

## Zpráva z desáté Mezinárodní vědecké konference Mezinárodní asociace pro hodnocení výsledků vzdělávání Dublin, 27.–30. června 2023

Výroční desátou Mezinárodní vědeckou konferenci (IRC) Mezinárodní asociace pro hodnocení výsledků vzdělávání (IEA), která se letos konala Dublinu (Irsko), hostilo Centrum pedagogického výzkumu (The Educational Research Centre, ERC), což je výzkumná organizace zřizovaná irským ministerstvem školství. Kromě realizace výzkumu různých aspektů vzdělávání napříč všemi úrovněmi vzdělávací soustavy a podpory školám a učitelům v oblasti hodnocení a evaluace ERC také úzce spolupracuje s IEA na realizaci mezinárodních šetření. Bylo tak nasnadě, že se ve spolupráci s IEA jeho zástupci ujali organizace této mezinárodní konference, jejímž cílem je zejména prezentovat a kriticky diskutovat zjištění ze studií IEA (TIMSS, PIRLS, ICCS, ICILS, či REDS) a také metodologické aspekty těchto šetření. Konference se účastnilo více než 200 odborníků z celého světa (dosud nejvíce ze všech konferencí IEA) různých odborných zaměření a profesních profilů, zastoupeny byly země Severní i Latinské Ameriky, Afriky, Asie i Evropy.

Během hlavního programu, kterému předcházely dva dny metodologických workshopů, zazněly tři plenární příspěvky. Již tradičně se jednoho vždy ujímá místní expert, který účastníkům přibližuje význam a úlohu IEA studií v hostitelské zemi. V případě této konference šlo o Gerryho Shiela (seniorní výzkumník z ERC), který posluchačům detailně přiblížil, jakou roli historicky i v současné době hrály a hrají tyto studie při formování vzdělávací politiky. Irsko se do IEA studií zapojilo začátkem devadesátých let (TIMSS 1995), nicméně poté se dalších kol šetření neúčastnilo. Svůj postoj ale přehodnotilo v důsledku „PISA šoku“, neboť v roce 2009 došlo k prudkému zhoršení výsledků irských žáků v matematické i čtenářské gramotnosti. Shiel to přičítá nepříznivému dobovému socioekonomickému kontextu, kdy Irsko obzvláště významně pocítovalo negativní dopady globální finanční krize. „PISA šok“ vedl irskou decizní sféru k rozhodnutí zapojit se do TIMSS a PIRLS 2011, aby byly snáze identifikovány problémy vzdělávacího systému již v časnějších fázích vzdělávací dráhy žáků. Byly také vypracovány nové koncepční dokumenty vzdělávací politiky na roky 2011–2020. Strategie zahrnovala množství různých opatření (mimo jiné například zavedení standardizovaného testování ve 2., 4. a 6. ročníku) a stanovila kvantifikovatelné cíle, kterých mělo být pomocí těchto opatření dosaženo (např. podíly žáků absolvujících vyšší úroveň maturity). Průběžná evaluace této strategie v roce 2017 se opírala o velké množství informací získaných právě ve studiích PIRLS a TIMSS, jejichž

90 výsledky dokládají jak realizaci procesních reforem (např. zvýšení množství vyučovacíh hodin angličtiny mezi lety 2011 a 2016), tak i dopady na výsledky žáků odrážející náročnější požadavky.<sup>1</sup> Revidované cíle stanovené pro rok 2021 se nicméně naplnit nepodařilo, což Shiel přičítá negativnímu dopadu opatření spojených s pandemií covid-19. Irsko se chystá účastnit i budoucích studií TIMSS a PIRLS.

Druhou významnou vystupující odbornicí se stala účastníkem IEA konferencí dobře známá Leslie Rutkowski (profesorka metodologie kvantitativního výzkumu, Indiana University, USA). Pro svoji přednášku ne zvolila partikulární metodologické nuance různých statistických metod, ale širší a docela provokativní téma, neboť problematizovala používání výsledků IEA studií jako indikátoru inteligence, jak je prováděno některými hojně citovanými akademiky a mnohdy publikováno i ve vysoce impaktovaných časopisech. Svůj příspěvek začala právě metodologickou kritikou vybraných autorů a prací, aby následně posluchače provedla historií měření inteligence. Přiblížila historii eugenického hnutí v USA a první průkopníky v oblasti měření inteligence (často se jednalo o velmi dobré matematiky a statistiky, jejichž statistické testy či postupy jsou dodnes využívány – Pearson, Spearman, Fischer) a upozornila na zneužití některých nástrojů vytvořených pro měření inteligence (např. na základě Stanford-Binetovy škály pro měření inteligence bylo začátkem 20. století z USA deportováno velké množství imigrantů „nesplňujících normy“, blíže viz Allen, 2006). Ústup od testování inteligence a zobecňování výsledků na skupiny (například příslušníky různých národů) je zaznamenán zejména po 2. světové válce (v důsledku hrůzné aplikace některých antroporasyových teorií), avšak jak si Rutkowski všimá, v poslední dekádě opět nabírá na síle. Důrazně upozorňuje, že není možné chápat výsledky mezinárodních šetření jako indikátor inteligence, neboť do výsledků žáků se promítá velké množství různých jiných faktorů, mimo jiné i skutečnost, že testy jsou tzv. „low stake“, a sami žáci na konci testování mnohdy indikují, že se při vypracovávání testů příliš nesnažili.

Třetí plenární přednáška se týkala dokladů, jak studie IEA (ale i OECD) mohou přímo či nepřímo pomáhat učitelům zlepšovat svoji výuku. Předcházel jí krátký panel, v němž odborní pracovníci IEA představili novou sérii publikací určenou právě učitelům (Teacher snippets a Research for educators), v nichž jsou vybrána relevantní data ze studií IEA a přístupným způsobem interpretována tak, aby je učitelé mohli využít ve své praxi. V samotné plenární přednášce pak prof. Caroll Campbellová (University of Toronto) představila konkrétní zjištění z šetření TIMSS a PIRLS týkající se například efektivních vyučovacích postupů, způsobů hodnocení žáků, učitelského leadershipu a jiných aktivit, které učitelé mohou používat ke zlepšení výsledků žáků.

S povděkem jsem kvitoval, že oproti IEA IRC konferenci v roce 2017, z níž jsem odcházel s pocitem, že většina výzkumníků kladla důraz především na statistické postupy správného zpracování dat, se většina jednotlivých příspěvků i diskuzní pa-

<sup>1</sup> Samozřejmě existuje zde i riziko zpětného působení a pokud se tvůrci vzdělávací politiky zaměří jen na výsledky vzdělávání a jejich kontext, který je v těchto studiích zjišťován, riskují, že jiné důležité výstupy či relevantní kontextuální charakteristiky nezjišťované v těchto studiích zůstanou zcela mimo jejich pozornost.

nely v mnohem větší míře zaměřovaly na věcné interpretace s ohledem na domácí kontexty jednotlivých zemí a také na praktické implikace zjištění z mezinárodních šetření IEA pro učitele, školy i tvůrce vzdělávacích politik. Samozřejmě to neznamená, že by metodologické aspekty stály úplně stranou pozornosti, nicméně i metodologické příspěvky měly širší přesah (oproti předchozím IRC konferencím byly příspěvky nově rozděleny do tří kategorií podle toho, zda se týkaly praxe, politiky či metodologie). Důležitý mi v této souvislosti přišel zejména diskuzní panel o možnostech (korektního) vyvozování kauzálních závěrů z dat mezinárodních šetření (za předpokladu použití některých statistických metod simulujících kvaziexperimentální design, např. propensity score matching). Jeho účastníci byli více či méně skeptičtí k možnosti takového usuzování, a to zejména proto, že metodologický design a vzorkování studií IEA ze své podstaty neumožňuje naplnit Rubinovy (1974) předpoklady. Jako nadějnější spatřovali možnost navázat v národních kontextech na studie PIRLS či TIMSS buď kvalitativními studiemi, nebo vytvořit panel umožňující sledovat vybraný vzorek žáků a jejich průchod vzdělávacím systémem longitudinálně (dobrým příkladem z České republiky může být longitudinální výzkum CLoSE realizovaný v letech 2012–2018). Kromě spolehlivějších možností pro vyvozování kauzálních závěrů by takové studie podle účastníků panelu mohly posílit zájem veřejnosti o výsledky mezinárodních studií. I na této konferenci zaznělo několik příspěvků, které kromě dat z mezinárodních šetření vytěžovaly také data sebraná v návazných empirických šetřeních. Například ve studii ICCS v Mexiku bylo zjištěno, že by více než dvě třetiny respondentů podpořily diktaturu, pokud by společnosti přinesla řád a stabilitu, a mexičtí výzkumníci se proto rozhodli kvalitativně prozkoumat, jaké důvody žáky k tomuto tvrzení vedou.

Kromě metodologických a empirických příspěvků byly na konferenci prezentovány také přehledové studie. Jedna z nich navazovala na předchozí práce, ale zaměřila se na IEA studie o informačních technologiích (kromě novější ICILS také jejího předchůdce COMPED z devadesátých let 20. století). Její autorka zjistila, že v porovnání s jinými IEA šetřeními jsou využívány v mnohem menší míře, což se odráží mimo jiné v kvantitě empirických výstupů v mezinárodních recenzovaných časopisech. Výzkumníkům na základě svých výsledků doporučila také více vytěžovat data z dotazníků pro ICT koordinátory, které byly v dosud publikovaných statích využívány sporadicky či vůbec.

Konference IEA probíhají od roku 2004 každé dva roky a tato byla výroční – desátá. Vybízela tak k hlubšímu zamyšlení nad významem mezinárodních šetření a jejich budoucností. Zvyšující se zájem odborné veřejnosti (indikovaný dosud nejvyšším počtem účastníků) trochu kontrastuje s do jisté míry vadnoucím zájmem veřejnosti, a to jak té laické, tak i profesní. Nejen že zveřejnění výsledků mezinárodních šetření nebývá očekáváno s takovým napětím či nedočkavostí jako dříve, ale i sami ředitelé či učitelé jsou často velmi unavení z množství dotazníkových šetření, kterých se musí účastnit, mnohdy bez jasné představy, co jejich účast přinese pozitivního když ne jim, tak alespoň škole či vzdělávacímu systému jako celku. I z toho důvodu vnímám velmi pozitivně, že se celou konferencí jako červená nit táhla myšlenka či

- 92 nevyřčená snaha, aby výzkumná zjištění založená na datech IEA pomáhala zejména těm v „první linii“ (učitelům a ředitelům) a také aby vedla k pozitivním změnám ve vzdělávacích politikách. Snad i díky této konferenci se myšlenka využitelnosti a užitečnosti pro praxi promítne i do každodenní práce nejen účastníků konference, ale všech analytiků a odborníků, kteří s těmito daty z mezinárodních šetření pracují.

### Poděkování

Má cesta na konferenci byla umožněna díky projektu České školní inspekce číslo CZ.02.02.XX/00/22\_005/0002901 *Datově-analytická podpora pro hodnocení a řízení vzdělávací soustavy ČR*. Autor zprávy děkuje za poskytnou podporu.

Vít Štastný  
vit.stastny@pedf.cuni.cz

### Reference

- Allen, G. E. (2006). Intelligence tests and immigration to the United States, 1900–1940. In *Encyclopedia of the life sciences*, 1–5. <https://doi.org/10.1002/9780470015902.a0005612>
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.