

Příklad 3.9 Adaptivní testování v projektu NIQUES

Projekt NIQUES v roce 2012 i 2013 k počítačovému adaptivnímu testování určitým způsobem směřoval tak, že podle výsledků první části testu žáci řešili odlišné varianty ve druhé části. „Z toho důvodu došlo u každého z testů po vyřešení úvodní, všem žákům společné části úloh s obtížností na úrovni minimálního standardu ke větvení – pokud žák nevyřešil správně alespoň dvě třetiny úloh, pokračoval v testu řešením dalších úloh opět na základní úrovni minimálního standardu. Pokud byla úspěšnost žáka v úvodní části úloh alespoň dvoutřetinová, zobrazily se žákovi ve zbytku testu úlohy vyšší úrovně. I po rozvětvení se žák mohl libovolně vracet k úlohám úvodní části, případně opravy řešení ale již nevedly ke změně zobrazených úloh. V testech z matematiky a českého jazyka byly v úvodní společné části testů i v pokračovacích částech testů za rozvětvením zastoupeny všechny tematické části úloh, v testech z cizích jazyků byly obsahem úvodní společné části testů pouze úlohy sledující porozumění poslechu a porozumění čtenému textu, naopak v pokračovacích částech testů již žáci řešili pouze úlohy sledující základní gramatické a lexikální dovednosti.“

Zdroj: Závěrečná zpráva o přípravě, průběhu a výsledcích druhé celoplošné generální zkoušky ověřování výsledků žáků v počátečním vzdělávání.

Obě uvedené moderní možnosti tvorby testů se efektivně využívají díky novější teorii testování, tzv. IRT (Item Response Theory). Uměla by si s nimi poradit i klasická teorie testů, která je dominantně popisována v této příručce, ale cesta je více krkolomná. Tato příručka se věnuje IRT v kapitole 4, ale jen okrajově, neboť její aktivní používání dosud při státních projektech testování chybí. Používána je v mezinárodních výzkumech.

3.2.5 Hodnocení škol z výsledků testování

Pro hodnocení škol Českou školní inspekcí je potřeba uvážit obdobné aspekty, jako byly uvedeny u hodnocení vzdělávací soustavy. Rozdíl je v tom, že u hodnocení vzdělávací soustavy je škola součástí vzorku a do šetření zahrnuta buď je, nebo není s ohledem na to, aby byly získány spolehlivé informace o celé vzdělávací soustavě. Škola je jen jednotkou v uvažované stratifikaci výběru. Z hlediska hodnocení škol Českou školní inspekcí je potřeba o školách zahrnutých do testování uvažovat jinak, a sice z perspektivy každé školy. Pokud se realizuje testování plošně a povinně, tak jsou tyto úvahy nadbytečné. Při testování jen vzorků škol je důležité, aby se v určité definované časové periodě dostalo na každou školu a bylo zajištěno spravedlivé rozložení testovaného obsahu a zapojení ročníků do testování. Jestliže komplexní inspekční činnost ve školách je realizována přibližně jednou za šest let, pak maximální interval testování školy by měl být jednou za šest let a zhruba v době před komplexní inspekcí, aby mohly být výsledky do inspekční zprávy zahrnuty a současně vhodně zasazeny do kontextu podmínek a dalších kvalit školy a nebyla tak kvalita školy redukována na výsledky v testech. Též by se inspektoři neměli nechat ovlivnit znalostí výsledků školy v testech při posuzování a hodnocení jiných stránek kvality školy.⁴⁶ Při hodnocení vzdělávací soustavy byly náročnější formy zkoušky např. u cizích jazyků uvažovány jen na zúženém vzorku škol. Pokud by měly být

⁴⁶ Z metodologie výzkumů v psychologii, sociologii i pedagogice je známo, že posuzovatel je ve svých úsudcích ovlivňován hodnocením jiných stránek posuzovaného objektu, a to i při jeho upřímné snaze o maximální objektivitu. Proto se organizují tzv. slepé či dvojité slepé experimenty. V případě činnosti České školní inspekce by inspektoři vykonávající inspekční činnost na dané škole neměli znát předem výsledky školy v testech, protože by je mohly i nevědomky ovlivnit při posuzování dalších kvalit školy při inspekční návštěvě. Důležité je případně uvážlivě a v kontextu zasazené uvedení výsledků testů do inspekční zprávy po zhodnocení školy v ostatních kritériích.