

Proto by bylo vhodné vedle reliability a specifík testovaného vzorku žáků uvádět i odhady chyby měření.

Kdyby mělo být na reliabilitu nahlíženo pouze jako na vlastnost testu (ne vlastnost vzorku testovaných žáků), pak by bylo potřeba uvést, pro jakou žákovskou populaci toto platí, a sběr dat by musel být realizován na dobře vybraném (náhodném) vzorku z dané populace.

4.6.3 Vztah mezi validitou a reliabilitou

Obecně platí, že reliabilita je nutnou podmínkou validity. Test nemůže být validní, pokud dostatečně přesně neměří to, co chceme měřit. Ale sebelepší reliabilita validitu testu nezajistí, test může měřit velmi přesně, ale něco úplně jiného, než autor testu zamýšlel. Někdy se mohou dostat nároky na validitu do rozporu s nároky na reliabilitu, resp. v teoretické rovině tento rozpor nastat nemůže, v empirické však ano. Např. test komplexnějšího charakteru, který pokrývá větší šíři obsahu i ověřovaných znalostí a dovedností žáků, může vykazovat vysokou míru obsahové validity. Odhad reliability na základě Cronbachova alfa však může být nižší, protože úlohy v testu jsou si méně podobné a schopnost žáka vyřešit jednu úlohu „nezaručuje“ vyřešení úlohy jiné. Matematicky řečeno, korelační koeficienty mezi úlohami by byly nižší, což se promítne do nižší vnitřní konzistence. Pro tento případ by bylo vhodnější použít metodu půlení s rozdělením na sudé a liché úlohy, pokud by byly úlohy v testu uspořádány do obsahově blízkých bloků.

Na základě empirického prokazování validity a reliability testu lze i matematicky vyjádřit vztahy mezi koeficienty, které validitu a reliabilitu dokladují.

Představme si výpočet kritériální validity testu na základě korelačního koeficientu mezi výsledkem testu t a vnějším kritériem k . T a K jsou pravá skóre měřené dovednosti v testu a daného kritéria, e a f jsou odpovídající chyby měření. Potom platí:

$$t = T + e \quad \text{a} \quad k = K + f.$$

Dá se dokázat⁹⁵, že platí:

$$r_{tk} = r_{TK} \sqrt{r_T r_K},$$

kde r_{tk} je empiricky zjištěná kritériální validita testu, r_{TK} je empiricky nezjistitelná skutečná korelace mezi pravými skóre testované dovednosti a kritériem, r_T je reliabilita testu, r_K je reliabilita měření kritéria. Protože reliabilita se pohybuje v intervalu mezi 0 a 1 a r_{TK} může být nanejvýše rovno 1 (to nastane tehdy, když vnější kritérium je právě to, které má měřit daný test), pak lze odvodit následující nerovnosti:

$$r_{tk} \leq \sqrt{r_T r_K} \leq \sqrt{r_T}$$

Tyto nerovnosti nám říkají, že empiricky prokazovaná kritériální validita testu bude vždy menší a nanejvýše rovna odmocnině z reliability testu. Nízká reliabilita testu tedy oslabuje potenciální korelační koeficienty výsledku testu s jinými vnějšími kritérii.

4.6.4 Chyba měření a její interpretace

Určení chyby měření se opírá o již výše uvedený vzorec

$$\sigma_e = \sigma_t \sqrt{1 - r},$$

⁹⁵ Viz např. Zvára, K., & Štěpán, J. (2002). *Pravděpodobnost a matematická statistika*. Praha: Matfyzpress.