

Rozdíly mezi pořadím na základě škál schopnosti a průměrnými pořadími zemí podle procenta správných odpovědí

Mezinárodní výsledky PISA jsou založené na počtu bodů žáka v dané otázce skrze model založený na teorii odpovědi na položku (IRT) (viz *Technická zpráva PISA 2012*, OECD, 2014). Toto škálování se provádí z několika důvodů. Zaprvé podporuje sestavování popsaných škál schopnosti. Zadruhé tento přístup shrnuje odpovědi žáků na mnohé otázky několika málo indexy. Tím zajišťuje, že jsou indexy srovnatelné u žáků, kteří odpovídali do různých testových sešitů, jež obsahují různé dílčí sady otázek (Adams a kol., 2010). Škálování žakovských skóre odráží přístup PISA, který spočívá v budování mezinárodně podporovaných rámců hodnocení a následném vývoji základů otázek, které stejným způsobem tvoří vzorky z těchto rámců.

Přístup založený na průměru procent správných odpovědí použitý v Kapitole 3 tohoto dílu přináší další způsob, jak porovnat výsledky země v šetření. Výhoda tohoto přístupu je, že se dá snadno replikovat na libovolné dílčí sady otázek.

Když se pořadí zemí založená na tomto přístupu, používajícím všechny otázky, porovnájí s pořadím zemí založenými na obvyklém škálovacím přístupu, objeví se malé rozdíly. Existuje pro to šest důvodů: zaprvé, metodika založená na průměru procent správných odpovědí přiřazuje libovolnou hodnotu (typicky 0 nebo 0,5) všem odpovědím s částečným počtem bodů; procenta správných odpovědí jsou tudíž založena na menší sadě informací o výkonu žáků v testu než u výsledků na škále, kde je každá odpověď s částečným počtem bodů škálována podle konkrétní obtížnosti. Zadruhé, tato metodika nebere v potaz žáky, kteří nezodpověděli žádnou otázku v testu řešení problémů, ačkoliv jim byl přidělen sešit s testovými úlohami a alespoň částečně vyplnili žakovský dotazník. Protože není možné zjistit, proč nezodpověděli otázky v testu řešení problémů (např. zda došlo k technickému selhání počítačového systému nebo úmyslně odmítli plnit úkoly), jsou jejich odpovědi kódovány jako „nevyplněno“, ne jako nesprávné odpovědi, a je jim přiřazena hodnota pro chybějící odpověď. Obvyklý škálovací přístup naproti tomu napravuje situace, kdy se žák rozhodne test nevyplnit, imputací výkonu z dostupných informací o něm, včetně jeho výkonu v jiných testech. Zatřetí, metodika založená na procentu správných odpovědí váží všechny otázky stejně, zatímco podle škálovacího přístupu jsou otázky váženy počtem testových sešitů, do kterých byly zařazeny. Začtvrté, přístup založený na procentu správných odpovědí neřeší efekt testových sešitů, který byl šetření v PISA pozorován. Zapáté, škálovací metodika převádí procentní hodnoty, které jsou omezené nulou a 100, do logitové škály. Tato transformace má účinek „natahování“ velmi nízkých a velmi vysokých procentních hodnot v porovnání s hodnotami, které se blíží 50 %. Zašesté, pokud se po uskutečnění testování v nějaké zemi odhalí problém, například chyba v překladu jedné otázky, je jí přidělena hodnota pro chybějící odpověď v testech u všech žáků v této zemi. Pořadí zemí podle procenta správných odpovědí mohou tedy vycházet z menšího počtu otázek než škálované výsledky. V šetření PISA 2012 řešení problémů byla jedna otázka (CP018Q05) stažena po testování ve Francii, protože do národní verze omylem nebyl zahrnut jeden zásadní pokyn pro žáky.