se základy pro psaní testových položek. Uvádějí také, co vlastně testovými položkami měřit. Zmiňují se tedy o jejich validitě a porovnávají otevřené a uzavřené položky. Dále popisují jednotlivé typy testových položek, jejich tvorbu, výhody a nevýhody.

V oblasti testování jsou významné zejména ty práce, které se zaměřují na vyšší úroveň vnímání, schopností a dovedností a nezabývají se pouze ověřováním znalostí. Tomuto problému se věnuje Haladyna (1997,1999) a také Morrison aj. (1996) či Irvine a Kyllonen (2001). Například Haladyna (1997) předkládá určitý návod na testování motivace, postojů, sebedůvěry a dalších součástí citové oblasti vnímání. V souvislosti s touto problematikou můžeme najít také mnoho forem testových položek, které se vztahují právě k ověřování vyšších úrovní myšlenkových činností. Jsou to např. soubory otázek vyplývajících z kontextu (context - dependent item sets), položky s vícenásobnou volbou odpovědi postihující tuto oblast, portfolia atd.

Na závěr se dá konstatovat, že práce v oblasti testování se dají rozdělit do dvou skupin. První skupinou jsou spíše teoretické publikace, které se zabývají testováním v celé jeho šíři a tvorbou testových položek na teoretické úrovni. Takto se postupuje i v práci od Irvine a Kyllonena (2001), která obsahuje různé modely tvorby položek, a také se zabývá postupy, jak od teorie přecházet k praktickému využití. Druhou skupinu tvoří práce zaměřené na praktickou pomoc tvůrcům a uživatelům testů a testových položek. Ty se tedy snaží o zlepšení vědomostí o testování. Také udílejí rady, jak vybrat nejlepší typ testové položky pro ten který účel a ukazují velké množství příkladů.

4.2 Základní typy didaktických testů

V pedagogické praxi se používají didaktické testy nejrůznějšího druhu, formy i kvality. Je možno je třídit podle různých kritérií. Tabulka č. 4 uvádí klasifikace od Byčkovského na základě Niemierka.

Největší význam pro praxi má třídění podle stupně dokonalosti testu a jeho příslušenství. Podle tohoto kritéria označujeme didaktické testy důkladněji připravené s úplnější výbavou jako testy standardizované. Didaktické testy, u nichž nebyly důsledně realizovány všechny kroky obvyklé při přípravě a ověřování testů standardizovaných, označujeme jako testy nestandardizované - učitelské, neformální (Obst, 2006). Učitelé si sami pro svoji individuální potřebu připravují tyto nestandardizované testy. Někdy se užívá termínu testy kvazistandardizované, čímž se rozumí testy připravované dokonaleji než testy učitelské, u nichž ale standardizace nebyla provedena beze zbytku. Tímto testem je např. didaktický test, zjišťující úroveň vědomostí žáků v daném předmětu na určité škole (několik paralelních tříd) nebo na několika školách (Chráška, 1999).

Tab. č.4: Klasifikace didaktických testů na základě Niemierka (Byčkovský, 1982)

Klasifikace didaktických testů			
Klasifikační hlediska	Testy		
Měřená charakteristika výkonu	rychlosti; úrovně		
Dokonalost přípravy a vybavení testu	standardizované	kvazistandardizované	nestandardizované
Povaha činnosti testovaného	kognitivní; psychomotorické		
Míra specifčnosti učení	výsledků výuky; studijních předpokladů		
Interpretace výkonu	rozlišující (relativního výkonu), ověřující ("absolutního výkonu")		
Časové zařazení do výuky	vstupní	průběžné (formativní)	výstupní (sumativní)
Rozsah obsahového zaměření	monotematické; polytematické (souhrnné)		
Míra objektivnosti skórování	objektivně	kvaziobjektivně	subjektivně

Zdroj: CHRÁSKA (1999), HRABAL aj. (1994)

Technicky nejpropracovanějším a původním typem testu je *standardizovaný na normy orientovaný didaktický test* (NR - norm referenced test). Pro měření ve škole se postupem času vyvinul *kriteriální test* (CR - criterion referenced test), který se s NR testem vzájemně doplňuje. Většina principů a technik standardizovaného testu se v pozměněné formě uplatňuje i v *písemných zkouškách testového typu - esej test* (E test).

Centrum pro zjišťování výsledků vzdělávání (CERMAT) rozlišuje didaktické testy podle ověřovaného výsledku učení, měřené charakteristiky výkonu, formy zadání a interpretace testových výsledků (viz. Tab. č. 5).

Velmi důležitým kritériem z hlediska vytváření didaktického testu je způsob interpretace testových výsledků. Ten je podstatný nejen v závěrečných fázích používání didaktického testu, ale je velmi důležitý už při volbě postupu konstrukce testu, včetně volby obsahu a typů užitých testových položek. V praxi pedagogové často funkce rozlišujících a ověřujících testů kombinují.

Tab. č. 5: Možné dělení didaktických testů (upraveno podle CERMATu)

Kritérium třídění	Testy	Popis
Forma zadání	Zadané na papíře	tištěné; text, graf
	Zadané ústně	ústně, audiozáznam
	Speciální	na PC
	Kombinované	kombinace např. písemného a ústního zadání
Ověřovaný výsledek učení	Kognitivní	ověřuje znalosti a intelektové dovednosti
	Psychomotorické	ověřuje psychomotorické dovednosti
Měřená charakteristika výkonu	Rychlost	úkoly méně komplexní, počet překračuje možnosti dané časovým limitem
	Úroveň	úkoly náročnější, čas na řešení je dostatečný
Interpretace výsledků	Rozlišující	cílem je vzájemně porovnat výsledky žáků
	Ověřující	cílem porovnat výsledky žáků s předem stanovenými kritérii

Zdroj: www.ceremat.cz (14.4.2007)

Na normy orientovaný (NR) test

Standardizované testy (NR testy, testy rozlišující, testy relativního výkonu) se používají ke zjištění individuálních rozdílů v míře měřeného znaku. Vydávají je většinou specializované instituce. Součástí příslušenství je testová příručka (manuál), která informuje o vlastnostech testu a o jeho správném používání. Většinou je také k dispozici standard (testová norma) pro hodnocení výsledků žáků (Obst, 2006). K základním idejím tohoto testu patří snaha dosáhnout maximální možné objektivitu a diferencovanosti hodnocení testového výkonu žáka. Výkon jednotlivého žáka se srovnává s výkonem populace – všech žáků stejného věku nebo ve stejném školním ročníku. Populaci představuje reprezentativní vzorek (obvykle 400 – 500) žáků, se stejnými základními charakteristikami jako populace. Objektivitu se dosahuje tím, že mají všichni testovaní žáci stejné podmínky při testování. Žáci jsou stejného věku, ze stejného ročníku téhož typu školy. Také jim jsou předkládány testy s jednoznačnými položkami a jednotnou instrukcí. Na vypracování testu mají všichni stejnou časovou dotaci a testy jsou hodnoceny jednotným způsobem. NR testy se často používají u přijímacích zkoušek na vysoké školy.

Kriteriální (CR) test

Kriteriální testy (testy ověřující, testy absolutního výkonu) se vyvinuly do určité míry jako kritická odpověď na NR testy. Ty neumožňují samy o sobě získat informaci o stupni zvládnutí učební látky jednotlivcem a nebo skupinou. Můžeme srovnávat a diferencovat jak všeobecně slabé, tak i velmi dobré testové výkony populace.

Kritériem úspěchu u CR testů je obsah výuky, učivo tj. předem stanovený stupeň jeho zvládnutí. Hodnocení výsledků testu vyjadřuje, zda žák zná to, co znát má a musí nebo nikoliv. Rozhodnutí o učivu nutnému pro bezpodmínečné zvládnutí a z toho plynoucí výběr požadavků a tvorby položek jsou kritickými body konstrukce CR testů.

Základní kroky konstrukce tohoto testu jsou obdobné jako u NR testu, ale s pozměněným důrazem na jednotlivé etapy (např. chybí tvorba norem) a kvantifikace nemá tak dominantní postavení. Aby mohl tento test bezpečně zjistit ovládnutí vybraných jevů, je třeba každý testovaný jev pokrýt větším počtem položek. Analýzou otázek a odpovědí při opakovaném testování téhož jevu v průběhu výuky se zjišťuje i citlivost položek na učení.

Tyto testy je vhodné a výhodné použít v předmětech, kde je hierarchizované učivo, kde další postup závisí na zvládnutí předchozího kroku, např. v přírodních vědách, matematice, gramatice jazyků, na základní škole v čtení, psaní a počtech. Ověřující testy jsou velmi důležitou částí vyučovacího postupu u učitelů a ve školách, které se řídí principy "mastery learning" – to je učení až do úplného ovládnutí základního učiva.

Esej test (ET)

Esej test je písemná zkouška testového typu s menším počtem širších otázek, na které žáci odpovídají vlastními formulacemi - „svými slovy“ (Hrabal aj., 1994). E-testy jsou považovány za kompromis mezi běžnou písemnou zkouškou a objektivním testem. Je možné je také považovat za testy s širokými otázkami nebo za písemné zkoušky testového typu. Tyto testy poskytují větší prostor pro uplatnění individuality žáka, širší možnost aktivizovat a hodnotit náročnější myšlenkové operace, uplatnit a rozvíjet dovednosti (zejména formulační a stylistické). Zachovávají výhody běžné písemné zkoušky a na druhé straně umožňují výrazně objektivnější hodnocení (aplikací obecných testových pravidel a postupů v E - testech se zmírňují slabiny běžných písemných zkoušek).

E-test je vhodné používat hlavně u předmětů a látky, která vyžaduje a umožňuje uplatnění spíše vyšších kognitivních než pamětních dispozic. Jako typický učitelský test klade v porovnání se standardizovaným testem menší požadavky na finance.

4.3. Základní typy testových položek

Základem každého didaktického testu je testová položka (otázka, úloha). Klasifikaci testových položek uvádí např. Hniličková aj. (1972), Půlpán (1991) nebo Hrabal aj. (1994).

Chráska (1999) a Centrum pro zjišťování výsledků ve vzdělávání (CERMAT) rozlišují testové položky na otevřené a uzavřené. Uzavřené položky rozdělují na úlohy s dvoučlennou volbou (dichotomické), s výběrem z více odpovědí, přiřazovací a uspořádací. Otevřené dále dělí na úlohy se širokou odpovědí (nestrukturované; se strukturou - vymezenou nebo danou konvencí) a se stručnou odpovědí (produkční; doplňovací).

4.3.1 Uzavřené položky

V uzavřených položkách je žákovi nabízeno několik alternativ, z nichž žák vybírá jednu nebo více správných. Následující charakteristiky uzavřených i otevřených testových položek jsou upraveny podle www.cermat.cz (14.4.2007), Hniličková aj. (1972), Hrabal aj. (1994) a Půlpán (1991).

➤ Úlohy s dvoučlennou volbou (dichotomické)

Dichotomické testové položky jsou často označovány také jako Ano - Ne položky, alternativní úlohy nebo true - false items. Žákovi jsou předkládány dvě alternativy odpovědi s tím, že jedna je správná a tu má označit (např. podtržením, zakroužkováním,...).

Úlohy tohoto typu kladou velký důraz na memorování. Mají obrovskou pravděpodobnost uhádnutí (50%), proto se užívají značně omezeně. Aby se spolehlivost výsledků získaných testem s dichotomickými úlohami zvýšila, je potřeba, aby tento test obsahoval dostatečný počet těchto úloh. Proto bývají dichotomické úlohy řazeny do svazků složených z více jednotlivých dichotomických podúloh. Svazky úloh s dvoučlennou volbou odpovědi se také někdy používají ve vazbě na výchozí text, v němž žák může nalézt potřebné informace k správným odpovědím.

Jejich vytvoření se jeví rychlé a snadné. Při tvorbě těchto úloh se musí dbát na to, aby každé tvrzení mělo jen jednu centrální myšlenku a bylo formulováno jednoznačně. Tvůrce by se měl vyvarovat také používání negativních otázek, dvojitě negace a jiných logicky komplikovaných konstrukcí.

➤ **Uzavřené položky s výběrem z více odpovědí**

Tyto položky jsou nejčastěji používané, v zahraničí se označují jako Multiple Choice Items (položky mnohonásobné volby, s mnohonásobným výběrem). Obsahují *jádro* (kmen), ve kterém je definován problém a více než dvě *alternativy*, které představují jeho možné řešení. Alternativy jsou většinou vyjmenovány, ale mohou jimi být i obrázky, či grafické symboly. Správně může být jedno řešení, více ale i žádné. Nesprávná řešení se nazývají *distraktory*.

Tyto položky mají dvě základní podoby:

- ukončená otázka - kmen je běžnou otázkou a alternativy jsou možné odpovědi.
- nedokončená věta - kmen je nedokončenou větou a nabízené alternativy tvoří různá možná zakončení. Žák vybere vhodnou alternativu tak, aby tvrzení bylo správné a pravdivé.

Pravděpodobnost, že žák vyřeší správně celý test složený z těchto položek hádáním je velmi malá. Šance uhádnutí se žákovi snižuje se zvětšujícím se počtem odpovědí. Test složený z těchto položek je závislý na přehlednosti, čase, srozumitelnosti a čitelnosti textu. Pravděpodobnost, že žák správnou odpověď pouze uhádne je při klasickém počtu alternativ (4) poměrně vysoká a to 25%.

➤ **Úlohy přiřazovací**

Přiřazovací úlohy (angl. Matching Items) obsahují dva seznamy pojmů a instrukci. Žák má správně v těchto seznamech najít odpovídající si dvojice.

Počet pojmů v pravém sloupci je úmyslně vyšší než počet pojmů v levém sloupci. Tato zásada by se měla u tohoto typu úloh dodržovat vždy, aby nedošlo k tomu, že žák bude mít usnadněnou situaci tím, že zná některé z nabízených přiřazení. Počet možných přiřazení by se mu tak zmenšoval.

Výhodou těchto úloh je, že omezují možnost uhodnutí správné odpovědi na minimální míru. Použit se však dají jen v poměrně omezeném okruhu učiva.

➤ **Uspořádací úlohy**

V uspořádacích testových úlohách (ordering items) se od žáka požaduje, aby předložené pojmy uspořádal podle určitého kritéria do řady. Instrukce uvádí, podle kterého kritéria a jakým způsobem se mají prvky uspořádat. Prvky se mohou řadit např. podle velikosti, významu, stupně obecnosti nebo chronologicky. V instrukci je nezbytně nutné také

zmínit, který údaj má být uveden jako první, tj. směr uspořádání - od nejmenšího po největší, od nejmladšího po nejstarší apod.

4.3.2 Otevřené položky

Otevřené úlohy vyžadují po žákovi, aby sám odpověď vytvořil. Odpovědi u otevřených položek může být slovo, číslo, výpočet, jedna věta nebo i delší text.

➤ Otevřená položka se stručnou odpovědí

Dalším velmi často používaným typem položek je otázka s krátkou odpovědí (*short - answer item*), nejlépe s jedním výrazem, slovem, číslem, značkou, symbolem, vzorcem, jednoduchým grafem nebo krátkou větou. Tyto úlohy se většinou používají pro zjišťování znalosti faktů, vzorců, vztahů či specifických informací. Poměrně snadno se konstruují a jsou objektivně skórovatelné. Vzhledem k tomu, že odpovědi může být nekonečně mnoho, je pravděpodobnost uhádnutí té správné obvykle nulová. Nevýhodou těchto úloh je to, že žák mnohdy odpovídá správně, ale jinak, než si představoval autor testu. Autor testu by měl pamatovat na to, aby otázky byly jednoznačně a dobře formulované.

Úlohy se stručnou odpovědí mohou mít charakter:

- produkční - od žáka se čeká vyprodukování odpovědi na položenou otázku
- doplňovací (*close, completion* nebo *supply items*) - žák má doplnit neúplné tvrzení

➤ Otevřené položky se širokou odpovědí

Tyto úlohy se v anglosaské literatuře označují jako *Essay - Items*. Patří mezi nejčastěji používané položky v běžných písemných zkouškách testového typu. Kmen je záměrně formulován velmi široce, což umožňuje individuální přístup k odpovědi. Požaduje se např. pojednání na určité téma, vyřešení problému, popis určitého procesu apod. Požadovaný rozsah odpovědi se žákovi může naznačit velikostí vynechaného místa v testovém zadání. Někdy se vymezuje i struktura požadované odpovědi.

Hranice mezi stručnou a širokou odpovědí není striktně vymezena. Obvykle se má za to, že pokud je odpověď v rozsahu dvou a více vět, pak znamenají širokou odpověď. Zvláštní kategorií představuje strukturovaná písemná práce – kdy žák píše rozsáhlejší slohový útvar.

Essay - položky se používají když chceme zjistit, jak umí žák myslet v určitém předmětu. Naproti tomu v případech, kde jde o zjišťování konkrétních znalostí či dovedností, tyto otázky zbytečně komplikují zkoušení. Pozitivem je, že tyto položky vedou žáka k co

nejvyšší efektivitě v organizaci a vyjadřování vlastních myšlenek. Poskytují možnost projevu vlastní individuality. Učitel tak má prostor k nahlédnutí do procesu myšlení a uvažování jednotlivých žáků a zjistit úroveň jejich myšlenkových operací např. zobecňování, abstrakce, analýza, hodnocení.

Zdánlivou předností je snadná tvorba otázek. Nepromyšlená formulace textu může nadměrně zkomplikovat či znemožnit objektivní hodnocení, které je největší slabinou těchto položek. Testy z těchto úloh jsou subjektivně skórovatelné. To znamená, že dva či více nezávislých posuzovatelů nedojde při hodnocení odpovědi ke stejným závěrům. Dokonce se může stát, že závěry téhož učitele jsou různé, hodnotí – li odpověď s časovým odstupem.

4.4 Porovnání uzavřených a otevřených položek

Proč testy s výběrem odpovědi ano?

Velkou výhodou je objektivita skórování. Dobře vytvořený test, který obsahuje jen úlohy s výběrem odpovědi, má vysokou reliabilitu (výsledky měření jsou stabilní při opakovaném zadání testu) a validitu (měří skutečně ten výkon, pro který byl vytvořen a použit). Výhodou takového testu je i rychlé a levné zpracování výsledků pomocí výpočetní techniky, jejich vyhodnocování a standardizování (Berová, 1998). Tyto testy jsou také vhodné pro žáky, kteří mají potíže s formulováním vlastní odpovědi nebo pomalu píší.

Proč testy s výběrem odpovědi ne?

Odborníci na tvorbu testů si stále více uvědomují, že některé vědomosti a dovednosti žáků se velmi těžko testují pomocí uzavřených položek (Berová, 1998). Objektivně skórovatelné testy s výběrem odpovědi nemohou plně a komplexně zachycovat všechny vzdělávací cíle (Šteffl, 1998).

Testy s výběrem odpovědi mohou posloužit jedinému cíli: prostřednictvím dotazů na nejzákladnější a tedy nesporné věci oddělit nositele podprůměrných výkonů. Jedině za tímto účelem jsou takové testy používány v Nizozemí či ve Spojených státech. Představa, že testování odliší také například studenta výborného od chvalitebného, je mylná. U testů s náročnějšími, a tedy nejednoznačnými úlohami totiž závisí dosažení maximálního počtu bodů nikoli na probantově zdatnosti, ale na míře, v níž se ztotožní se způsobem uvažování autora testu (Malina, 1999).

U testů s výběrem odpovědí je snadnější opisování, při nevhodně zvolených distraktorech mohou žáci správnou odpověď jen uhádnout a nejde zde vysledovat myšlenkový postup žáka ke konečnému řešení.

Testy by měly zjišťovat nejen znalosti, ale také kreativitu žáků a schopnost využít získané poznatky a dovednosti. Proto se do testů s uzavřenými otázkami vkládají často i úlohy aplikačního charakteru.

Proč otevřené úlohy ano?

Otevřené úlohy, na které žáci odpovídají vlastními slovy, jsou vhodné pro uplatnění individuality žáka. Dávají širší možnost aktivizovat náročnější myšlenkové operace (aplikaci, srovnání, analýzu,...) a uplatnit a rozvíjet formulační a stylistické dovednosti žáků.

U problémů, jejichž řešení vyžaduje několik kroků, výsledek informuje ve kterém z nich došlo k omylu. Pravděpodobnost uhádnutí je zde na rozdíl od úloh s výběrem odpovědi nulová.

Proč otevřené úlohy ne?

Obtížně se hodnotí. Odpovědi na tutéž otázku se mohou značně lišit. Problémy by mohly nastat, kdybychom chtěli použít testy s otevřenými otázkami (zejména esej testy) na nějaké celoplošné testování. Bylo by potřeba přesně stanovit kritéria pro hodnocení, proškolit, vycvičit a zaplatit tisíce hodnotitelů a zajistit složitou logistiku předávání esejí (každou esej musí hodnotit dva nezávislí hodnotitelé a v případě neshody ještě třetí). Bylo by také potřeba, aby žáci byli s kritérii hodnocení těchto testů seznamováni během celého studia.

Z předchozích úvah plyne, že při tvorbě otevřených úloh by autor měl především dbát na jejich jednoznačnost, jasně definovat požadavky a přesně určit způsob jejich vyhodnocování a skórování. Chceme-li opravdu poznat žákovu osobnost je vhodné i přes velkou náročnost při vyhodnocování testu použít testy s úlohami otevřenými.

4.5 Konstrukce didaktických testů

Vytvořit dobrý didaktický test, který je určený pro široké použití je poměrně náročný proces. Podílí se na něm celá řada odborníků a jeho příprava trvá i několik měsíců. Pro vytvoření desetiminutového testu do třídy k ověření minulé látky je potřeba jiná doba než pro přípravu souhrnného, např. celostátního maturitního testu.

Pro konstrukci didaktického testu je také nutné si nejprve definovat cíl testu (k čemu mají sloužit získané výsledky), jeho obsah (jaké znalosti a dovednosti se mají v testu ověřovat), určit si cílovou populaci žáků (typ školy, studijní obor, ročník studia), druh testu, vypracovat specifikční tabulku (obsahuje očekávané znalosti a dovednosti žáka, tematické celky učiva a specifické cíle), stanovit název testu, odhadnout počet spolupracovníků pro přípravu testu a časovou a finanční náročnost přípravy. Při úvahách o tom, jakou výslednou úroveň osvojení poznatků úlohy testují, jsou velmi dobrou pomocí různé taxonomie výukových cílů např. Tollingerové taxonomie učebních úloh (Obst, 2006) nebo Bloomova taxonomie výukových cílů (Skalková, 1999). Poté mohou autoři přistoupit ke konstrukci vlastních testových položek a jejich následnému ověřování v praxi.

Autor testu by měl být odborníkem v testované oblasti, ale také dobrým pedagogem a psychologem, aby se dokázal vcítit do pozice testovaných žáků. Při samotné konstrukci testu by se mělo dbát na to, aby testové položky byly jednoznačné, věcně správné, správně formulované a přiměřeně dlouhé. Také by měly mít nezpochybnitelné odpovědi a skutečně měřit to, co se vytyčilo v cílech. Jakmile má autor k dispozici dostatek testových úloh, může se pustit do vlastní konstrukce testu. Vybere vhodné úlohy, zařadí je do testu, odhadne časovou náročnost, zvolí způsob hodnocení jednotlivých úloh ale i celého testu. Dále by měl sestavit záznamový arch, do něhož by mohli žáci zapisovat své odpovědi a popsat organizaci testování (jak má být učebna vybavena, povolené pomůcky, podrobný popis průběhu testování včetně zpracování písemných instrukcí pro zadavatele).

Test určený pro širokou veřejnost musí projít fází ověřování. Recenzenti jej hodnotí po stránce obsahové, konstrukční a jazykové. V tzv. pilotáži jsou části testu zkoušeny na vybraném vzorku studentů. Z této pilotáže se dále určuje, které úlohy jsou vhodné zařadit do testu a které ne.

Doporučení při tvorbě testových položek se u jednotlivých typů liší (upraveno podle www.ceremat.cz, 14.4.2007).

Doporučení pro tvorbu dichotomických testových položek

Odpověď na tvrzení v úloze musí být jednoznačná (ANO nebo NE, PRAVDA nebo NEPRAVDA) a musí být jasně uvedeno, jakým způsobem se má správná odpověď označit. Pravdivých i nepravdivých tvrzení by mělo být v testu zhruba stejný počet a měla by se střídát, aby nevznikala schémata (např. ANO, ANO, ANO, NE, NE, NE). V tvrzeních by se tvůrce měl vyhýbat výrazům jako např. vždy, často, téměř, nikdy a zřídka. Neměly by se používat věty vytržené z kontextu učebnice. Tvrzení by měla být vždy formulována stručně a jasně.

Doporučení pro tvorbu uzavřených položek s výběrem z více odpovědí

U těchto úloh musí být jasně řečeno, zda jsou v testu použity úlohy s jedním, žádným nebo větším počtem správných odpovědí. Nejvhodnější počet alternativ je 3 - 5, přičemž nejčastěji se používají 4 alternativy. Nižší počet nahrává hádání a u vyššího je potřeba větší čas na vypracování testu a žáci se mohou snáze v úkolech ztrácet. Otázka by se měla konstruovat až na základě přesného znění řešení. Správná řešení musí být jednoznačně správná, je potřeba posoudit všechny distraktory, zda jsou skutečně jednoznačně nesprávné. V zadání úlohy se vyhýbáme všem slovům, údajům či tvarům, které by mohly sloužit jako nápověda.

Pro formulaci distraktorů by autor měl využít chyb, kterých se sami žáci dopouštějí nejčastěji, používat slova, která mají asociaci s výchozím textem nebo kmenem úlohy, formulace z učebnice apod. Alternativy by si měly být podobné svým obsahem i formou, správná odpověď by se neměla výrazně lišit od distraktorů. Distraktory by měly být stejně atraktivní, aby nedošlo k vyloučení zcela nesmyslných distraktorů a tím zmenšení možných variant.

Používání záporu by mělo být jen zřídka a měl by se graficky zvýraznit (např. tučně nebo podtržením). Distraktory by se měly umisťovat zcela náhodně, nebo např. seřadit podle abecedy. Bývají sklony nedávat správnou odpověď na první místo a žáci na to často mohou přijít. Zřídka by se měly používat alternativy jako „všechny výše zmíněné odpovědi“ nebo „žádná z výše uvedených odpovědí“.

Doporučení pro tvorbu testových položek přiřazovacích

Seznamy pojmů by měly být poměrně krátké a stejnorodé. Každý bod v prvním seznamu se čísluje a každé alternativě se přiřazuje písmeno. Možnosti by měly být uspořádány logicky. Alternativ se uvádí vždy více než položek výchozího seznamu, aby se eliminovala vylučovací metoda. Vždy se definuje způsob vyznačení dvojic a pokud lze alternativy přiřadit více než jednou, výslovně se na to musí upozornit.

Doporučení pro tvorbu testových položek uspořádacích

U těchto úloh je také velmi důležité přesně definovat způsob, jakým mají žáci vyznačit pořadí. Musí se uvést, který údaj má být uveden jako první, tzn. směr uspořádání (např. od nejmenšího po největší, od nejmladšího po nejstarší, od nejbližšího po nejvzdálenější apod.).

Doporučení při tvorbě testových položek se stručnou odpovědí

Volí se vždy úlohy, na něž se dá stručně odpovědět. V odpovědích se nevyžaduje přesné opakování textu z učebnice. Vždy se musí uvažovat všechny možné odpovědi. Při velkém počtu se úloha musí přeformulovat. Na doplnění odpovědi je nutné ponechat dostatečně velké místo. U doplňovacích úloh se vynechávají pouze podstatné údaje, ale jen v takové míře, aby byla věta smysluplná. Pokud se doplňují číselné údaje, vždy je nutné příslušnou jednotku předem vypsát.

Doporučení při tvorbě testových položek se širokou odpovědí

Na začátku by si měl autor promyslet, jakou rozumovou operaci bude po žákovi požadovat. V instrukci by se měly používat slova jako vysvětlete, zdůvodněte, popište, obhájte, porovnejte, zhodnoťte, ukažte na příkladu, odhadněte, co by se stalo, kdyby... apod. Vyhnout by se mělo slovním spojením pokuste se nebo dokážete popsat? Úloha by se také neměla začínat slovy co, kdy, kdo, protože tyto slova vedou k pouhému vypisování faktů.

Úloha musí být formulovaná tak, aby žák měl jasnou představu co se od něj očekává, v jaké formě nebo stylistickém útvaru má být odpověď napsaná, v jakém rozsahu a zda se hodnotí i pravopis. Podstatné je také vymezení, co se má v odpovědi objevit.

Tyto testové položky by neměly hodnotit správnost postojů či názorů zastávaných žáky. Testové položky se širokou odpovědí by se měly využívat pouze tam, kde není možné použít k uspokojení specifických cílů ostatní typy úloh.

5. Závěr

V práci byl podán přehled o historii testování, jehož počátek se klade do 20. let minulého století. Snahy o racionální řízení výchovně vzdělávacího procesu vedly k potřebě získat objektivní obraz o úrovni a kvalitě osvojovaných vědomostí. Výzkumy ukázaly, že testování pomocí vhodně konstruovaných didaktických testů tento požadavek plní, testování se ujalo a stalo se běžnou součástí evaluačního systému po celém světě.

Dále byly v práci charakterizovány vybrané instituce a projekty, zabývající se hodnocením vyučovacího procesu v zahraničí i u nás. Celonárodním projektem v USA je NAEP, který se již od roku 1969 zabývá soustavným sledováním úrovně vědomostí a dovedností amerických studentů. Další podobnou americkou organizací je ETS. Mezi mezinárodní organizace patří např. IEA, která realizuje celou řadu projektů (např. TIMSS a SITES) a ACER (Australská rada pro pedagogický výzkum), realizující např. projekt PISA. V západní Evropě se testováním zabývá např. nizozemský institut CITO. V Evropě východní má velmi dobrou úroveň testování v Polsku, nejvýznamnější organizace jsou CKE ve Varšavě a PTDE v Krakově.

Mezi nejznámější organizace, které se zabývají testováním u nás patří CERMAT, Scio a Kalibro. V práci jsou popsány projekty těchto institucí, zejména pak příprava a realizace nové maturitní zkoušky a projekt hodnocení výsledků vzdělávání žáků 9. tříd a odpovídajících ročníků víceletých gymnázií, který je řešen v rámci projektu Kvalita 1. CERMAT od roku 2001 připravuje a realizuje dlouhodobý cyklus programů pod názvem Krok za krokem k nové maturitě, zaměřený na přípravu škol a žáků na novou podobu maturitních zkoušek. Kvalita 1 je projekt Evropského sociálního fondu (ESF), jehož cílem je postupné vytvoření nového komplexního systému monitorování a hodnocení výsledků základního a středního stupně vzdělávání. Projekt připravilo MŠMT a podílí se na něm CERMAT, NÚOV a ÚIV.

Velmi důležitým nástrojem, používaným při testování jsou kvalitní didaktické testy a testové položky. Jsou zde charakterizovány jednotlivé typy didaktických testů, testových položek a způsob jejich konstrukce.

Přínosem práce je pohled na zahraniční organizace a realizované projekty v rámci přírodovědného vzdělávání. Uvedený přehled může pomoci stávajícím i budoucím učitelům biologie a pracovníkům zabývajícím se testováním v přípravě kvalitních úloh pro své žáky, které by vedly k rozvíjení a hodnocení klíčových kompetencí žáků.

V rámci diplomové práce bych ráda využila poznatků z bakalářské práce k vytvoření souboru úloh, které by rozvíjely a následně ověřovaly u studentů gymnázia především kompetenci k řešení problémů souvisejících s využitím biologických poznatků v praxi.

6. Použitá literatura a internetové stránky

citováno podle ČNS ISO 690 (1996)

- BEROVÁ, M. Meranie a hodnotenie výsledkov vzdelávania v mezinárodných komparatívnych štúdiách (Vedomosti slovenských žiakov z prírodovedných predmetov v mezinárodnom porovnaní). *Pedagogická revue*, 1998, roč. 50, č. 3, s. 234-247.
- BYČKOVSKÝ, P. 1982. *Základy měření výsledků výuky. Tvorba didaktického testu*. Praha: ČVUT, 1982. 149 s.
- HALADYNA, T., M. 1999. *Developing and validating multiple-choice test items*. Lawrence Erlbaum Associates, 1999. 228s. ISBN 0805837479.
- HALADYNA, T., M. 1997. *Writing test items to evaluate higher order thinking*. Allyn & Bacon, 1997. 272 s. ISBN 0205178758.
- HNILIČKOVÁ, J., JOSÍFKO, M., TUČEK, A. 1972. *Didaktické testy a jejich statistické zpracování*. 1. vydání. Praha : Státní pedagogické nakladatelství, 1972. 200 s.
- HRABAL, V., LUSTIGOVÁ, Z., VALENTOVÁ, L. 1994. *Testy a testování ve škole*. Praha: Pedagogická fakulta UK, 1994. 101 s.
- CHRÁSKA, M. 1999 *Didaktické testy- Příručka pro učitele a studenty učitelství*. Brno: Paido, 1999. 91s. ISBN 80-85931-68-0.
- CHRÁSKA, M. Co to je didaktický test? *Učitelské listy*, 1999, roč. 6, č.2, s.16-17.
- IRVINE S., H., KYLLONEN P., C. 2001. *Item generation for test development*. Lawrence Erlbaum Associates, 2001. 412s. ISBN 0805834419.
- KUKLOVÁ, J. Ověřování specifických cílů tematického celku biologie živočichů [Diplomová práce]. Praha: PřF UK, 2003.
- MALINA, T. Sezame, otevři se aneb o sondě. *Učitelské noviny*, květen 1999, roč. 102, č. 20, s.13.
- METELKA, M. *Soubor testů z učiva biologie člověka pro gymnázia* [Diplomová práce]. Praha: PřF UK, 2001.
- MORRISON, S., SMITH, P., BRITT, R., BOYD, D., MURRAY, A. 1996. *Critical thinking and test item writing*. Health Education System, 1996. ISBN 0965667804.
- MRÁČKOVÁ, G. 1999. *Soubor testových úloh z učiva přírodopisu* [Diplomová práce]. Praha: PřF UK, 1999.
- OBST, O. 2006. *Didaktika sekundárního vzdělávání*. Olomouc: Univerzita Palackého v Olomouci, 2006. 195 s. ISBN 80-244-1360-4.

OSTERLIND, S. J. 1998. *Constructing test items*. Kluwer Academic Publishers, 1998. 352 s. ISBN 0792380770.

PALEČKOVÁ, J. Mezinárodní výzkum matematického a přírodovědného vzdělávání - výsledky českých žáků. In: *Nové možnosti vzdělávání a pedagogický výzkum* (sborník příspěvků IX. Mezinárodní konference ČAPV s mezinárodní účastí). Ostrava: 2001. s.283-287. ISBN 80-7042-181-9.

PŮLPÁN, Z. 1991. *Základy sestavování a klasického vyhodnocování didaktických testů*. Hradec Králové: Kotva, 1991. 148 s. ISBN 80-900254-4-7.

PRŮCHA, J. 1996. *Pedagogická evaluace - hodnocení vzdělávacích programů, procesů a výsledků*. 1. vydání. Brno: Masarykova univerzita, 1996. 168 s. ISBN 80-210-1333-8.

SKALKOVÁ, J. 1999. *Obecná pedagogika*. Praha: ISV, 1999. 292 s. ISBN 80-85866-33-1.

SLAVÍK, J. 1999. *Hodnocení v současné škole: východiska a nové metody pro praxi*. Praha: Portál, 1999. 192 s. ISBN 80-7178-262-9.

STRAKOVÁ, J. Měření matematické gramotnosti a analytických dovedností patnáctiletých žáků ve výzkumu OECD PISA. In: *Nové možnosti vzdělávání a pedagogický výzkum* (sborník příspěvků IX. Mezinárodní konference ČAPV s mezinárodní účastí). Ostrava: 2001. s.466-469. ISBN 80-7042-181-9.

STRAKOVÁ, J., TOMÁŠEK, V., PALEČKOVÁ, J. *Třetí mezinárodní výzkum matematického a přírodovědného vzdělávání: souhrnné výsledky žáků 8. ročníků*. Praha: Výzkumný ústav pedagogický, 1996. 76s.

ŠTEFFL, O. K sondě Maturant 98. *Učitelství noviny*, září 1998, roč. 101, č. 33, s. 17-20.

THORNDIKE, R. L. a HAGEN, E. 1967. *Measurement and Evaluation in Psychology and Education*. 2. vydání. New York: John Wiley, 1967.

UN. Testování. *Učitelství noviny*, 1997, roč. 100, č. 25, s.14-20.

Internetové stránky:

www.acer.edu.au (27.4.2007)
www.ceremat.cz (14.4.2007)
www.cito.com (10.7.2007)
www.cito.nl (27.4.2007)
www.esf-kvalita1.cz (14.7.2007)
www.ets.org (21.4.2007)
www.iea.nl (17.7.2007)
www.kalibro.cz (13.7.2007)
www.kvalita1.cz (14.7.2007)
www.nationsreportcard.gov (28.7.2007)
www.nces.ed.gov (28.7.2007)
www.oke.krakow.pl (17.7.2007)
www.ptde.org (20.7. 2007)
www.scio.cz (27.4.2007)
www.uiv.cz (14.7.2007)