

musí rozumět tomu, kdy a za jakých podmínek se test používá, jaký charakter mají pomocí něj získané výsledky, a musí je umět přetlumočit do řeči těch, komu je interpretace určena.

Podobně „důkazy“<sup>91</sup> zde nemají charakter původního a z matematiky převzatého významu jako logicky čistého a nezpochybnitelného postupu odůvodnění. Takovéto pojetí by totiž nedovolovalo adjektiva jako „kvalitní“ či „přesvědčivý“, jaké se ve spojení s nimi ve *Standardech* objevují. „Důkazy“ jsou zde spíše různé postupy ústící v argumentaci pro použití testů v té či oné oblasti za těch či oněch podmínek s nabízenou variabilitou interpretací výsledků. Důkaz tedy má svého čtenáře a nárokuje si jeho porozumění. Tím čtenářem je uživatel testu a je na něm, aby posoudil, zda jsou „důkazy“ validity pro jeho konkrétní případ k interpretaci dostatečně přesvědčivé a nápomocné.

Současné *Standardy* navrhují **členění zdrojů důkazů validity založených na analýzách:**

- obsahu (dříve obsahová v.)
- postupů odpovídání na testové položky (analýza procesu měření)
- vnitřní struktury testu
- vztahu k jiným proměnným (dříve kriteriální v.)
- důsledků testování<sup>92</sup> (např. zahrnutí informací o prospěšnosti měření a o oprávněnosti rozhodnutí na základě měření učiněných)

#### 4.6.2 Reliabilita a chyba měření

Reliabilita přímo souvisí s chybou měření. Čím má test vyšší reliabilitu, tím měří přesněji. Existují určité statistické metody odhadu reliability závislé na způsobu shromáždění dat o testu (viz níže). Pokud se tedy podaří na základě empirických dat odhadnout reliabilitu testu, lze z ní přímo určit chybu měření. Chybou měření se myslí nepřesnost, s jakou je určen výsledek žáka v určitém testu, jak moc se může lišit jeho testový výsledek od jeho skutečné dovednosti, kterou má test měřit.

Následující podrobné vysvětlení se opírá o tzv. klasickou teorii testů. Zmínka o modernějších teoriích testování je uvedena v závěru tohoto oddílu. K vysvětlení potřebujeme zavést některé pojmy:

- pravé skóre
- naměřené skóre
- chyba měření

**Pravé skóre** je hypotetická hodnota vlastnosti, kterou chceme zjišťovat a kterou má žák v čase testování. Tato hypotetická hodnota je myšlena v měřítku, které platí pro použitý test. Pokud by test měřil absolutně přesně, bylo by toto pravé skóre právě výsledek žáka v testu.

**Naměřené skóre** je výsledek žáka v testu.

**Chyba měření** je rozdíl mezi skutečnou hodnotou vlastnosti (pravé skóre) a výsledkem v testu.

Matematicky lze zapsat takto:

$$t = T + e,$$

kde  $t$  je naměřené skóre,  $T$  pravé skóre a  $e$  chyba měření.

<sup>91</sup> Možná by bylo vhodnější překládat „doklady“ validity (takto výjimečně ve *Standardech* přeloženo na s. 24).

<sup>92</sup> Konstruktová validita je zde zahrnuta v různých zdrojích důkazů a její pojetí v novém významu validity je shledáno jako nadbytečné.