

#### 4. Vyhodnocení testu

Postup spojený s vyhodnocením testů PIRLS/TIMSS je opětovně obdobný metodice šetření PISA. Z tohoto důvodu i zde uvádíme pouze shrnutí hlavních metodických procedur:

- Rozdílná pravděpodobnost výběru žáka pro šetření je řešena přiřazením různé váhy odpovědím žáků.
- Kvalita testových položek je zjišťována výpočtem základních statistických a psychometrických parametrů. Potenciálně problémové testové položky jsou označeny výstražnou značkou a následně kontrolovány vzhledem k dosahované kvalitě. Poznamenejme, že při ověřování kvality testových položek jsou mimo jiné sledovány úvahy, že země s lepšími celkovými výsledky dosahují lepších výsledků i v hodnocení jednotlivých položek a že odpovědi na testové položky zůstávají v čase konzistentní.
- Skóre žáka je počítáno prostřednictvím IRT<sup>92</sup> teorie – Raschova modelu, přičemž vyjadřuje pravděpodobnost, že žák bude odpovídat na testovou položku shodně s kalibrovaným modelem ve vazbě jednak na své schopnosti a jednak na charakteristiky testové položky. Skóre žáka je rovněž transformováno na škálu s průměrnou hodnotou 500 a směrodatnou odchylkou 100. Využíván je vyšší počet škál v návaznosti na základní metodický rámec šetření PIRLS/TIMSS.
- Propojující testové položky umožňují kalibrovat původní model výpočtu škály šetření PIRLS/TIMSS a tímto způsobem zachytit vývoj výsledků v čase.
- V rámci definovaných škál jsou stanoveny 4 úrovně výkonnosti – pokročilá, vysoká, střední a nízká, přičemž v roce 2001 byly hraniční body stanoveny pomocí percentilů (90, 75, 50 a 25) a v roce 2006 pomocí pevně stanovených hranic. Popis kompetenčních znalostí a dovedností žáků na dané výkonnostní úrovni byl proveden prostřednictvím hodnocení úspěšnosti jednotlivých testových položek.

Zdůrazněme, že, stejně jako šetření PISA, respektuje i šetření PIRLS/TIMSS komplexní systém managementu kvality vyhodnocení výsledků.

#### 5. Reporting výsledků testu

Reporting výsledků šetření PIRLS/TIMSS probíhá podobně jako v případě šetření PISA ve vztahu k definované škále šetření, tentokrát se čtyřmi definovanými úrovněmi výkonnosti. Výsledky zemí jsou prezentovány zejména prostřednictvím průměrné hodnoty dosaženého skóre testovaných žáků jednotlivých zemí respektive prostřednictvím rozdělení četností výsledků žáků vzhledem k definovaným úrovním škály. Zachyceny jsou rovněž změny dosažených výsledků v čase, rozdíly ve výsledcích vzhledem k charakteristikám žáků, rozdíly ve výsledcích žáků mezi školami a žáků uvnitř školy. Při hodnocení je respektována statistická významnost rozdílů.

- ICCS

Čtvrtým mezinárodním šetřením našeho zájmu je šetření ICCS, které bylo v roce 2009 realizováno IEA ve spolupráci s australským *Australian Council for Educational Research* (dále jen ACER). Hlavním cílem šetření ICCS je poznat, jak jsou mladí lidé jako občané schopni plnit svou roli ve společnosti (viz Schulz,

---

<sup>92</sup> Item Response Theory