

# A je ta zkouška spravedlivá? aneb Spravedlivost v jazykovém testování

Kateřina Vodičková

Vzhledem k profesionalizaci jazykového testování v posledních desetiletích a vzhledem ke skutečnosti, že různé zkoušky představují mnohdy podklad pro rozhodnutí vysoké důležitosti, klade se na tzv. spravedlivost zkoušek (*test fairness*) značný důraz. Ačkoliv mezi odbornou veřejností nepanuje v pohledu na spravedlivost naprostá shoda, je zřejmé, že se jedná o pojem v jazykovém testování zásadní.

O nejednotnosti chápání spravedlivého přístupu ke kandidátům ve zkouškách z češtiny pro cizince svědčí už samotná rozkolísanost nejen v překladu pojmu *fairness*, ale též v jeho chápání. Termín *fairness* bývá v českém kontextu užíván jako *férovost*, *spravedlnost*, *spravedlivý/rovný přístup ke kandidátům* apod.<sup>1</sup> Zatímco v případě zkoušek z jiných cizích jazyků, např. angličtiny, se již stalo dobrou praxí nejen spravedlivý přístup zajišťovat, ale též odbornou i širokou veřejnost seznamovat s tím, jak bylo spravedlivosti dosaženo, v případě češtiny pro cizince se mnohdy s tímto pojmem a příslušnou dokumentací teprve pracovat začíná.

Pojem spravedlivosti v testování se vyvíjel a rozbíhal různými směry, dokonce se může zdát, že si některá pojetí odporují. Většinou se pod spravedlivostí rozumí dostupnost zkoušky, rovnost, neexistence znevýhodnění, důvěryhodnost, spolehlivost, „spravedlnost pro všechny“ (Saville, 2002, s. 12). Popisný je v tomto směru název druhé části standardů v testování *The Standards for Educational and Psychological Testing* (1999), jenž zní *Spravedlivost v testování: spravedlivost v testování a užití testů, práva a povinnosti kandidátů, testování jednotlivců s odlišným zázemím nebo se zvláštními potřebami*.

Tradičně byla (a dosud dle našeho názoru je) spravedlivost spojena se zkoumáním validity a reliability. Například podle slovníku *Dictionary of Language Testing* (2001, s. 199–200) se spravedlivost týká důsledků pro testovaného jedince, skupiny a společnost jako celek a vztahuje se k validitě daného testu i celého procesu testování. Spojitost s validitou i důsledky nejen pro kandidáty, ale také pro společnost se odráží např. v *Rámci pro spravedlivé testování* (Kunnan, 2014, s. 37–39), který mapuje spravedlivost v následujících pěti oblastech: validita, nepřítomnost znevýhodnění, dostupnost, administrace a důsledky pro společnost.

Odborné diskuze o spravedlivosti se v souvislosti s hodnocením objevují od 80. let minulého století. V hodnocení ve vzdělávání byl tento termín poprvé použit v souvislosti s testovou baterií *General Aptitude Test* (Hartigan, Wigdor, 1989), v souvislosti s jazykovým testováním o termínu v roce 1997 diskutoval Kunnan na

---

<sup>1</sup> V tomto příspěvku užíváme v češtině termínu *spravedlivost*.

mezinárodní konferenci *Language Testing Research Colloquium*. V roce 1999 byla spravedlivost zařazena do standardů pro testování *The Standards for Educational and Psychological Testing* do sekce s názvem *Spravedlivost v testování (Fairness in Testing)*, dílčí sekce se věnovaly spravedlivosti v testování a užití testu, právům a povinnostem kandidátů, testování jedinců s různým lingvistickým zázemím a testování jedinců se speciálními potřebami.<sup>2</sup> Od roku 2000 se spravedlivostí již zabývá několik publikací (např. Kunnan, 2004; Kunnan, 2014; McNamara, Roever, 2006; celé jedno číslo *Research Notes* z roku 2005 bylo pod názvem *Ethics and Fairness in Language Testing* věnováno spravedlivosti) či studií a článků.

Spravedlivost představuje komplexní, pro mnohé odborníky obtížně uchopitelný pojem. Tato skutečnost vede k nejednotnosti chápání spravedlivosti. Její pojetí i přístup k ní se vyvíjely, např. Kunnan (2010) hovoří o období, kdy byla spravedlivost chápána především jako absence znevýhodnění (60.-80. léta 20. století), následovalo období kodexů (80. léta) a standardů (90. léta). Pro 21. století je podle Kunnana (2010) typické řešení etiky a zásad nebo filozofický přístup.<sup>3</sup>

Příkladem kodexů a standardů, jež se užívají dodnes a se spravedlivostí pracují, jsou např. standardy Americké psychologické asociace (American Psychological Association: *The APA Standards for Educational and Psychological Testing*<sup>4</sup>), Kodex dobré praxe Evropské asociace pro jazykové testování (European Association of Language Testing and Assessment: *The EALTA Code of Good Practice*<sup>5</sup>), pokyny a etický kodex mezinárodní testerské asociace ILTA (International Language Testing Association: *The ILTA Guidelines for Practice & Code of Ethics*<sup>6</sup>) či systém pro monitorování kvality, jež vyvinula Asociace jazykových testerů v Evropě ALTE (Association of Language Testers in Europe: *The ALTE Quality Management System – Code of Practice and Minimum Standards*<sup>7</sup>).

*The Code of Practice* asociace ALTE a *The Code of Ethics* asociace ILTA reprezentují široce pojaté zásady profesionálního chování, jak dosáhnout dobré praxe, spravedlnosti a spravedlivosti pro všechny. Například ALTE vystavělo *The Code of Practice* na 18 širokých prohlášeních, která pokrývají vývoj zkoušky, problematiku výsledků, spravedlivost a vztah ke kandidátům, pozornost se soustřeďuje zejména na roli různých skupin uživatelů zkoušky vzhledem k usilování o spravedlivost. *The Code of Ethics* asociace ILTA oproti tomu představuje spíše morální zásady a ideály jazykového testování jakožto profese.

---

<sup>2</sup> Podrobnější informace viz např. Kunnan (2014).

<sup>3</sup> Podrobně viz Kunnan (2010).

<sup>4</sup> <http://www.apa.org>

<sup>5</sup> <http://www.ealta.eu.org>

<sup>6</sup> <http://www.iltaonline.com>

<sup>7</sup> <http://www.alte.org>

U výše uvedených materiálů se jedná převážně o doporučení, jež mají zajistit (alespoň minimální) kvalitu testu. Otázkou, zda by takovýto kodex dobré praxe neměl vzniknout i v českém kontextu, se zabývá např. Hulešová (2015).

Zajímal nás pohled na danou problematiku nejen (české) odborné veřejnosti, ale též studentů, resp. potenciálních kandidátů různých zkoušek a testů. Provedli jsme tedy mezi odbornou veřejností a uživateli zkoušek malý průzkum prostřednictvím dotazníkového šetření.<sup>8</sup>

V první otázce měli respondenti odpovědět, jak by do češtiny přeložili pojem *test fairness*, případně *fairness in testing*. Zde jednoznačně převažoval překlad spravedlnost, mnohdy upřesněný: spravedlnost při testu, při testování/zkoušení, spravedlivé testování. V jednom případě se objevil návrh užívat výrazu férovost, tři respondenti uvedli, že nevědí, jak by anglický pojem přeložili.<sup>9</sup>

Ve druhé otázce nás zajímalo, co si účastníci dotazníkové „sondy“ představují pod pojmem „spravedlivost ve zkoušce/testu“. Zde respondenti uváděli především rovné podmínky pro všechny účastníky testu, dále objektivitu, srovnatelnost variant testu („stejně varianty“), spravedlivé hodnocení, nezávislost na vnějších faktorech jako prostředí, osobnost examinátora atd. Podle jednoho respondenta je spravedlivá zkouška ta, která neznevýhodňuje, ani nezvýhodňuje kandidáta na základě jeho věku, kultury či mateřského jazyka, resp. jazyků, které ovládá. Zmiňuje se i dodržování stanovených kritérií a podmínek zkoušky.

Pozornost si dle našeho názoru zaslouží i zamyšlení dvou respondentů nad rozdílem mezi objektivností a spravedlivostí, přičemž spravedlivost chápe jeden z účastníků dotazníkového šetření jako pojem hlubší a uvádí následující přístup: „Pokud učitel učí žáka dlouhodobě a zná ho jako pilného a schopného studenta, může mu dát vyšší známku u zkoušky, přestože z důvodu nemoci nevykázal nejlepší výsledek<sup>10</sup>. Bude to spravedlivé, ale nebude v tomto případě objektivní.“ Obdobně druhý respondent zvažuje, že by spravedlivost měla zahrnovat následující: 1) Individuální přístup k žákovi – hodnotit Petra vůči Petrovi a neporovnávat ho s Honzou (osobní růst), 2) přípustnost hodnocení vyšší známku jako odměna za snahu a odvedenou práci. Dle našeho názoru se v těchto případech jedná spí-

---

<sup>8</sup> Průzkumu se zúčastnilo 10 respondentů, z toho jeden muž, devět žen. Šest respondentů uvedlo národnost českou. Co se týče profesí, jednalo se o učitele, a to především na vysokých školách, a/nebo o jazykové testery. Čtyři respondenti byli studenty, dva respondenti označili jako profesi „jiné“, ale blíže ji nespecifikovali.

<sup>9</sup> Pozornost si zaslouží i podrobnější vyjádření jednoho z respondentů, který se domnívá, že „pokud jde o zavedení určitého pojmu, vyjadřujícího objektivitu a nezaújatý postoj zkoušejícího, lze použít i anglický výraz *test fairness* bez překladu“. S tím je samozřejmě možné souhlasit, na druhou stranu jsme si vědomi možných komplikací, mj. (ne)začlenění anglického výrazu do českého deklinačního systému a skutečnosti, že ne každému musí být anglický pojem srozumitelný.

<sup>10</sup> Dle našeho názoru míněno „nepředvedl nejlepší výkon“.

še o motivaci než o spravedlivý přístup v hodnocení konkrétní zkoušky či testu. Možnost zohlednit uvedené aspekty také souvisí s účelem konkrétního hodnocení, např. takovýto přístup jistě nebude možný v případě mezinárodně uznávaných zkoušek úrovně (proficiency test).

Není bez zajímavosti, které zkoušky účastníci dotazníkového šetření považovali za spravedlivé. Uváděli např. mezinárodní zkoušky DELE, DaF, TOEFL, zkoušky Cambridge English Language Assessment (např. FCE), ale též státní jazykovou zkoušku a zkoušky z konkrétních předmětů na vysokých školách<sup>11</sup>. Mezi zkoušky, které respondenti nepovažovali za spolehlivé, zařadili zkoušku z ruského jazyka TRKI, státní maturitu, starou podobu státní maturity a řadu zkoušek z některých předmětů na konkrétních vysokých školách<sup>12</sup>, jeden respondent uvedl „jakýkoli jazykový test sestavený testovacími společnostmi (SCIO apod.), který není sestaven dle individuálních požadavků jednotlivých pracovišť“.

Dotazníková „sonda“ potvrdila, že v očích odborné i široké veřejnosti hraje spravedlivost v jazykových zkouškách významnou roli. Respondenti si vytvářejí vlastní názor na to, které zkoušky lze považovat za spravedlivé, které nikoliv a proč. Spravedlivost je pro účastníky průzkumu spojena s dostupností zkoušky pro všechny, s rovným a spravedlivým přístupem ke kandidátům, spolehlivostí zkoušky a celkově s její důvěryhodností. Předpokládáme, že bude panovat značná shoda v tom, že by mělo být v testování cílem usilovat o celkovou spravedlivost, jednat vždy čestně a vyhýbat se negativnímu vlivu, aby byly kandidátům a dalším uživatelům zkoušky zajištěny vysoce kvalitní služby (Saville, 2005, s. 2). Pokud mají být zajištěny kvalitní služby a spravedlivost testu i celého procesu testování, je zřejmé, že se spravedlivost bude odrážet ve všech fázích procesu konstrukce testu, tj. již od identifikace potřeby testu až po interpretaci výsledků a komunikaci s uživateli zkoušky. Za spravedlivost v testování tak zodpovídají nejen tvůrci testu, ale všichni zúčastnění, tj. uživatelé zkoušky, poskytovatelé zkoušky a další.

Spravedlivost prostupuje všemi fázemi procesu a je součástí validace každé zkoušky. Je nutné nejen při vývoji zkoušky respektovat jistá pravidla a doporučení, ale také v rámci validačního argumentu dokumentovat, jak bylo spravedlivosti dosaženo. V neposlední řadě se domníváme, že by měli být především uživatelé zkoušky a odborná veřejnost informováni o tom, jak je spravedlivosti u konkrétních zkoušek dosahováno.

## Literatura

DAVIES, A. ET AL. (2002). *Dictionary of Language Testing*. Cambridge, UK: CUP. ISBN 0-521-65876-4.

<sup>11</sup> Jednalo se o závěrečný písemný test z ruštiny Jazyková cvičení, zkoušku z angličtiny na FS ČVUT, jazykovou zkoušku z angličtiny (univerzitní kurz) i konkrétní přijímací oborové testy.

<sup>12</sup> Např. zkouška z němčiny na FS ČVUT, zkouška z latiny na vysoké škole.

- HARTIGAN, J. A., WIGDOR, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues and the General Aptitude Test Battery*. Washington, DC: National Academy Press. ISBN 0-309-04030-2.
- HULEŠOVÁ, M. (2016). *Kodex dobré praxe v jazykovém testování – a proč?* Dostupné na <http://ujop.cuni.cz/strediska/vyzkumne-a-testovaci-centrum/vyzkumne-a-testovaci-centrum-vyzkum> [online 29. 9. 2016]
- KUNNAN, A. J. (2004). Test fairness. In autor Milanovic, M., Weir, C. (Eds). *European Language Testing in Global Context*. Cambridge, UK: CUP, s. 27–48. ISBN 0-521-53587-5.
- KUNNAN, A. J. (2010). *Test fairness and ALTEs model of quality management*. ALTE 39th meeting, Prague. Dostupné na: <http://www.alte.org> (pro členy ALTE).
- KUNNAN, A. J. (2014) Fairness and Justice in Language Assessment. In KUNNAN, A. J. (Ed.) *The Companion to Language Assessment*. Malden, MA: Wiley, s. 27–45. ISBN 978-0-470-65533-7.
- McNAMARA, T., ROEVER, C. (2006). *Language Testing: The Social Dimension*. Malden, MA – Oxford: Blackwell Publishing. ISBN 1-4051-5543-4.
- SAVILLE, N. (2002). Striving for Fairness – The ALTE Code of Practice and Quality Management Systems. *Cambridge ESOL Research Notes*, č. 7, s. 12–14.
- SAVILLE, N. (2005). Setting and monitoring professional standards: A QMS approach. *Cambridge ESOL Research Notes*, č. 22, s. 2–5.
- WEYANT, K., CHISHOLM, A. (2014). Safeguarding fairness principles through the test development process. *Cambridge English: Research Notes*, č. 55, s. 3–6.

## Autorka

**Mgr. Kateřina Vodičková, Ph.D., MA**, e-mail: [katerina.vodickova@ujop.cuni.cz](mailto:katerina.vodickova@ujop.cuni.cz), Ústav jazykové a odborné přípravy, Výzkumné a testovací centrum, Univerzita Karlova v Praze.

Autorka vyučuje češtinu pro cizince, působí v oboru čeština jako cizí jazyk a v současné době se věnuje především jazykovému testování.