

Typy pro uživatele externích evaluačních služeb

Jana Straková, Petr Suchomel

Na co si dát pozor a co vyžadovat

V České republice působí několik organizací, které nabízejí školám nástroje k hodnocení výsledků vzdělávání. Je určitě užitečné, když se učitelé a ředitelé naučí těchto nástrojů využívat a doplní jimi hodnocení, které pravidelně provádějí ve svých školách. Není třeba, aby se učitelé učili konstrukci testů ani konstrukci komplexnějších evaluačních nástrojů a veškeré evaluační nástroje si vyráběli sami. Je však třeba, aby byli v oblasti hodnocení výsledků vzdělávání poučeni natolik, aby kladli na organizace, které se evaluací zabývají, dostatečně vysoké nároky. Jen budou-li učitelé a ředitelé náročnými zákazníky, bude možno docílit, aby nástroje, které jsou školám nabízeny, byly opravdu kvalitní a pro školy maximálně užitečné. K tomu je třeba, aby věděli, jaké vlastnosti má mít kvalitní evaluační nástroj, podle čeho lze volit vhodný nástroj pro daný účel a jaké užitečné informace je možno z evaluace získat a jak s nimi dále nakládat.

Hlavním cílem tohoto textu je poskytnout základní informace o charakteristikách kvalitních testů, okrajově se zmiňujeme i o jiných evaluačních nástrojích. V samostatném dokumentu pak ukazujeme, jaké výsledky by měly přinášet externí evaluační aktivity a jak je s nimi možno pracovat.

1. Testy

Nejčastějším důvodem pro vyhledání služeb organizací zaměřených na evaluaci je potřeba porovnat výsledky žáků naší školy s výsledky žáků jiných škol. Toto porovnání se provádí zpravidla prostřednictvím testů. V tomto odstavci ukazujeme, podle jakých kritérií je možno testy vybírat a na co je třeba dávat pozor při výběru testu a při interpretaci výsledků. Rada náležitostí, které zde uvádíme, není v českém prostředí běžná. V zahraničí však představují tyto požadavky minimální standard. Jen budou-li pedagogové vyžadovat řádnou tvorbu testu a jeho řádnou dokumentaci, získají nástroje, které jim poskytnou užitečnou a spolehlivou informaci o jejich práci.

Podle jakých kritérií bychom měli vybírat test

Validita testu

Základním hlediskem při tvorbě a hodnocení testu je jeho validita. Validita nám říká, zda můžeme výsledky daného testu interpretovat uvedeným způsobem. Je založena na teoretických a empirických poznacích, které se týkají měřené oblasti. Považujeme za důležité zdůraznit, že se nevztahuje k testu samotnému, ale ke způsobu interpretace výsledků. Například test z matematiky určený pro žáka 9. ročníku může být koncipován tak, aby na základě jeho výsledků bylo možno usoudit, jak žák zvládl učivo matematiky základní školy. To neznamená, že ten samý test je automaticky možno považovat za validní ve vztahu k rozhodování o dalším studiu žáka. Jedná se o dvě rozdílné interpretace: v jednom případě usuzujeme na zvládnutí učiva, ve druhém případě na vhodnost určité vzdělávací dráhy. Každý test by měl obsahovat informaci o tom, jaká z možných interpretací je validní, tedy jak mají být jeho výsledky používány a vhodnost tohoto využití by měla být teoreticky a/nebo empiricky doložena. Například na způsobilost žáka s určitým testovým výsledkem pro akademické studium může být usuzováno teoreticky z nároků tohoto studia nebo například empiricky z toho, že žáci s určitým testovým výsledkem v tomto studiu dobře prospívají.

Reliabilita

Vysoká validita zaručuje, že můžeme výsledek testu využít předepsaným způsobem. Potřebujeme se ještě dozvědět, jak kvalitní výsledek získáváme, jak moc se na něj můžeme spolehnout. Tuto informaci o kvalitě testu nám poskytuje údaj o jeho spolehlivosti, reliabilitě. Čím vyšší je reliabilita, tím větší je pravděpodobnost, že bychom při opakovaném měření dosáhli stejného výsledku, tedy že výsledek není náhodný¹. Pokud jsou součástí testu úlohy, které jsou hodnoceny hodnotiteli, měla by být součástí dokumentace i informace o shodě mezi hodnotiteli. Z dokumentace k testu bychom měli také získat údaj o chybě měření. Je-li například výsledkem testu přiřazení žáků k určitým klasifikačním stupňům, měli bychom získat informaci o tom, jaký podíl žáků je při opakovaném měření přiřazen ke stejnému klasifikačnímu stupni.

Popis testu

Povinnou součástí všech kvalitních testů vědomostí a dovedností žáků, se kterými se setkáváme v zahraničí, je přesný popis toho, co zjišťují, na jaké vědomosti a dovednosti jsou zaměřeny. Tento popis bývá zpravidla provázen rovněž zdůvodněním, proč se test zaměřuje právě na tyto aspekty vzdělávání, proč právě tyto aspekty jsou tvůrci testu považovány za důležité - z hlediska porozumění testované disciplíně, dalšího vzdělání nebo potřeb dospělého života.

Testy jsou zpravidla koncipovány tak, aby jejich řešení trvalo žákům jednu, maximálně dvě vyučovací hodiny. Za tuto dobu žáci vyřeší jen omezené množství testových úloh. Chceme-li srovnávat výsledky jednotlivých žáků, zadáváme všem žákům stejný test, a tedy můžeme ověřit jen velmi omezené množství vzdělávacích cílů. Z tohoto důvodu je mimořádně důležité, aby tyto cíle byly dobře voleny. Aby byly ověřovány základní vědomosti a dovednosti, nikoli okrajové, zbytné vzdělávací cíle². Na to, co je pro danou oblast stěžejní, mohou existovat odlišné názory. Proto je třeba vyžadovat, aby bylo vždy přesně vymezeno, na co se daný test zaměřuje, abychom mohli volit takový test, který hodnotí to, co my považujeme za zásadní a na co se zaměřujeme ve své výuce.

Součástí dokumentace k testu by tedy měl být popis toho, jak jsou ověřované vzdělávací cíle operacionalizovány: jaké dílčí cíle jsou zjišťovány a prostřednictvím jakého typu úloh a jak jsou jednotlivé dílčí cíle v testu pokryty. Součástí informace o testu by měly být i pokyny k administraci, případně k vyhodnocení žákovských odpovědí. Tyto pokyny by měly být striktně dodržovány.

Tyto informace nám pomáhají vybrat si správný test pro daný účel i vytvořit si názor na kvalitu testu.

Mnoho testů nabízených českými organizacemi zabývajícími se evaluací není opatřeno ani informací o tom, co daný test zjišťuje, ani informací o tom, jak to měří. To staví uživatele do situace, kdy musí tvůrcům testu plně důvěřovat, že zjišťují to, co on považuje za důležité a na co se zaměřuje ve své výuce, a že to měří dobře. Tato důvěra v českém prostředí ovšem nemusí být zcela na místě: setkáváme se i s testy, které více méně nahodile poskládají úlohy ze školních učebnic, z nichž některé testují důležité vědomosti a dovednosti a některé vědomosti a dovednosti zcela dílčí a okrajové, přičemž objem učiva, který chtějí zahrnout, je vzhledem k délce testu nesmyslně široký. Výsledek takového testu má pro učitele jen velmi omezenou informační hodnotu a často je obtížné jej správně interpretovat. Proto je velice důležité, aby uživatelé testů vyžadovali od organizací zabývajících se evaluací, aby testy řádně dokumentovali. Organizace, na které budou kladeny vyšší nároky, budou také produkovat kvalitnější testy, neboť budou muset o své práci více přemýšlet.

Testové skóry

Test by měl obsahovat popis stupnic, na kterých jsou stanoveny výsledky, jsou-li testování na základě výsledků rozdělování do kategorií, mělo by být zdůvodněn způsob rozdělování. Pokud je výsledek testu vyhodnocován vzhledem k normám, měl by být podrobně zdokumentován způsob

¹ Spolehlivost je zpravidla charakterizována koeficientem Cronbachovo alfa, který nabývá hodnot 0-1. Test je tím spolehlivější, čím více se hodnota blíží jedničce. Za přijatelné jsou zpravidla považovány hodnoty vyšší než 0,75.

² Zjišťujeme-li vědomosti a dovednosti žáků prostřednictvím výběrových šetření /jako jsou například mezinárodní srovnávací výzkumy PISA a TIMSS), jejichž cílem není hodnotit jednotlivé žáky, zadáváme různým žákům různé testy, což nám umožňuje ověřit daleko širší rozsah vzdělávacích cílů. V případě hodnocení jednotlivých žáků je však nutno se zaměřit na cíle, které jsou pro danou vzdělávací oblast stěžejní.

vytváření norem. Vznikají-li dvě varianty testu (A, B)), mělo by být podrobně popsáno, jak byla zajišťována jejich ekvivalence.

Rozlišovací a ověřovací testy³

U každého testu musí být zřejmé, zda se jedná o test, jehož cílem je rozdělit žáky podle úrovně jejich vědomostí a dovedností (rozlišovací test) nebo pouze ověření, zda si žáci osvojili jistou minimální úroveň vědomostí a dovedností (ověřovací test). Rozlišovací test je vhodný typicky pro účely přijímacího řízení. Ověřovací testy se naopak velmi dobře hodí jako zpětná vazba pro učitele: poskytnou informaci o tom, jaký podíl žáků si osvojil vědomosti a dovednosti, které jsou považovány za důležité, respektive na které se jeho výuka zaměřuje.

Poznámky k současné nabídce českých evaluačních agentur

Zde uvádíme několik aspektů spojených s nabídkou českých organizací zabývajících se hodnocením výsledků vzdělávání, které nám připadají z hlediska využití školami matoucí a považujeme za důležité na ně upozornit. Náš výčet určitě není vyčerpávající.

Testy studijních předpokladů, obecných dovedností

Školám jsou často nabízeny všelijaké testy studijních předpokladů či tzv. obecných dovedností. V případě těchto testů je třeba se ještě více zajímat o to, co vlastně měří, než u testů z dílčích vzdělávacích oblastí, abychom byli schopni jejich výsledky správně interpretovat. Tyto testy zpravidla posuzují základní jazykové, respektive matematické dovednosti žáků a jejich logické uvažování. Základní problém při interpretacemi jejich výsledků je ovšem skutečnost, že často nejsou dostatečně teoreticky ani empiricky podloženy. V zahraničí (např. ve Spojených státech) jsou tyto testy používány pro přijímání studentů k vysokoškolskému studiu. Existují doklady o tom, že studenti s vyšší úrovní těchto dovedností jsou úspěšnější při vysokoškolském studiu. Využití testu studijních dovedností k tomuto účelu má tedy dobrý smysl.

U nás jsou však výsledky testů používány k hodnocení žáků, škol a učitelů. A zde již je třeba maximální obezřetnosti: Nevíme, do jaké míry působí na rozvoj těchto testovaných dovedností škola a jak se přirozeně vyvíjejí s věkem. Máme málo dokladů o tom, do jaké míry je možno je systematicky rozvíjet cíleným cvičením, a zpravidla se na tyto dovednosti ve výuce také nezaměřujeme. Proto je pro nás obtížné interpretovat informaci o úrovni těchto dovedností, případně jejich nárůstu u našich žáků. Zůstává zde řada nezodpovězených otázek: Je tento nárůst způsobem tím, že je lépe vzděláváme? Nebo tím, že zestárlí? Je dobře, když rozdíl mezi obecnými dovednostmi a vědomostmi a dovednostmi v jednotlivých vzdělávacích oblastech je co největší, neboť to svědčí o tom, že jsme byli schopni žáky v dané oblasti maximálně rozvinout bez ohledu na jejich přirozené schopnosti? Nebo je naopak dobře, když je rozdíl co nejmenší, neboť to svědčí o tom, že jsme maximálně schopni rozvíjet u žáků tyto obecné dovednosti i v rámci předmětové výuky? Je vysoká úroveň obecných dovedností výsledkem naší práce nebo pouze genetickým vkladem našich žáků?

Tyto otázky si samozřejmě můžeme klást i v případě jednotlivých vzdělávacích oblastí, ani tam není zcela zřejmé, že zlepšení výsledků je důsledek naší práce, ale souvislost předmětových vědomostí a dovedností s výukou je přece jen více zřejmá. Z toho důvodu doporučujeme pečlivě zvážit, zda pro nás testování obecných dovedností má nějakou informační hodnotu a pokud se rozhodneme, že ano a test si objednáme, být maximálně obezřetní při interpretaci jeho výsledků.

Ještě diskutabilnější v tomto ohledu jsou tzv. testy všeobecného přehledu. Zde je zpravidla zjišťováno, zda žák zná vybrané fakty, které jsou tvůrcem testu považovány za důležité a vypovídající z hlediska všeobecného vzdělání. Hodnocení toho, co je a co není důležité, však může být velmi individuální. Zde by měl uživatel v každém případě vyžadovat podrobné zdůvodnění zacílení testu a jeho operacionalizace a měl by být maximálně obezřetný při posuzování validity.

Složení škol ve vzorku

Školy zpravidla získávají od organizátorů testování informaci o výsledcích žáků své školy a informaci o „celostátním průměru“. Tento „celostátní průměr“ je hodně ovlivněn tím, jaké školy jsou do toho

³ Norm-referenced, resp. criterion-referenced tests.

kterého testování zapojeny. Součástí zprávy o výsledcích testování by měla být podrobná informace o obecných charakteristikách ostatních škol (regionu, velikosti místa působnosti školy, velikosti školy, zaměření, v ideálním případě i složení žáků školy – charakterizované např. nejvyšším vzděláním rodičů). Na základě těchto informací si mohou učitelé vytvořit představu o tom, s kým se srovnávají. Analýzy podle typu škol by měly být rovněž samozřejmou součástí školních zpráv.

Na jaké otázky je třeba si odpovědět při výběru testu?

Následující formulář uvádí seznam otázek, na které bychom si měli odpovědět, než se rozhodneme pro využití testového nástroje. U každé otázky je uveden kurzívou vysvětlující komentář. Pokud bude více otázek zodpovězeno záporně, je třeba se opětovně zamyslet nad cílem evaluace a uvažovaným řešením.

Otázka, na kterou je třeba si odpovědět.	Ano/Ne
<p>1. Potřebuji pro daný evaluační účel srovnání výkonů mých žáků s výkony žáků z jiných škol?</p> <p><i>Například v situaci, kdy budeme chtít žákům poskytnout co možná nejkvalitnější zpětnou vazbu k jejich práci a podrobně určit úroveň jejich vědomostí a dovedností, zvolíme velmi pravděpodobně nějakou komplexnější evaluační úlohu, která nám umožní posoudit širokou škálu žakových vědomostí a dovedností. V situaci, kdy budu chtít vědět, jak moji žáci ovládají základní matematické operace ve srovnání s podobnými žáky v jiných školách, zvolíme test, který nám tuto komparaci umožní.</i></p>	
<p>2. Potřebuji ohodnotit vědomosti a dovednosti žáků týkající se širšího celku učiva?</p> <p><i>Autentické evaluační úlohy nám umožní získat hlubší vhled do dílčích vědomostí a zejména dovedností a poskytnout žákům kvalitní zpětnou vazbu, která jim pomůže se dále zlepšovat. Testy se hodí k (povrchnějšímu) hodnocení širšího učebního celku.</i></p>	
<p>3. Je tento test vhodný pro účel, pro který ho potřebuji? Je doporučená interpretace výsledků shodná s mými potřebami?</p> <p><i>Vždy bychom měli volit test, který je určen přesně k tomu účelu, pro který ho potřebujeme. Neměli bychom například používat přijímací test k dalšímu studiu k hodnocení práce žáka na nižším stupni. Přijímací test by měl být validizován ve vztahu k úspěšnosti žáka v dalším studiu, při ověřování testu by mělo být prokázáno, že žáci, kteří mají v testu dobré výsledky, jsou úspěšní v dalším studiu. To však nutně neříká nic o tom, kde své vědomosti a dovednosti získali, tedy jak pracovali na nižším stupni.</i></p>	
<p>4. Je dostatečně dobře popsáno, jaké vzdělávací cíle test ověřuje? Je dostatečně dobře zacílen, neklade si příliš velké ambice, neověřuje příliš mnoho cílů?</p> <p><i>Test musí být opatřen podrobným popisem cílů, jejichž dosažení ověřuje, a zdůvodněním jejich důležitosti. Záběr testu by neměl být příliš široký, neboť žák v omezeném testovacím čase může soustředěně zodpovídat pouze malé množství otázek. Příliš ambiciózní záběr testu vede na velmi povrchní a nespolehlivé měření.</i></p>	
<p>5. Poskytne mi tento test validní informaci o výsledcích mých žáků? Zjišťuje to, na co se ve výuce zaměřuji a co považuji za důležité?</p> <p><i>Vždy si musím položit otázku, zda cíle, které test ověřuje a které autor testu považuje za důležité, jsou důležité také pro mne, zda skutečně stojí v centru mé výuky. Pouze pokud je zacílení testu shodné se zacílením mé výuky, je možno uvažovat o použití testu k ověření výsledků mé práce.</i></p>	
<p>6. Je testový nástroj dostatečně kvalitní? Je dostatečně hodnověrně doloženo, že testové úlohy spolehlivě ověřují stanovené vzdělávací cíle? Má dostatečně vysokou reliabilitu?</p> <p><i>Autor testu musí být schopen doložit, jak přesně daný test ověřuje stanovené vzdělávací cíle a dokumentovat jeho kvalitu.</i></p>	

Otázka, na kterou je třeba si odpovědět.	Ano/Ne
<p>7. Je testový nástroj finančně dostupný? Jsou nároky na administraci přijatelné? Je přijatelná jeho časová náročnost?</p> <p><i>Čas, který věnujeme testování, bychom mohli využít k nějakým výukovým aktivitám. Proto bychom si měli vždy položit otázku, zda bude tento čas dobře využit, zda výsledky testování jsou pro nás a pro naše žáky natolik užitečné, že za tuto investici stojí.</i></p>	
<p>8. Je test určen pro žáky podobné těm, které vyučují? Byl na takových žácích pilotován?</p> <p><i>Autor testu by měl vždy poskytnout podrobnou informaci o tom, pro jaké respondenty je test určen, na jakých respondentech byl ověřován, případně standardizován. Například test určený pro dobře situované žáky z majoritní společnosti nemusí být kulturně vhodný pro romské žáky z vyloučených lokalit a může o úrovni jejich vědomostí a dovedností přinášet zkreslené informace.</i></p>	
<p>9. Budu mít dostatečně podrobnou informaci o tom, s kým jsou moji žáci porovnáváni? Je srovnání s ostatními zúčastněnými školami relevantní?</p> <p><i>V ČR jsou školám často nabízeny testy, které srovnávají výsledky jejich žáků s výsledky žáků z ostatních zúčastněných škol. Pro interpretaci výsledků je třeba mít dostatečné informace o tom, jaké další školy se testování zúčastnily. Je důležité vědět, s kolika školami jsou výsledky našich žáků porovnávány, zda se jedná o školy městské či venkovské, v jakých regionech se nacházejí, jaké je složení jejich žáků.</i></p>	
<p>10. Je zpráva, kterou obdržím, dostatečně informativní? Poskytne mi užitečné a dostatečně podrobné informace, na základě kterých budu moci modifikovat svoji výuku? Poskytne užitečné informace mým žákům?</p> <p><i>Při výběru testu je užitečné si učinit představu o podobě a obsažnosti zpětné vazby, která bude zapojeným školám a žákům zaslána. Tato informace je zásadní pro posouzení, zda bude investice do testu efektivní.</i></p>	

2. Autentické evaluační nástroje

Považujeme za důležité upozornit na to, že testy nejsou zdaleka jedinými nástroji, které by mohly školy od externích agentur poptávat. V současné době jsou instituce zabývající se evaluací vyhledávány zejména proto, že poskytují testy a umožňují komparaci výsledků žáků jednotlivých tříd a škol s výsledky žáků z ostatních škol, které jsou do evaluační aktivity zapojeny. Učitelé by nicméně mohli požadovat i další typy evaluačních nástrojů, které tuto komparaci v principu také umožňují. Typickým pomocníkem pro každodenní práci učitelů jsou jednotná zadání úloh s návody na jejich vyhodnocení.

Tyto nástroje mají tu výhodu, že umožňují postihnout širší spektrum žákovských vědomostí a dovedností, než postihují tradiční testy, a sledovat jejich vývoj. Rovněž mohou být voleny tak, aby daleko lépe odpovídaly výuce, tedy tomu, co se žáci aktuálně ve škole učí. Pro profesionální agentury zabývající se evaluací není zároveň problém nástroje standardizovat a poskytnout tak učitelům informaci o tom, jaké jsou hodnocené vědomosti a dovednosti jejich žáků ve srovnání s celostátním průměrem, respektive se specifickými skupinami žáků a škol.

Učitelé si samozřejmě v principu mohou tvořit podobná zadání sami. Na webových stránkách amerických a britských organizací je možno nalézt mnoho pěkných příkladů zadání úloh a zejména schémat pro jejich vyhodnocování⁴. Tvorba je však časově náročná a vyžaduje určitou zkušenost a samostatná práce učitele neumožňuje srovnání s jinými školami, tedy mu neposkytuje dostatečnou zpětnou vazbu o jeho práci. Proto se domníváme, že by se evaluační agentury měly zaměřit i na tento typ služeb, respektive že by učitelé měli od nich tento typ služeb vyžadovat. Pro lepší představu o tom, jaký typ nástrojů máme na mysli, uvádíme v příloze ukázkou dvou zadání pro společenskou vědu a návody pro vyhodnocování žákovských odpovědí. Uvedené úlohy umožňují při vyhodnocování jistou subjektivitu, ale dobře ilustrují šíři dovedností, které je možno s jejich pomocí hodnotit. Návody by bylo samozřejmě ještě možno konkretizovat a jejich objektivitu zvýšit. Dalším příkladem úloh, které by mohly posloužit učitelům k objektivnímu posouzení vědomostí a dovedností

⁴ V angličtině jsou schémata pro hodnocení vědomostí a dovedností zpravidla označována jako „rubrics“.

svých žáků, jsou čtenářské úlohy⁵. Pěknou ukázkou takové úlohy je možno nalézt na stránkách Ústavu pro informace ve vzdělávání: <http://www.uiv.cz/soubor/1266>.

Na jaké otázky je třeba si odpovědět při výběru autentické evaluační úlohy?

Autentické evaluační úlohy jsou zpravidla časově náročné jak pro učitele, tak pro žáka. Proto je třeba je vybírat pečlivě tak, aby čas s nimi strávený byl pro žáky co nejužitečnější a aby byla co nejužitečnější informace, kterou z nich žák i učitel získají. Tyto úlohy mohou být zadávány primárně jako zpětná vazba pro žáky, v tom případě není zásadní maximální objektivita hodnocení, je však důležité, aby úloha byla pro žáky maximálně stimulující. Pokud je hlavním smyslem evaluace srovnat výsledky žáků s nějakými národními standardy nebo s výsledky žáků jiných škol, je důležitá podrobnost a jednoznačnost návodu na vyhodnocování, která určuje objektivitu vyhodnocování žákovských odpovědí.

Učitel, který volí evaluační úlohu, by si měl při jejím výběru odpovědět na následující otázky. U každé otázky je uveden kurzívou vysvětlující komentář. Pokud bude více otázek zodpovězeno záporně, je třeba se opětovně zamyslet nad cílem evaluace a uvažovaným řešením.

⁵ V angličtině bývají označovány jako reading inventories.

Otázka, na kterou je třeba si odpovědět.	Ano/Ne
<p>1. Vyžaduje vyřešení této úlohy vědomosti a dovednosti, které tvoří důležitou součást kurikula?</p> <p><i>Jedná-li se o větší časovou investici, je třeba ji zacílit na vědomosti a dovednosti, které jsou základní a měly by si je z výuky odnést všichni žáci. Musí se jednat o vědomosti a dovednosti, které jsou ve výuce rozvíjeny a na které je kladen velký důraz.</i></p>	
<p>2. Stojí úloha za čas a energii, která je nezbytná k jejímu řádnému vyřešení?</p> <p><i>I když je úloha zacílena na základní vědomosti a dovednosti, může být neefektivní: žáci při jejím řešení mohou ztrácet čas zbytečnostmi, které je nijak nerozvíjejí a jejichž informační hodnota je minimální.</i></p>	
<p>3. Je úloha pro žáky dostatečně náročná a motivující?</p> <p><i>Řešení úlohy by mělo být pro žáky intelektuálně stimuluující.</i></p>	
<p>4. Pomůže zpětná vazba, kterou získají v průběhu práce na úloze, žákům vyhodnotit svoji práci a přizpůsobit ji danému úkolu?</p> <p><i>Otázky vycházejí z předpokladu, že by evaluace měla být v první řadě užitečná pro učení žáků. A také z předpokladu, že při řešení evaluační úlohy jsou žáci motivováni se učit. Tato otázka předjímá, že žákům bude v průběhu řešení úlohy poskytnuta podpora, která jim umožní rozpoznat slepé uličky a své řešení přehodnotit. Učitel se samozřejmě může rozhodnout, že v průběhu řešení podporu poskytovat nebude a poskytne zpětnou vazbu až na základě celkového výsledku.</i></p>	
<p>5. Pomohou žákům známky, které dostanou za danou úlohu, a s nimi spojené informace získat zpětnou vazbu a zlepšit jejich práci?</p> <p><i>Vzhledem k tomu, že se jedná o evaluační úlohu, je odpověď na tuto otázku zcela zásadní. Vyhodnocení úlohy musí žákům poskytnout důležité informace o tom, co již zvládají a kde by se měli ještě zlepšit a jak.</i></p>	
<p>6. Je návod na vyhodnocování žákovských odpovědí dostatečně podrobný a jednoznačný?</p> <p><i>Podrobnost a jednoznačnost návodu na vyhodnocování odpovědí je důležitá v případě, že bude úspěšnost žáků v úlohách srovnávána s normami nebo s výsledky žáků z jiných škol.</i></p>	
<p>7. Je úloha určena pro žáky podobné těm, které vyučují? Byla na takových žácích pilotována?</p> <p><i>Autor úlohy by měl vždy poskytnout podrobnou informaci o tom, pro jaké respondenty je úloha určena, na jakých respondentech byla ověřována. Je důležitý nejen věk žáků, typ vzdělávacího programu, ale i jejich rodinné zázemí. Pokud by byla úloha ověřována pouze na žácích určité skupiny, mohla by ve vztahu k jiným skupinám poskytnout zavádějící informaci.</i></p>	
<p>8. Budu mít dostatečně podrobnou informaci o tom, s jakým výkonem je výkon mých žáků porovnáván? Byla úloha standardizována na dostatečném počtu dostatečně různorodých žáků? Je srovnání s ostatními zúčastněnými školami relevantní?</p> <p><i>Pokud je pro nás důležité porovnání výsledků našich žáků v úloze s výsledky nějakých celorepublikových norem, musíme mít jistotu, že tyto normy jsou skutečně vytvořeny na vzorku žáků, který dobře reprezentuje žáky dané věkové kategorie v ČR. Stejně tak provádíme-li porovnání s žáky několika škol, musíme mít vědět, o jaké školy se jedná, zda jejich výběr je dostatečně reprezentativní.</i></p>	

Literatura

Standardy pro pedagogické a psychologické testování.2001. Testcentrum: Praha.

Košťálová, H., Straková, J. (Eds). 2008. Hodnocení, důvěra, dialog, růst. SKAV: Praha.

Příloha – Ukázka úloh pro hodnocení v dějepise

Historický projekt (skupinová a samostatná práce) – 1. ročník střední školy

Žáci jsou rozděleni do malých skupin, ve kterých analyzují primární a sekundární zdroje, které vyjadřují různé pohledy na danou historickou událost. Po přečtení a prodiskutování materiálů každá skupina zaujme nějaký postoj, který prezentuje zbytku třídy. Učitelé a žáci poté dávají skupině otázky týkající se jejich postoje. Následně mají žáci za úkol napsat samostatně krátký text a poté reakci na konkrétní aspekt daného problému, který od žáků vyžaduje podpořit nebo naopak odmítnout nějaký názor. Žáci někdy mohou svůj názor prezentovat i jiným způsobem než formou textu (např. prostřednictvím komiksu).

Celá práce je ohodnocena 100 body. Body se přidělují podle následujících oblastí:

Skupinové učení a spolupráce (20 bodů)

17-20 výjimečný výkon

- skoro všichni žáci se nadšeně účastní
- všichni cítí zodpovědnost za úkol
- žáci respektují a vnímají názory ostatních a odkazují se na ně
- otázky a odpovědi vypovídají o dobré přípravě a promyšlenosti

13-16 výborný výkon

- žáci vykazují zběhlost ve spolupráci
- aspoň tři čtvrtiny žáků se aktivně účastní
- probíhá živá diskuze související s úkolem

8-12 chvalitebný výkon

- žáci vykazují určitou schopnost spolupráce
- alespoň polovina žáků diskutuje a říká své nápady
- probíhá pozorné čtení dokumentů a naslouchání
- alternativy jsou částečně prodiskutovány

5-9 základní výkon

- silná závislost na jednom mluvčím
- jen jeden nebo dva žáci se aktivně účastní
- spolupráce je sporadická
- konverzace není zcela zaměřena na téma

1-4 minimální výkon

- velmi silná závislost na jednom mluvčím
- malá spolupráce
- velmi krátké konverzace
- někteří žáci projevují nezáměr a nesoustředěnost

Kritické myšlení (30 bodů)

25-30 výjimečný výkon

- prokáže jasné a správné pochopení problému a s ním souvisejících aspektů v celé šíři
- k vytvoření postoje využívá všechny informace z výchozích materiálů a rozsáhlé osobní znalosti, které souvisejí s problémem, jsou správné, přesné a konzistentní
- úsudek zakládá na důkladném prozkoumání faktů

19-24 výborný výkon

- *prokáže jasné pochopení problému v celé šíři a aspoň dvou hlavních aspektů souvisejících s problémem*
- *k vytvoření postoje využívá nejdůležitější informace z výchozích materiálů a osobní znalosti, které souvisejí s problémem a jsou konzistentní*
- *úsudek zakládá na prozkoumání informací a částečném zvážení důsledků*

13-18 chvalitebný výkon

- *prokáže rámcové pochopení rozsahu problému a nejméně jednoho souvisejícího aspektu*
- *k vytvoření postoje využívá nejdůležitější informace z výchozích materiálů a některé vlastní myšlenky*
- *úsudek zakládá na prozkoumání informací a částečném zvážení důsledků*

7-12 základní výkon

- *prokáže jen velmi rámcové porozumění problému*
- *zaměřuje se pouze na jedno téma/aspekt*
- *využívá pouze poskytnuté informace*
- *názor zakládá pouze na názoru, nikoli na faktických informacích*
- *úsudek zakládá na částečném prozkoumání faktů s malým zohledněním důsledků*

1-6 minimální výkon

- *prokáže malé porozumění problému nebo jeho aspektů*
- *využívá jen ty nejzákladnější části poskytnutých informací*
- *při vytváření vlastního postoje míchá dohromady fakta a názory*
- *úsudek zakládá pouze na zohlednění jednoho nebo dvou aspektů celého sdělení*
- *chybí zohlednění důsledků*

Sdělování myšlenek (20 bodů)

17-20 výjimečný výkon

- *zaujme výrazný přesně definovaný postoj*
- *argumentace je systematická a přesvědčivá a jasně doložená fakty*
- *zabývá se všemi aspekty problému a prokáže, že do hloubky rozumí důležitým souvislostem*
- *zkoumá problém z různých úhlů pohledu*

13-16 výborný výkon

- *zaujme jasný postoj*
- *argumentace je systematická, případně s malými chybami v předložených faktech*
- *zabývá se hlavními aspekty a prokáže jisté porozumění souvislostem*
- *zauvažuje také nad zkoumáním více než jednoho úhlu pohledu*

8-12 chvalitebný výkon

- *zaujme definovaný, ale ne příliš vyhraněný postoj*
- *argumentace je méně systematická*
- *užívá obecné/nekonkrétní názory a doložení fakty je nedostačující*
- *zabývá se jen několika aspekty*
- *vidí problém jen z jednoho pohledu*

5-9 základní výkon

- zaujme nedefinovaný a nekonkrétní postoj
- prezentace myšlenek je chaotická, nesystematická
- svůj postoj dostatečně nedoloží, používá všeobecná tvrzení
- zdůrazňuje pouze jeden aspekt problému

1-4 minimální výkon

- postoj je neurčitý
- prezentace je stručná a zahrnuje všeobecná sdělení nesouvisející s tématem
- celkový pohled na problém je nejasný
- sdělení odbíhají se od tématu

Využití historických znalostí (30 bodů)

25-30 výjimečný výkon

- vykoná přesnou analýzu výchozích informací a tématu
- hlavní i dílčí témata a pojmy pojednává s využitím řady faktických údajů
- k důkladnému pochopení problému využívá dříve nabytých vědomostí a na jejich základě jej propojuje s minulými, případně budoucími událostmi

19-24 výborný výkon

- vykoná přesnou analýzu výchozích informací
- hlavní témata pojednává s využitím faktických údajů
- ke zkoumání daných témat využívá všeobecné znalosti historie

13-18 chvalitebný výkon

- uspokojivě propojuje hlavní fakta se základními tématy
- aspoň jedno téma nebo koncept dokládá fakty
- uspokojivě využívá všeobecné znalosti historie

7-12 základní výkon

- prezentuje jen základní fakta s jistou mírou nepřesnosti
- využívá informace pro obecné vysvětlení aspoň jednoho tématu nebo myšlenky
- nevyužívá dostatečně historické znalosti
- opírá se hlavně o dodané informace

1-6 minimální výkon

- opakuje jeden nebo dva nepřesné údaje
- tématy a myšlenkami se zaobírá pouze nepřesně a stručně
- historické znalosti nepoužívá skoro vůbec
- opírá se výlučně o dodané informace

Zpracování dějepisného tématu formou referátu (základní škola)

Znalosti

4 Problém vysvětlí jasně, logicky a konsistentně. Jasně vymezí téma práce.

Shromáždí vhodná a relevantní fakta a jasně je strukturuje.

3 Problém vysvětlí důkladně a více méně přesně. Vhodně vymezí téma práce. Shromážděná fakta jsou správná, ale je jich málo

2 Částečně vysvětlí problém, ale pomine některé klíčové prvky. Zadání práce není vymezeno správně. Fakta jsou zmatená a neuspořádaná

1 Téma není vysvětleno vůbec. Nedostatečně vymezené zadání. Základní informace je nedostatečná

Kvalita prezentace

4 Téma práce odpovídá zadání a je prezentováno nápaditě. Práce je velmi kreativní, je věcná, smysluplná, výstižná. Prezentace je plynulá, má logickou strukturu. Na obrazových materiálech je patrná důkladnost, důraz na detaily jako překlepy, úprava.

3 Téma práce souvisí se zadáním, práce je zajímavá a smysluplná. Prezentace je v zásadě plynulá. Obrazové materiály pěkně vypadají. Mírně nedostatečný důraz na detaily jako překlepy, úprava apod.

2 Téma práce příliš nesouvisí se zadáním. Sdělení je nejasné a chybějí podstatné prvky. Prezentace není smysluplná. Chyby v detailech.

1 Téma nesouvisí se zadáním, postrádá srozumitelnost. Velké chyby v detailech

Kvalita analýzy

4 Práce je jasná, smysluplná, věcná, cílená. Má závěr. Jsou použity výstižné argumenty. Závěry jsou přesvědčivé a založené na důkazech.

3 Práce je smysluplná. Jasně formuluje problémy a řeší je. Argumenty jsou do značné míry uspořádané a výstižné. Závěry jsou celkem přesvědčivé

2 Práce není dostatečně smysluplná, neřeší jasně vymezený problém. Argumenty jsou zmatené a nejasné. Řešení problému je nedostatečné.

1 Práce není smysluplná, neformuluje problém. Argumenty nejsou podloženy fakty. Chybí řešení.

Zdroje

4 V práci bylo využito neobvyklých zdrojů. Bylo konzultováno několik institucí. Literatura obsahuje množství primárních i sekundárních zdrojů.

3 Použití dostatečných zdrojů, konzultována jedna nebo dvě instituce. Literatura obsahuje dostatečné množství primárních i sekundárních zdrojů.

2 Chybí dostatečné zdroje. Konzultována byla aspoň jedna instituce. Literatura obsahuje nedostatečné množství primárních a sekundárních zdrojů

1 Žádné zdroje nebo nedostatečné, nebyly konzultovány žádné instituce.