

Stálá komise pro testy a testování při EFPA¹

Tomáš Urbánek

Psychologický ústav AV ČR, v.v.i., Brno
tour@psu.cas.cz

Abstrakt

EFPA Standing Committee on Tests and Testing (Stálá komise pro testy a testování při Evropské federaci psychologických asociací) je odborně-profesním tělesem, sdružujícím zástupce jednotlivých členů EFPA. Jak napovídá název, do agendy této komise spadají otázky související s tvorbou, adaptací a používáním psychologických testů v zemích, ve kterých působí členové EFPA – odborné a profesní psychologické společnosti zastupující jednotlivé státy.

V minulosti tato komise řešila otázky nutných kompetencí uživatele psychologického testu, otázky evropských standardů pro testy a testování, návrh modelu recenze testu jako součásti kontrolních mechanismů napomáhajících zkvalitnění testové psychodiagnostiky atd.

V České republice zatím neexistuje podobná skupina, která by se otázky testování a diagnostiky z odborného a profesního hlediska zabývala. Součástí příspěvku je návrh na ustanovení fóra, které by sloužilo k zveřejňování materiálů SCTT odborné veřejnosti a diskusi o těchto otázkách.

Klíčová slova: testy a testování, recenze testu, standardy užívání testu, certifikace uživatele testu

Úvodem

Každý psycholog v České republice, který je členem některé z řady odborných a profesních společností sdružených v Unii psychologických asociací ČR, je prostřednictvím této Unie také členem Evropské federace psychologických asociací (European Federation of Psychologists' Associations, zkráceně EFPA). Tato organizace v současné době sdružuje celkem 34 členských států. O jejich cílech, stanovách, etických zásadách, aktivitách a dalších důležitých tématech je možné si přečíst na internetu (viz www.efpa.eu). Z důvodu omezeného rozsahu tohoto příspěvku je možné uvést aspoň to, že dokumenty EFPA jsou výsledkem diskuse představitelů evropské psychologie a do značné míry tak vypovídají o praxi převládající v jednotlivých členských zemích. Doporučení vydávaná touto federací jsou založena na standardech, na kterých se shodnou přední evropští a světoví odborníci nejen z akademické, ale i aplikované sféry, a představují dost často pravidla a postupy, které jsou přijaté a uplatňované v některé z evropských zemí, nebo jejich kombinaci.

¹ Článek vznikl jako součást výzkumného záměru Psychologického ústavu AV ČR, v.v.i., reg. číslo: AV0Z70250504.

Jedním ze způsobů práce běžných v podobných společnostech (podobně jako je tomu např. v Česko-moravské psychologické společnosti) je zakládání tematicky zaměřených pracovních skupin. Ty jsou v EFPA nazývány stálé komise (Standing Committee) a je jich v současné době pět. Jednou z nich je také Stálá komise pro testy a testování (Standing Committee on Tests and Testing, zkr. SCTT), jejímž členem jsem se stal jako zástupce ČR v roce 2005 v Granadě, kde tato komise zasedala v rámci programu Evropského kongresu psychologie. Tento příspěvek je pokusem přiblížit moje zkušenosti z této práce a poskytnout psychologické veřejnosti u nás základní informace o tom, čím se tato komise zabývá, popř. vyvolat diskusi o těchto otázkách.

Stálá komise pro testy a testování při EFPA (dále SCTT) není jediným odborně-profesionálním tělesem, které se otázkami testování zabývá. V některých národních psychologických společnostech existují obdobné pracovní skupiny jako SCTT a na celosvětové úrovni existuje Mezinárodní testová komise (International Test Commission), která existuje delší dobu (od roku 1978) a jejími hlavními výstupy jsou např. doporučení týkající se správných postupů při převodu diagnostických metod (Van de Vijver, Hambleton, 1996) nebo doporučení týkající se počítačového a přes internet distribuovaného testování (viz www.intestcom.org). Kromě toho se otázkám psychodiagnostiky věnují odborné společnosti jako např. European Association of Psychological Assessment (EAPA) a další.

Muñiz et al. (2001) provedli studii u několika tisíc evropských psychologů převážně ze Španělska, Velké Británie a Nizozemska (celkem více než 7000), a dalších několika stovek psychologů ze Slovinska, Chorvatska a Belgie (celkem přes 700). V odpovědích respondentů se mimo jiné objevoval názor, že odborné a profesní instituce by měly aktivněji prosazovat správnou praxi testování. To se v mnoha zemích skutečně děje, přičemž se uplatňují dvě základní strategie – jednou z nich je snaha omezit přístup k testům (umožnit ho pouze členům určitých skupin) a druhou je snaha zvyšovat kvalifikaci uživatelů testů (Muñiz et al., 2001). Omezení přístupu k testům na základě získané kvalifikace je často inspirována klasifikací testů po vzoru APA na tři úrovně. Toto dělení je zhruba následující: úroveň A tvoří výkonové a znalostní testy, úroveň B testy schopností a inteligence úroveň a C individuálně používané testy inteligence a osobnosti a další komplexní testy, např. testy projektivní. V některých zemích je oprávnění používat tyto testy spojeno s dosažením určité úrovně akademického vzdělání v oboru psychologie, jinde (jako např. ve Velké Británii) je pro získání oprávnění pro používání určitého testu nutné získat oprávnění a seznamy akreditovaných uživatelů jednotlivých testů jsou spravovány (a aktualizovány) v rámci Britské psychologické společnosti (British Psychological Society - BPS).

To už je ale příklad druhé strategie, jak zvyšovat úroveň testování. Fremer (1996) cituje jakési ústní vystoupení přední představitelky teorie testování A. Anastasiové v Marylandu v roce 1987, podle níž nejčastější příčinou nesprávného používání testů je nedostatek znalostí (další dva nejčastější důvody jsou snaha zkreslit výsledky a nedbalost). Nedostatek znalostí je naštěstí možné napravit vzděláváním a jeden z možných nástrojů kontroly, který navrhuji někteří autoři, je pravidelná certifikace uživatelů testů.

V neposlední řadě je nutné zmínit také hodnocení samotných testů a dotazníků. V České republice máme k dispozici český překlad amerických Standardů pro pedagogické a psychologické testování (AERA, APA, NCME, 2001).

V některých evropských státech (např. ve Velké Británii, Nizozemí a Španělsku) i v USA² se testy recenzují a model recenze vycházející z této praxe je jedním z výstupů SCTT, který bych chtěl prezentovat v následující části příspěvku.

Hlavní aktivity komise v posledních letech

V předchozích odstavcích jsem se pokusil představit kontext, ve kterém členové SCTT pracují. Hlavním cílem komise je tedy zlepšovat praxi používání testů, a to tím, že se stanoví minimální požadavky týkající se testů samotných, ale uživatelů testů a postupů, které při testování uplatňují. Uvedu stručně pouze hlavní výstupy SCTT za posledních několik let, které se vztahují k těmto cílům. Jedná se o Evropský model recenze testu (EFPA, 2005), Evropské standardy uživatele testu pro oblast psychologie práce a organizace (EFPA a EAWOP, 2005) a Evropské certifikáty pro používání psychologických testů v oblasti práce a organizace (zatím v 3. neveřejné verzi návrhu).

Model recenze testu

Na rozdíl od Standardů pro pedagogické a psychologické testování (AERA, APA, NCME, 2001), které stanoví obecné požadavky na testy v podobě jakéhosi statického „zákoníku“, kladou autoři Evropského modelu recenze testu (EFPA Review Model for the Description and Evaluation of Psychological Tests, EFPA, 2005) důraz spíše na proces a obsah recenze. To ale neznamená, že by tyto dva dokumenty byly v rozporu. Evropský model recenze testu (dále ERM) ale naznačuje, jak konkrétně postupovat při hodnocení testu, a je tedy spíše nástrojem než standardem. Modelem se budu zabývat ze dvou hledisek – procesuálního a obsahového.

Proces recenze

Každý test je podle ERM recenzován dvěma nezávislými recenzenty (reviewers), jejichž recenze kontroluje odborný redaktor (consulting editor) a na celý proces recenzování dohlíží vedoucí redaktor (senior editor), který dbá především na zajištění jednotnosti kritérií používaných v recenzích. V návrhu ERM se píše, že každá členská země si tyto funkce může stanovit na základě lokálních podmínek, a to i tak, že se budou personálně překrývat (EFPA, 2005).

² Na tomto místě je vhodné aspoň zmínit existenci ústavu, který je součástí University of Nebraska-Lincoln. Je jím Buros Institute of Mental Measurement (viz <http://www.unl.edu/buros/>), který už více než 70 let vydává známé ročenky obsahující základní informace o psychodiagnostických nástrojích a v současné době nabízí na svých internetových stránkách k prodeji (kromě materiálů, které jsou volně dostupné) množství recenzí psychodiagnostických metod. Jedná se ale o americké verze těchto metod a Evropský model recenze testu je v našich podmínkách užitečnější, protože je určen pro hodnocení lokálního převodu dané metody.

Obsah recenze

Výsledná recenze by měla být zaměřena na „kvalifikované praktické uživatele testu“, počítá se ale i se zájmem „akademiků a specialistů na psychometriku a psychologické testování“ (EFPA, 2005, str. 1).

Model recenze má s poznámkami a vysvětlujícím textem celkem 29 stran, z toho důvodu zde bude uveden jen stručný popis. Většina obsahu má podobu strukturovaného formuláře (v podstatě komplexního dotazníku), ve kterém mají recenzenti stanoveny konkrétní položky, na které mají odpovídat. Položky jsou nejrůznějšího formátu, jaký si lze představit – od volných položek, ve kterých recenzent píše svůj hodnotící text, přes položky typu „multiple-choice“ až po ordinální hodnotící škály.

Kromě obecného popisu recenzovaného testu a údajích o recenzentech a redaktorech (oddíl 1) je součástí recenze klasifikace testu (oddíl 2) týkající se jeho zaměření, zamýšlené oblasti použití, zamýšleného způsobu použití, populace, pro kterou je nástroj určen, detailního popisu obsahu z hlediska škál, počtu a formátu položek, způsobů administrace včetně možnosti existence různých forem, časových nároků atd. Další údaje (oddíl 3) se týkají způsobů skórování a následné transformace hrubých skórů na skóry odvozené. Oddíl 4 se týká výstupů generovaných počítačem (pokud jsou součástí testu), např. jejich formy, propracovanosti, struktury, citlivosti vůči kontextu vyšetření, modifikovatelnosti, míry jejich dokončenosti, transparentnosti (tzn. do jaké míry je jasné, jak byl takový výstup vytvořen na základě získaných výsledků), stylu a zamýšleného příjemce. V oddíle 5 hodnotí recenzenti to, jak je test dodáván, co je součástí testových materiálů, jaké jsou pořizovací náklady a jaké jsou náklady na nákup dalšího materiálu, popř. na zpracování dat na serverech vydavatele. Hodnotí se zde také, jakou kvalifikaci požaduje dodavatel testu od toho, kdo test nakupuje, a jaké jsou profesionální požadavky na uživatele testu.

Další oddíly tvoří podle mého názoru jádro hodnocení nástroje. Oddíl 6 se zabývá hodnocením testových materiálů z různých hledisek. Jednak jsou zde materiály samotné, jak je publikoval vydavatel testu. Dále jsou zde informace publikované v odborné literatuře. Další zdroje mohou být dosud nepublikované materiály, které má vydavatel testu k dispozici a poskytne je pro účely recenze, a konečně důvěrné materiály vydavatele, které mohou být v některých případech důležité pro provedení recenze, a které může vydavatel poskytnout recenzentovi za přesně stanovených podmínek. Pokud recenzent hodnotí testový materiál na základě zdrojů nedostupných běžnému uživateli, měl by to v recenzi uvést. Samotné hodnocení v této sekci se týká např. teoretického vymezení měřeného konstruktů, procesu vývoje nástroje, detailnosti analýzy položek, vysvětlení obsahové validity, shrnutí dosavadního výzkumu, kvality standardizace, norem, reliability, instrukcí atd.

V oddíle 7 jsou hodnoceny normy, reliabilita a validita metody. Je zde poměrně rozsáhlá poznámka týkající se hlavních psychometrických zásad, které je nutné při zpracovávání této části mít na zřeteli. Hodnotí se adekvátnost těchto technických informací o testu, vhodnost norem, rozsah výběrového souboru, posuzuje se možnost případných odlišností např. ve skupinách menšin, vlivů věku a pohlaví atd. Podobná hlediska se uplatňují při hodnocení validity testu, tzn. způsobů, jakými byla zkoumána, jaké hodnoty byly zjištěny. Pochopitelný důraz je zde kladen především na konstruktové a empirické důkazy o validitě. Podobně se postupuje u hodnocení reliability.

V oddíle 8 je hodnocena kvalita počítačově generovaných výstupů, pokud je test poskytuje, a to z hlediska jejich relevance, spravedlnosti, přijatelnosti pro příjemce, rozsahu a možnosti, aby do nich

uživatel testu zasáhl a upravil si je podle svých představ. Poslední oddíl 9 podává závěrečné hodnocení v podobě závěrečné hodnotící zprávy a bodových doporučení týkajících se možného používání recenzovaného nástroje (podrobnosti viz EFPA, 2005).

V současné době připravuji český překlad Evropského modelu recenze testu, který bych rád zveřejnil např. na internetových stránkách Českomoravské psychologické společnosti.

Evropské standardy uživatele testu v oblasti práce a organizace

Tento materiál (European Test User Standards for test use in Work and Organization settings, EFPA a EAWOP, 2005) byl vytvořen pracovní skupinou složenou ze zástupců EFPA a EAWOP (European Association of Work and Organizational Psychologists). Jak je z názvu patrné, jedná se o standardy, které by měly zajímat především pracovní psychology. Motivace pro tvorbu těchto materiálů je dána především faktem, že psychologické testy jsou v Evropě ve značné míře používány ne-psychology. Standardy pro uživatele testu pak nastavují určitou úroveň, kterou by měl každý uživatel splňovat, aby bylo možné např. řešit sporné případy zneužití metod.

Text Standardů je opět poměrně rozsáhlý (26 stran), takže zde budou pouze uvedeny jednotlivé standardy s krátkým komentářem.

Tabulka č. 1: Evropské standardy pro používání testů v oblasti práce a organizace

Část 1:	Přijmout zodpovědnost za etické použití testu
Standard:	1.1 Chovat se profesionálním a etickým způsobem
Standard:	1.2 Zajistit si kompetenci pro používání testů
Standard:	1.3 Přijmout zodpovědnost za vlastní používání testů
Standard:	1.4 Zajistit bezpečné uložení testových materiálů
Standard:	1.5 Zajistit důvěrnost zacházení s výsledky testů
Část 2:	Postupovat při používání testů správným způsobem
Standard:	2.1 Vyhodnotit potenciální užitečnost testování v diagnostické situaci
Standard:	2.2 Vybrat testy vhodné v dané situaci
Standard:	2.3 Dobře uvážit otázky spravedlnosti (rovnosti) při testování
Standard:	2.4 Správně analyzovat a interpretovat výsledky
Standard:	2.5 Jasně a přesně sdělit výsledky relevantním příjemcům
Standard:	2.6 Kriticky posoudit vhodnost testu a jeho použití

Část 3:	Postupovat správným způsobem při administrování testů
Standard:	3.1 Učinit nutné přípravy pro testové sezení
Standard:	3.2 Administrovat test správně
Standard:	3.3 Skórovat výsledky testu přesně

V případě každého standardu je brán zřetel ke kontextu, v jakém je nutné o daném standardu uvažovat. Jedná se např. o zadání, které dostal uživatel testu od zadavatele úkolu, který má splnit, o obecných teoriích, modelech (např. teoretických, praktických) a principech (např. etických), v souladu s nimiž musí pracovat na splnění tohoto úkolu, faktech a datech (např. pravidla a zákony), které musí brát v potaz, a metodách, popř. konkrétních postupech, které musí použít. V mnoha případech se také jedná o základních znalostech a dovednostech (např. z konkrétních oborů psychologie, psychometrie apod.), které musí uživatel testu ovládat, aby mohl konkrétní standard splnit.

Procedura certifikace uživatele testů v oblasti psychologie práce a organizace

V současné době členové SCTT pracují na návrhu procedury, kterou musí projít uchazeč, který chce získat oprávnění k používání psychologických testů v oblasti práce a organizace (tzv. *EuroTest*). Procedura, která byla navržena v době posledního jednání komise v červenci 2008 v Berlíně (jedná se zatím o 3. nepublikovanou verzi), navazuje přímo na výše uvedené Evropské standardy uživatele testu v oblasti psychologie práce a organizace. Získaný certifikát by měl mít tři úrovně podle náročnosti, zatím poměrně volně definované, a výsledkem pro uchazeče by bylo získání certifikátu a zveřejnění jména v seznamu spravovaném příslušnou komisí. Takový certifikát by platil omezenou dobu (v návrhu je nyní 7 let). Po skončení platnosti bude certifikát nutné obnovit na základě dokladů o pokračující profesionální praxi a trvajících kompetencích.

Cílem je tedy vytvořit takovou proceduru, aby komise pověřená udělováním tohoto certifikátu na národní úrovni mohla postupovat jednotně v celé Evropě. Dá se totiž očekávat pohyb psychologů v rámci Evropy a profesní organizace by měly uchazeče hodnotit na základě stejných nebo podobných kritérií a hlídat vysoký standard používání testů. Očekává se přitom, že to, co začíná být v současné době aktuální v oblasti psychologie práce, může být za nějakou dobu zajímavé i pro další oblasti psychologické praxe.

Nejde přitom o snahu nahradit stávající praxi v jednotlivých státech Evropy, pokud se osvědčila. Z toho důvodu jsou jednotlivá pravidla formulována dostatečně volně, aby je bylo možné např. v České republice přizpůsobit existující praxi.

Důsledky pro psychologickou praxi v ČR

Nyní je nutné zamyslet se nad tím, co zmíněný vývoj na úrovni Evropy bude znamenat pro Českou republiku. Bez ohledu na to, že tyto iniciativy jsou podle mého názoru užitečné a prospěšné pro udržení a zlepšování kvality poskytovaných psychologických služeb, je pravděpodobné, že v několika následujících letech se u nás nebude z tohoto hlediska dít nic zásadního. Pokud se ale certifikace uživatelů testů (a s nimi spojené standardy uživatele) v Evropě prosadí a Česká republika bude atraktivní pro psychology z jiných částí Evropy nebo světa, dá se zřejmě očekávat, že přicházející psychologové budou žádat o udělení tohoto certifikátu. Pak bude žádoucí, aby tyto postupy byly uplatňovány, přinejmenším jako určitá forma ochrany úrovně kvality praxe v naší zemi.

Pokud se výše zmíněný proces certifikace a standardy v Evropě prosadí, bude rozhodně žádoucí učinit je součástí vzdělávání psychologů. Spíše se dá totiž očekávat, že naopak čeští psychologové budou mít zájem o práci v zahraničí, a znalost evropských standardů a procedur jim v takovém případě bude velmi k užitku. Další hledisko spočívá v tom, že sami zaměstnavatelé a zákazníci psychologů, především z řad velkých firem, by mohli začít zmíněné certifikáty a standardy vyžadovat a jejich získání by pak představovalo konkurenční výhodu, ne-li přímo existenční nutnost.

Tento spíše letargický závěr ale neplatí pro Evropský model recenze testu. V jeho případě se domnívám, že je vysoce žádoucí, aby se začal používat co možná nejdříve a začaly se publikovat recenze testů – zatím aspoň těch nejpoužívanějších, ale postupně pokud možno všech testů, které se používají v České republice. Je podle mého názoru žádoucí, aby takovou aktivitu podpořila Českomoravská psychologická společnost, popř. další profesní organizace, např. ustavením pracovní skupiny pro testy a testování, z jejichž členů by se mohli rekrutovat redaktoři recenzí a samotní recenzenti.

Ustavení takové pracovní skupiny je žádoucí také z toho důvodu, že národní zástupce v SCTT (ať už to bude kdokoli) potřebuje nějakou skupinu kolegů, s nimiž bude o těchto tématech diskutovat, a se kterými se bude radit v případech, kdy bude SCTT žádat o názor jednotlivých národních zástupců, tzn. národní odborné komunity.

Literatura

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2001). *Standardy pro pedagogické a psychologické testování*. Praha: Testcentrum.

European Federation of Psychologists' Associations (2005). *EFPA Review Model for the Description and Evaluation of Psychological Tests: Test Review Form and Notes for Reviewers, Version 3.41*. Brussels, Belgium: EFPA. Dostupný též z WWW: <www.efpa.eu>.

European Federation of Psychologists' Associations & European Association of Work and Organizational Psychologists (2005). *European Test User Standards for test use in Work and Organizational settings, Version 1.92*. Brussels, Belgium: Dostupný též z WWW: <www.efpa.eu>.

Fremer, J. (1996). Promoting high standards for test use: Developments in the United States. *European Journal of Psychological Assessment*, 12, 160-168.

Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J. R., Zaal, J. N. (2001). Testing Practices in European Countries. *European Journal of Psychological Assessment*, 17, 201-211.

Van de Vijver, F., Hambleton, R. K. (1996). Translating Tests: Some Practical Guidelines. *European Psychologist*, 1, 89-99.