

Didaktické testy

Doplňkový text k předmětu Pedagogika ve studiu učitelství na MFF UK
Pavla Zieleniecová, únor 2013

Obsah:

- a) Didaktické testy a jejich účel
- b) Vlastnosti didaktického testu
- c) Druhy didaktických testů
- d) Druhy testových úloh
- e) Uplatnění testů ve výuce
- f) Pro a proti didaktickým testům

a) Didaktické testy a jejich účel

Didaktický test je typem zkoušky; jeho účelem je *zjišťování výsledků výuky*, tj. vědomostí a dovedností žáků.

Klasická zkouška taková, jakou ji známe ze školních tříd, má některé nedostatky, které z ní činí ne zcela objektivní a spolehlivý nástroj zjišťování výsledků, jichž žáci svým učením dosáhli. Může být například ovlivněna subjektivním vztahem učitele a žáka, nerovnými podmínkami různých žáků při zkoušení (často je různý obsah a různá obtížnost zkušebních otázek, různá doba zkoušení, různé prostředí), aktuálním psychickým stavem učitele, nejasnými kritérii hodnocení a mnohdy i ne zcela promyšlenými otázkami učitele, které nemusí reprezentovat právě tu vědomost či dovednost a právě na té úrovni, o jaké se učitel domnívá, že ji zjišťuje. Didaktický test se od klasické zkoušky liší především tím, že jsou v něm minimalizovány vnější vlivy, které by mohly nerovným způsobem ovlivnit výsledek žáků, ať už by plynuly ze subjektivního vlivu osob, které se zkoušky účastní (např. zkoušející, resp. zadavatel testu, hodnotitel), z okolního prostředí (např. hygienické a psychohygienické podmínky), z testu samotného (chyby, nejasnosti a dvojznačnosti zadání, špatné psychometrické vlastnosti zkoušky, nepřiměřenost zkoušky možnostem žáků), z průběhu administrace testu zadavatelem a řešení testu žáky (jasné instrukce při zadávání testu, zamezení podvádění). Didaktický test je při dodržení všech pravidel při jeho tvorbě, použití a hodnocení objektivnějším, validnějším, spolehlivějším, přesnějším a citlivějším nástrojem než klasická zkouška.

b) Vlastnosti didaktického testu

O kvalitě testu rozhoduje jeho objektivnost, reliabilita (přesnost a spolehlivost), validita (platnost) a citlivost (rozlišovací schopnost).

Objektivnost znamená, že výsledek testu závisí pouze na žáku, ne na dalších okolnostech mimo něj (tj. nepromítá se do něj především subjektivní vliv učitele). K zajištění vysoké míry

objektivnosti didaktického testu přispívá jednoznačná formulace úloh testu, jednotné podmínky při jeho zadávání a přesná pravidla hodnocení.

Reliabilita testu je mírou chyby, jíž jsou zatíženy výsledky testu. Podobně jako chyba při jakémkoli jiném měření, i reliabilita testu má dvě složky: *spolehlivost* (odpovídající náhodné chybě) a *přesnost* (odpovídající systematické chybě). Didaktický test je *spolehlivý*, není-li jeho výsledek ovlivněn náhodnými vlivy působícími na žáka při řešení. To znamená, že při opakovaném zadání testu stejným žákům (za hypotetického předpokladu, že si žáci úlohy testu a jejich řešení nepamatují) by test přinesl stejný výsledek. Jinými slovy, výsledek testu je stabilní. Zdrojem nízké spolehlivosti testu jsou různé proměnlivé, náhodné vlivy, které při řešení testu působí na žáka a které mohou mít původ jak ve vnějším prostředí, tak v žákovi samotném. Přesnost didaktického testu postihuje posunutí (zkreslení) výsledků testu v důsledku trvalejších, systematicky působících vlivů. Zdroje nízké spolehlivosti testu bývají především v testu samotném - může se například jednat o formulace testových úloh neodpovídající úrovni znalostí žáků, o terminologii, která nebyla používána, o nejednoznačné či nejasné zadání úloh.

Validita didaktického testu znamená, že test měří skutečně to, co měřit má, že zjišťuje tu vědomost či dovednost a na té úrovni, o jaké se tvůrci a uživatelé testu domnívají, že měří.

Citlivý didaktický test dobře svými výsledky rozlišuje výkon různých žáků. Krajním případem necitlivého testu je, dosáhnou-li v něm všichni žáci dobrých nebo všichni špatných výsledků. Citlivý test by měl rozložit výsledky žáků po celé možné škále hodnocení a rozlišovat i mezi blízkými výkony žáků.

Pro zajištění těchto vlastností se při konstrukci testu provádí jeho předběžné ověřování (tzv. pilotáž), kdy se statistickými metodami zkoumají výsledky vzorků žáků náhodně vybraných z populace, pro kterou je test určen.

Kromě uvedených vlastností je třeba zmínit i to, že didaktický test je ve srovnání s klasickou zkouškou časově úspornější, a to jak pro žáky, tak pro učitele, resp. zadavatele testu.

c) Druhy didaktických testů

Testy je možné rozlišovat podle řady hledisek: podle způsobu zadání (testy zadané písemně, ústně, praktickou ukázkou, na počítači, kombinovaně), podle způsobu řešení (testy řešené písemně, ústně, prakticky, na počítači, kombinovaně), podle dostupnosti (testy veřejné, neveřejné) aj. Pro učitele je důležité rozlišení testů podle způsobu, jakým byly zkonstruovány, na *objektivně skórovatelné* a *subjektivně skórovatelné*.

Objektivně skórovatelné testy se konstruují podle propracované metodiky, která zaručí jejich objektivnost, spolehlivost, validitu a citlivost v dané cílové populaci. Vyžadují profesionální přístup, jejich sestavování není v silách jednotlivého učitele. Východiskem je vždy analýza cílů výuky a příslušného obsahu učiva, jehož osvojení žáky má test měřit. Na jejím základě jsou z rozsáhlé zásoby vybírány testové úlohy. Provádí se zpravidla několikastupňové pilotování ve

vzorcích vybraných z cílové populace, při nichž se ověřují jak jednotlivé položky, tak základní vlastnosti testu jako celku, a test se koriguje.

Objektivně skórovatelné testy se zadávají, skórují (vyhodnocují) a interpretují podle jednoznačných a pro všechny stejných, předem daných kritérií a pravidel. Umožňují srovnání jednotlivých žáků ve třídě, tříd mezi sebou, škol, zemí.

Objektivně skórovatelný test může být *standardizovaný*, je-li pro něj vytvořen tzv. standardizovaný výsledek (standard, norma) v cílové populaci. Konečný test se zadá reprezentativnímu výběru žáků dané populace, jehož výsledek představuje standard. S tímto standardním výsledkem je pak možné srovnávat výsledky různých žáků a skupin žáků. U *nestandardizovaných* testů tento standard vytvořen není.

Subjektivně skórovatelné testy nemají důsledně zaručeny výše uvedené základní vlastnosti; neumožňují proto objektivní a spolehlivé srovnávání žáků a nemůže pro ně ani být vytvořen nějaký standard. Jsou zpravidla sestavovány učitelem, který je zadává svým žákům, a mají sloužit potřebám výuky konkrétního učitele u konkrétních žáků. Jejich výhodou je, že učitel může přesně zacílit zkoušku na míru svých žáků, cílů výuky, probraného učiva a účelu zkoušky.

Přesto, že není v silách učitele dodržet metodiku vytvořenou pro objektivně skórovatelné testy, měl by dbát alespoň na základní pravidla tvorby testů (nejprve podrobná rozvaha, co má test zjišťovat - jak co do obsahu vědomostí a dovedností, tak jejich úrovně, cílený výběr úloh s ohledem na tyto obsahy a jejich úroveň, vytvoření kritérií pro hodnocení řešení jednotlivých úloh, stanovení jasných pravidel pro zadání, a pokud možno i předběžné ověření testu).

Dodrží-li učitel základní pravidla pro sestavování testu, jeho zadání žákům a vyhodnocení, bývá výsledek testu zpravidla přesnější a spolehlivější ve srovnání s tradičně zadanou zkouškou. Nicméně učitel si při interpretaci výsledků testu vždy musí být vědom, že výsledky jsou zatíženy větší chybou (kterou navíc nemá možnost odhadnout), než výsledky testu objektivně skórovatelného. Velmi důležité je, aby výsledky interpretoval opatrně, bez jejich absolutizování, a měl na paměti, že jsou jen jedním z ukazatelů výsledků žáka, které je nutno zjišťovat i jinými způsoby.

d) Druhy testových úloh

Testové úlohy (testové položky) mohou být uzavřené nebo otevřené. *Uzavřené* úlohy mají jednu nebo více správných odpovědí, které jsou nabídnuty žákovi v textu úlohy. U *otevřených* úloh žák sám tvoří odpověď, přičemž spektrum přijatelných odpovědí bývá zpravidla širší. Tvořená odpověď může představovat jedno slovo, jednu větu nebo i delší pojednání. Čím je úloha otevřenější, tím je odpověď delší, a tím je také širší spektrum přijatelných odpovědí.

Mezi uzavřené testové položky patří:

- úlohy s *dvoučlennou volbou*: žák volí odpověď ze dvou možností, ano a ne,

- úlohy s *vícenásobnou volbou*: úloha obsahuje vedle zadání i nabídku několika odpovědí, tzv. alternativ, z nichž jedna nebo více je správných,
- úlohy *situační*: možné odpovědi jsou nepřímo nabídnuty v zadání, například v tabulce či grafu, který je součástí zadání,
- úlohy *přiřazovací*: zadání obsahuje dvě řady pojmů, čísel apod., úkolem je přiřadit k sobě dvojice z obou řad podle nějakého pravidla,
- úlohy *uspořádací*: úkolem je seřadit za sebou zadané pojmy, čísla apod. podle nějakého pravidla.

Otevřené položky mohou být tohoto typu:

- úlohy *doplňovací*: úkolem je doplnit do zadaného textu, vztahu apod. chybějící prvky (slova, čísla, symboly,...),
- úlohy s *krátkou tvořenou odpovědí*: vyžadují krátkou odpověď, např. číslo získané výpočtem, vztah získaný algebraickou úpravou, slovo nebo větu odpovídající na zadanou kvalitativní úlohu,
- úlohy s *širokou odpovědí*: jejich řešením je širší pojednání, obsahující například úvahu na dané téma, hodnotící posouzení určitého textu, matematické odvození či důkaz,
- *praktické* úlohy: jsou komplexní, představují zpravidla ucelený problém nebo projekt zahrnující získání potřebných dat, jejich vyhodnocení a interpretaci; například se může jednat o otevřenou laboratorní úlohu.

Čím je testová položka více otevřená, tím je obtížnější zajistit její objektivní skórování. Hodnocení těchto položek nemůže být jen "správně - nesprávně" nebo na jednorozměrné stupnici, jako je tomu u uzavřených úloh. Je potřeba použít více hodnotících kritérií, která vystihují různé aspekty řešení (tj. stupnice hodnocení je vícerozměrná), přičemž pro každou úlohu se musí vytvořit zvláštní soubor hodnotících kritérií. Testy s otevřenými položkami rovněž vyžadují, kromě této přípravy hodnotících kritérií a stupnic, také velmi pečlivý proces ověřování. Jeho účelem není jen výběr úloh, ale také ověření těchto hodnotících kritérií a způsobu, jakým je uplatňují hodnotitelé. Někdy se pro zajištění větší objektivity zadává vyhodnocování vyřešených testů ne jednotlivci, ale skupině hodnotitelů.

Tímto komplikovaným procesem přípravy a vyhodnocování se testy s otevřenými položkami prodražují; přesto však se k nim praxe testování stále více přiklání. Položky s uzavřenou odpovědí mohou totiž měřit výsledky výuky jen nižší úrovně; složitější dovednosti mohou postihnout jen otevřené položky.

e) Uplatnění testů učitelem ve výuce

Učitel může ve výuce uplatnit jednak testy, které si sám sestavil, nebo profesionálně sestavené testy, které je možné si objednat u komerčních subjektů. Testy může využít pro účely *diagnostické* – ke zjištění, čemu se žák naučil, kde má případné nedostatky; pro účely *procvičovací* – k zopakování a procvičení probraného učiva; pro účely *zkušební* – k hodnocení a případné klasifikaci žáka na základě dosaženého výsledku; pro účely *srovnávací* – k porovnání

výsledků žáků ve třídě, tříd ve škole, výsledků žáků s výsledky jiných škol nebo k porovnání se standardizovaným výsledkem v dané populaci.

Využívání testů ve škole má svoji morální stránku; učitel musí zaručit, že testy a jejich výsledky nebudou využívány proti zájmům žáků či skupin žáků. I proto je důležité, aby žáky nehodnotil jen na základě výsledků testů, byť by byly sebeobjektivnější, ale aby jako zdroje poznání a hodnocení žáků využíval i jiné zdroje informací, např. klasickou zkoušku, hodnocení prací žáků (portfolia) apod.

Učitel se setká i s testy, které jsou zadávány jeho žákům z rozhodnutí mimo učitele a někdy i mimo školu. Mohou to být např. testy sloužící výzkumným účelům v národním i mezinárodním měřítku, mapování výsledků vzdělávání pro potřeby rozhodovacích orgánů (ministerstva školství, krajů) a zkušebním účelům (společné maturity, zkoušky v 5. a 9. ročníku základní školy). Všechny tyto testy by měly mít vysoce profesionální charakter a měly by mít zajištěny všechny nezbytné základní vlastnosti kvalitních testů. Zadávající organizace (například CERMAT, Česká školní inspekce) by měly škole poskytnout podrobné vyhodnocení výsledků jejich žáků, i s možností (anonymního) srovnání s jinými školami. Učitel pak může využít výsledků těchto testů ve výuce pro diagnózu, procvičování, hodnocení, srovnávání.

f) Pro a proti didaktickým testům

Didaktické testy jsou – nejen u nás – předmětem neustávajících diskusí a sporů. Zastánci a kritici využívání tohoto prostředku hodnocení výsledků vzdělávání ve škole uvádějí řadu výhod a nevýhod, mezi nimiž mnohé mají zásadní povahu.

Jako nejdůležitější se mezi zmiňovanými výhodami didaktických testů uvádí srovnatelnost zjištěných výsledků vzdělávání. Na základě výsledků testů může škola, učitel i žáci porovnat své výsledky, měřené testem, se školami a žáky jinými a udělat si tak představu o jejich kvalitě, úrovni, náročnosti. Kritici v tomto ohledu odmítají především vytváření pořadí (žebříčků) škol na základě výsledků testů a zcela zásadně varují před tím, aby byla taková pořadí škol využívána k administrativním opatřením, například k výběru škol, které mají být zrušeny.

Zásadní výhodou je také možnost širokého použití jediného testu za podmínek, které zaručují vysokou objektivitu jeho výsledků. Poměrně levně a rychle je takto možné udělat snímek výsledků výuky v určité populaci žáků, v určitém segmentu vzdělávací soustavy. Taková informace může mít význam nejen pro veřejnost (z jejichž prostředků je vzdělávání financováno), ale také pro veřejnou správu, například jako součást mapování stavu vzdělávací soustavy a vytváření koncepcí jejího rozvoje.

Mezi zásadními nevýhodami didaktických testů je třeba jmenovat především dvě (které spolu souvisí): omezenost vzdělávacích výsledků, které jsou didaktické testy schopny postihnout, a deformaci vzdělávacích cílů a obsahů, k níž může při použití didaktických testů docházet.

Použitá a doporučená literatura:

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

<http://www.novamaturita.cz/didakticke-testy-1404033135.html>

Kritika testů jako hodnocení školního výkonu:

<http://www.nuov.cz/uploads/Periodika/ZPRAVODAJ/2004/Zp04pVIIa.pdf>