

- vyhodnocení s využitím bodové škály s průměrnou hodnotou 250 a směrodatnou odchylkou 50, na kterou jsou výsledky žáků převedeny prostřednictvím IRT teorie¹⁰⁹,
- vyhodnocení s využitím 5 úrovní úspěšnosti odpovědí žáků s tím, že hranice jednotlivých úrovní jsou stanoveny na základě percentilů žáků (10 % - 25 % - 30 % - 25 % - 10 %), kteří by měli být schopni správně zodpovědět otázky dané obtížnosti.

Poznamenejme, že hraniční body jednotlivých úrovní obtížnosti jsou staveny s využitím bodové škály tak, aby například 10 % testovaných žáků dosáhlo počtu bodů nižších než hraniční hodnota bodové škály pro úroveň nižší než 1, a dále že pro každou úroveň obtížnosti jsou definovány jednak typové testové položky, které by měl žák na dané úrovni výkonnosti správně odpovědět a jednak kompetence, které by měl žák na dané úrovni výkonnosti ovládat (viz Eivers et al. 2010a). Metodika irského národního hodnocení rovněž zohledňuje aspekt různých vah přiřazených jednotlivým odpovědím žáků pro zajištění reprezentativnosti šetření. Výsledky irského národního hodnocení jsou dále diferencovány vzhledem k dílčím proměnným, jako jsou charakteristiky rodičů a rodinného prostředí (např. jazyk v domácnosti, vybavenost domácnosti knihami, postoje rodičů ke vzdělávání a další) nebo charakteristiky školy (např. poloha školy, složení žáků vzhledem k pohlaví, jazyk vyučování, učební metody a další). Významnost identifikovaných rozdílů je hodnocena tradičními metodami inferenční statistiky (viz Eivers et al. 2010b).

5. Reporting výsledků testu

V oblasti reportingu výsledků testu je primárně potřeba vzít do úvahy cíle irského národního hodnocení v podobě zajištění dat pro formulaci vzdělávací politiky respektive strategie rozdělování zdrojů (srovnej s Eivers 2010a, Hislop 2013). Současně jsou sledovány faktory pro vysvětlení rozdílů ve vzdělávacích výsledcích. Takto reporting výsledků irského národního hodnocení není primárně orientován na skupiny žáků, rodičů či škol, nýbrž na úroveň národní. Za tímto účelem je rovněž z jednotlivých testování zpracována závěrečná zpráva s podrobnostmi celkových výsledků i výsledků podle sledovaných proměnných (Eivers et al. 2010b).

2.1.6 Kanada

V případě Kanady je předmětem našeho zájmu při sledování dosažené úrovně základních gramotností federální hodnotící program s názvem *Pan-Canadian Assessment Program* (dále jen PCAP). Obdobně nastavený program byl v Kanadě iniciován již v první polovině 90. let 20. století (např. Volante a Jaafar 2008), do podoby PCAP byl transformován v roce 2007. Cílem PCAP je zajistit potřebné informace o vzdělávacích výkonech kanadských žáků napříč provinciemi. PCAP přitom není šetření celoplošné, nýbrž výběrové, a takto poskytuje zejména systémové informace na úrovni federace respektive provincií.

PCAP probíhá pravidelně ve tříletém cyklu, cílovou skupinou jsou žáci v 8. roce studia, tj. na konci základní školy (viz např. CMEC 2011b). PCAP se zaměřuje na hodnocení výkonnosti žáků ve třech oblastech v podobě (viz např. CMEC 2009, CMEC 2011a):

- čtenářské gramotnosti,
- matematické gramotnosti,
- přírodovědné gramotnosti.

¹⁰⁹ *Item Response Theory*