

Fórum:Testy/Vyrovnávání obtížnosti testů

Pokud je test administrován opakovaně, například pokud slouží k ověřování úrovně vědomostí potřebné pro výkon nějaké odbornosti nebo povolání, může být vznesen požadavek na zajištění vzájemné porovnatelnosti jednotlivých běhů testu. Součástí standardizace se pak stává **vyrovnávání obtížnosti testů** (též harmonizace testů). Vyrovnávání obtížnosti (angl. Equating) je statistický proces umožňující přepočítat hodnocení studentů z jednotlivých běhů (forem) testu tak, aby výsledky studentů dosažené v jednom běhu mohly být porovnávány s výsledky studentů v jiných bězích testu ^[1].

Lze k tomu použít řadu metod založených buď na klasické testové teorii (CTT), nebo na teorii odpovědi na položku (IRT) ^[2].

Základem části těchto metod je tzv. *kotvení testu*. Do testu se zařadí určitý počet úloh, které jsou ve všech verzích stejné. Tyto tzv. *kotvící položky* pak slouží ke vzájemnému porovnání verzí testu. Kotvící položky by měly být reprezentativní, měly by pokrývat rozsah obtížnosti testu a jejich počet by měl dosahovat minimálně 20% z délky testu ^[3].

Pro vyrovnávání obtížnosti testů a jejich škálování na základě IRT je k dispozici volně dostupný software IRTEQ (<http://www.hantest.net/irteq>) ^[4].

Reference

1. KOLEN, Michael J, Robert L BRENNAN a Michael J KOLEN. Test equating, scaling, and linking: methods and practices. 2nd ed. New York: Springer, c2004, xxvi, 548 p. ISBN 0-387-40086-9.
2. DAVIER, Alina A. *Statistical models for test equating, scaling, and linking*. New York: Springer, c2011, xix, 367 p. ISBN 978-0-387-98138-3-.
3. JELÍNEK, Martin a Petr KVĚTON. *Testování v psychologii : Teorie odpovědi na položku a počítačové adaptivní testování*. 1. vydání. Praha : Grada, 2011. 160 s. ISBN 978-802-4735-153.
4. Han, K. T. (2009). IRTEQ: Windows application that implements IRT scaling and equating [computer program]. *Applied Psychological Measurement*, 33(6), 491-493.