



Zamyšlení nad testy a testováním

PhDr. Kateřina Jančaříková, Ph.D.

V poslední době probíhají v médiích, v politické sféře, ale i ve sborovnách a domácnostech diskuse týkající se plošného testování žáků 5. a 9. ročníků a také nepříliš vydařené letošní verze státní maturity. Atmosféra je napjatá. Naprosto primární je otázka, proč žáky testujeme. Jaký je vlastně smysl testování?

Veřejnost očekává od plošného testování nějaký vyšší smysl. Rozhodně by ji neuspokojilo „testování pro testování“, tedy odpověď, že cílem testování je otestovat žáky. Dle vyjádření MŠMT „nemá plošné testování sloužit k porovnání škol“, a nemá tudíž na-

pomáhat rodičům volit „tu správnou školu“ pro svého prvňáčka ani nemá být kritériem pro hodnocení školy zřizovatelem.

Ideálně by plošné testování mělo sloužit především žákům, jejich rodinám a učitelům. Přinést zpětnou

vazbu, napomoci při výběru budoucího povolání, upozornit na slabá místa a motivovat žáka ke zlepšení. Jsou to však cíle reálné? A přináší plošné testování plánovaný užitek?

V tomto příspěvku jsou vytčena některá rizika, která mohou hodnocení výsledků zkreslovat (a jistě je i zkreslují) a která – obávám se – nejsou řádně eliminována.

Rizika testů „tužka-papír“

Plošné testování proběhlo formou testu „tužka-papír“ (název typu testu je takový, třebaže samotné testování

se odehrávalo na počítačích). Při vyhodnocování výsledků je nutné si uvědomit (a pokusit se nějak zpracovat) skutečnost, že testy „tužka-papír“ s sebou přinášejí jistá rizika pro určité skupiny žáků. Howard Gardner, profesor pedagogických věd na Harvardské univerzitě, označuje testy „tužka-papír“ jako testy „nespravedlivé k rozličnými inteligencím“, a to proto, že tyto testy nehodnotí pouze znalost či neznalost odpovědi na danou otázku, ale hodnotí zároveň čtenářskou gramotnost a znalost jazyka, ve kterém byly zadány.

Žáci z určitých skupin (např. cizinci, žáci s poruchami učení, žáci nedoslýchaví a neslyšící) jsou v testech znevýhodněni. Jejich výsledek z matematiky a dalších předmětů totiž zrcadlí více jejich jazykovou neobratnost než jejich znalost/neznalost v daném oboru! Nejsnáze si problém vysvětlíme na příkladu.

Příklad 1

Otázka z písemky ze ZSV, sekunda osmiletého gymnázia

Nejprve podá ruku:

- žena muži,
- muž ženě,
- mladší staršímu,
- starší mladšímu,
- nadřazený pořízenému,
- podřazený nadřazenému.

Žák s dyslexií zaškrtnl 2 odpovědi špatně, a to ne proto, že by neznal odpověď anebo že by se v reálné situaci neuměl správně zachovat, ale proto, že si text přečetl nepozorně a „ztratil se v něm“.

Když byl okamžitě po napsání testu vyzván učitelkou, aby vlastními slovy vysvětlil pravidlo podávání rukou, tak vyjmenoval všechny možné situace a správně řekl, kdo komu ruku podává. Očividně v testu neuspěl nikoli z důvodu neznalosti.

Ano, test „tužka papír“ nepřináší přesné výsledky. Zpětnou vazbu neposky-

tuje spolehlivě, pro vyhodnocování jeho výsledků si musíme uvědomit, že:

- čím je žák mladší, tím bude menší korelace mezi výsledkem v testu a jeho reálnými studijními předpoklady;
- žáci znevýhodnění (např. s poruchami sluchu, dyslexií) budou v písemném testu chybovat i ve věcech, které dobře ovládají v praxi;
- děti cizinců, pro které není čeština mateřským jazykem, a české děti, v nichž se hovoří „jednoduchou češtinou“, dosáhnou slabších výsledků v předmětech přírodovědných, společenskovedních či matematice z důvodů nedostatečné znalosti jazyka.

Pro eliminaci chyb způsobených nedostatečnou znalostí jazyka nebo nepozorností může dobře posloužit grafické zpracování testu, např. uspořádání textu, velikost písmen vhodné ilustrace. Jenže vzhledem k tomu, že každý žák je osobností s individuálním vnímáním a s individuálním stylem učení, nikdy se nepodaří vytvořit takový test, aby maximálně vyhovoval (nevedl k nesprávným odpovědím, i když znalost tady je) každému.

Mnohem lepší, ale velice náročnou variantou testů typu „tužka-papír“ je testování v reálných situacích. Howard Gardner popisuje tvorbu takového testu (*test Spectrum*).

Test Spectrum

Test Spectrum je určen žákům mateřské školy a jeho cílem je zmapování inteligenčního spektra každého žáka a následné rozřazení do tříd, přiřazení k učitelce, tvorba individuálního studijního plánu.

Test Spectrum se skládá z 15 uměle vytvořených situací. Jednou z těchto situací je „Stavba domečku pro mimozemšťana“. Děti dostanou za úkol postavit na zahradě domeček pro mimozemšťana.

Examinátor sleduje dle metodiky celou řadu faktů: Pracuje dítě samo, nebo v hloučku? Využívá dítě jeden, nebo více typů stavebního materiálu?

Pracuje dítě jen s tím, co má na dosah, nebo odešlo pro stavební materiál mimo hřiště? atp.

V poloce testu examinátor dětem sdělí další skutečnost: „Mimozemšťan má moc rád dřevo“ a sleduje a zapisuje, jak děti na nový fakt zareagují. Zboří všechno, co dosud postavily, a začnou stavět znovu, tentokrát ze dřeva? Ignorují informaci a pokračují ve stavbě podle původního schématu? Zakomponují dřevo do stávající stavby?

Všechny materiály ze všech 15 situací jsou následně vyhodnoceny a zpracovány. Rodiče i děti jsou seznámeni s výsledky, jsou jim udělena doporučení (včetně zařazení do třídního kolektivu a vypracování individuálního studijního plánu).

Tvorba testu Spectrum trvala týmu odborníků sedm let! Ovšem dle Howarda Gardnera teprve test tohoto typu může poskytnout žákovi a jeho rodičům užitečnou zpětnou vazbu.

Rizika nedostatečné práce s chybou

Velký handicap plošného testování tkví v tom, že neumožňuje práci s chybou. Má-li učitel pochybnosti o korelaci mezi výsledkem testu a skutečnými znalostmi či dovednostmi žáka, může s žákem individuálně promluvit a zjišťovat, proč k chybě došlo, resp. ověřovat jeho znalosti ústně či názorně (jako učitelka v příkladu výše uvedeném). Význam práce s chybou si opět nejlépe vysvětlíme na příkladu (*příklad 2 na následující straně*).

Zpětná vazba, kterou testy „tužka-papír“ přinášejí, může být naprosto demotivující. Danielovi z výše uvedeného příkladu třeba sděluje: Pozor, nejsi jako ostatní, správně je odpovídat ne logicky, ale tak, jak se očekává, že bude odpovídat většina. Může se pak stát, že žáci se učí nikoli vědomostem, ale jakési pofiderní dovednosti správně odpovídat na tento typ testů.

To mimochodem dobře zná každá česká rodina z vyplňování SCIO testů – úspěšnost závisí mnohem více na

Příklad 2

Jedenáctiletý Daniel se připravoval na přijímací zkoušky na osmileté gymnázium. Ve standardizovaných testech (Kalibro), které vyplňoval s matkou při přípravě na přijímací zkoušky na osmileté gymnázium, byla tato otázka:

- Sněženky se v lese vyskytují:**
a) na jaře,
b) na podzim,
c) celý rok,
d) v zimě.

Daniel zaškrtl odpověď „celý rok“. Podle seznamu správných odpovědí to byla odpověď špatná – správná odpověď byla a). Daniel při rozhovoru s matkou nad testem uvedl: „Když si vezmu lopatku, tak tam přece cibulky najdu kdykoli...“ Dítě přírodovědně nadané může být ve standardizovaných testech vytvořených dospělými bez přírodovědného nadání znevýhodněno. Teprve práce s chybou (diskuse s matkou) přinesla jasno, že dítě projevilo nikoli nedostatek, ale spíše nadbytek znalostí.

zkušenosti se SCIO testy než na samotných dovednostech, znalostech a inteligenci žáka (podrobně se tomuto tématu věnuje i příspěvek doc. Kubanové a doc. Lindy na straně 26–31).

Rizika opisování a nápovědy

Při každém písemném testování jsou žáci v jistém morálním pokušení, aby si k lepšímu výsledku pomohli „nemorálně“. Čím více na výsledku testu záleží, tím je riziko toho, že budou opisovat od souseda, z taháku či iPadu, větší.

Příklad 3

Dvaadvacetiletá Alena, posluchačka vysoké školy, si myslí, že na opisování a různou podporu při psaní testů má právo. Tvrdí: „Když nás nezkoušejí ústně, ale testem, tak prostě budeme opisovat. Proč ne?! Dnes studium není o tom, nahučet se znalostmi. Je to o tom, mít kamarády a umět se orientovat.“ Nebo jak je důležité jít na zkoušku ve skupině a vzájemně si pomáhat při zaškrtávání možností a–d. Z jejího vyprávění záhy pochopíme, že vynalézavost studentů nezná mezí a že vyučující nemá možnost opisování či konzultování výsledků zamezit.

Alena popisuje domluvený systém:

- a) = podrhu se na hlavě, b) = zatahám se za ucho, c) = chytím se za bradu, d) = podrhu se na zádech... Přejít k nové otázce = hlasitě se nadechnu...

Při plošném testování navíc hrozí i riziko nápovědy od učitelů, protože učitelé si pochopitelně přejí dosáhnout co nejlepších výsledků svých žáků a školy. Je to skutečně nemorální, nebo jen chytré řešení navozené situace?

Při vyhodnocení testů není nikdy jasné, zda výsledek či známka vypovídají o znalostech studenta nebo o jeho jiných schopnostech. Nastavený systém znevýhodňuje studenty s vysokým vnitřním morálním kodexem (neopisující). Nebo naopak – nastavený systém vede studenty k těmto drobným podvodům.



Při vyhodnocování testů by mělo být toto riziko také zohledněno a eliminováno nějakým systémem, například v hlavičce testu bude kolonka pro jména sousedů.

Riziko netransparentnosti

Není možné testovat žáky každý rok úplně stejným testem, protože by byl

test pochopitelně okamžitě vyvěšen na nějakém serveru i se správnými odpověďmi. Tvůrci testů tedy stojí před velice obtížným úkolem: vytvořit každý rok stejně obtížnou variantu testu, ale s jinými otázkami. A to, jak ví každý, kdo musel při zkoušení písemný test použít, je obtížné, až nemožné. Každá varianta má svá úskalí. O tom jsme se přesvědčili při letošní státní maturitě, která byla značně obtížnější než státní maturita loňská.

Výsledky plošného testování lze k porovnávání napříč ročníky využít jen s nejvyšší opatrností!

Nastal efekt zklamaného očekávání?

Splnilo letošní plošné testování svou funkci? Nebo to bylo jen plýtvání penězi, časem, energií a trpělivostí učitelů, rodičů a žáků?

Vzhledem k tomu, že výše uvedená rizika nebyla dostatečně ošetřena a že plošné testy neposkytují zpětnou vazbu přímočaře (je zde příliš mnoho dalších faktorů, které výsledky testů ovlivňují a zkreslují), nezbyvá si než přiznat, že plošné testování pravděpodobně svůj primární cíl neplní.

Pokud plošné testování neplní svůj primární cíl, musíme se ptát: Jaké cíle tedy plní? Jaký má smysl? A zjišťovat jeho dopady na žáka, školu a celou společnost: K čemu ještě může plošné testování vést? Jaké jsou druhotné a skryté dopady plošného testování? Kdo na jeho organizaci vydělává? A hlavně – nebylo by lepší investovat peníze do jiného projektu? ◆

Autorka působí na katedře biologie a environmentálních studií PedF UK



Mléčný koktejl PLUS.

NOVÉ PŘÍCHUTĚ

VÝHODY:

- obsahuje vysoké množství mléka
- výborná chuť s bohatou a hustou pěnou
- bez konzervantů a umělých barviv
- jednoduchá a rychlá příprava, dobrá rozpustnost

PŘÍCHUTĚ: malina | lískový oříšek | smetana-kokos
 banán | čokoláda | jahoda | meruňka | pistácie vanilka

CATUS

CATUS spol. s r.o.
 Kyjovská 1598, 580 01 Havlíčkův Brod, ČR
 E-mail: info@catus.cz, www.catus.cz

... bez konzervantů a umělých barviv...