

Testové položky PCAP nabývají různých podob a zahrnují vedle multi-choice otázek (cca 70 % otázek) rovněž otevřené otázky s potřebou krátké či delší odpovědi. Malý prostor je věnován pocitovým otázkám s cílem konfrontovat poznatky z tohoto typu otázek v dotaznících na jedné straně a testů na straně druhé. Součástí testů jsou dále propojující testové položky, tj. položky použité i v předchozích šetřeních, které umožňují srovnání vývoje výsledků v čase¹¹² (viz CMEC 2011b). Konečně součástí PCAP je také dotazník zaměřený na cílové skupiny žáků, učitelů a ředitelů s cílem identifikovat kontextové faktory (např. demografické a socioekonomické charakteristiky, postoje ke vzdělávání) pro vysvětlení rozdílů v dosažených výsledcích ve vzdělávání (CMEC 2009, CMEC 2012, CMEC 2013).

3. Administrace testu

PCAP testování probíhá v tištěné formě ve spolupráci provincie, místního úřadu a školního koordinátora. Testovací materiál je místním úřadem doručován osobě školního koordinátora, a to včetně formuláře se jmény účastníků se žáků. Tento formulář podléhá kontrole a vyplnění skutečné přítomnosti žáků ze strany školního koordinátora (viz CMEC 2010b). PCAP testování probíhá v předem daném časovém rozpětí cca 10 dní s tím, že délka vyplnění testového sešitu všech tří gramotností je cca 90 minut s možností prodloužení o 30 minut v případě potřeby (viz CMEC 2013, CMEC 2010b). Po skončení testování školní koordinátor sesbírá testy a zasílá je zpět místnímu úřadu ke zpracování (viz CMEC 2010b).

Oprava testů probíhá externím týmem proškolených hodnotitelů v průběhu cca tří týdnů po realizaci PCPA. Proces opravy testů je realizován na jednom předem určeném místě, kam jsou soustředěny všechny testovací sešity. Hodnocení je prováděno hodnotiteli z různých místních úřadů s tím, že pro hodnocení kvality oprav je dvakrát denně prováděna evaluace spolehlivosti hodnotitele a určitý počet testů je opravován dvakrát (viz např. CMEC 2009, CMEC 2011b).

4. Vyhodnocení testu

Opravené testy jsou vyhodnoceny na bázi IRT teorie¹¹³, která současně utváří teoretické východisko pro převedení výsledků žáků na škály s průměrem 500 a směrodatnou odchylkou 100 (např. CMEC 2013, CMEC 2011a). Pro hlavní testovanou gramotnost jsou výsledky hodnoceny rovněž na třech úrovních úspěšnosti, jejichž hraniční skóre je definováno expertně ve vazbě na očekávané schopnosti a dovednosti žáků a na podstatu a obtížnost jednotlivých otázek (viz CMEC 2009, CMEC 2011b). Na těchto dvou metrikách jsou následně srovnávány mimo jiné dosažené výsledky (viz např. CMEC 2011a, CMEC 2013):

- žáků jednotlivých kanadských provincií a teritorií,
- žáků různých demografických a socioekonomických charakteristik,
- žáků v rámci definovaných kategorií obsahu a kognitivních procesů metodického rámce testovaných gramotností.

Konečně s využitím propojujících položek a transformace škál je hodnocen vývoj dosažených gramotností v čase (CMEC 2011b). Metody korelačního a regresního počtu pak hodnotí vazby mezi položkami kontextového dotazníku a dosažené výkonnosti žáků (CMEC 2011a).

¹¹² Poznamenejme, že v roce 2010 byly tyto propojující položky použity pouze pro testování čtenářské gramotnosti, a to s ohledem na nižší počet žáků testovaných v dalších dvou gramotnostech v PCAP 2007 (viz CMEC 2011b).

¹¹³ *Item Response Theory*