

Výhody a nevýhody obou přístupů hodnocení

Holistická škála⁸⁶ vyžaduje obvykle menší množství kroků při rozhodování, může být tedy pro hodnotitele kognitivně méně náročná, a pokud je vyžadováno hodnocení většího množství prací a/nebo hodnocení jednoho výkonu větším počtem hodnotitelů, může být její využití levnější. Na druhé straně je interpretace výkonů testovaných u holistického přístupu omezena na globální popis, škály neumožňují vytvořit „profil“ testovaného, neinformují o jednotlivých rysech výkonu, o tom, co přesně výsledné skóre o dovednostech testovaného říká a do jaké míry je možné vyvodit obecné závěry o tom, jak by testovaný fungoval v mimotestové situaci. Různí testovaní s odlišným výkonem mohou dostat stejné hodnocení z odlišných důvodů (např. jeden text může být výborný po obsahové stránce a z hlediska koherence a koheze textu, avšak velmi slabý z hlediska použití jazykových prostředků, druhý text může být skvělý po stránce volby jazykových prostředků, ale jeho nedostatkem je nedodržování tématu, slabá argumentace apod.). Tento problém dokážou analyticky koncipovaná kritéria pojmenovat a postihnout, jsou tedy vhodnější pro diagnostické účely a pro vypracování „profilu“ testovaného z hlediska výkonu; avšak jejich implementace a používání bývají časově i finančně náročnější.

Dále holistická kritéria předpokládají, že se všechny hodnocené aspekty jazyka vyvíjejí stejně rychle, a mohou být tedy shrnuty do jednoho skóre. Analytická kritéria naopak vyčleňují jednotlivé komponenty řečové dovednosti a posuzují je odděleně. Zejména v případě první ukázky kritérií může být obtížné při hodnocení jednoznačně stanovit, do jaké úrovně posuzovaný výkon patří, pokud výkon naplňuje perfektně některé deskriptory (např. přesnost vyjadřování), avšak odporuje jiným (např. organizace textu).

Analytická kritéria mohou poskytovat lepší kontrolu nad chováním hodnotitele, protože je zcela zřejmé, jaké hodnocení a v jakém aspektu přidělil. Lze tedy v případě problematičnosti hodnocení dohledat dokonce i deskriptor, který podle hodnotitele popisuje jím posuzovaný výkon. U holistické škály lze pravděpodobně očekávat vyšší míru shody hodnotitelů, avšak u škály analytické je možné posuzovat a vyhodnocovat také vnitřní konzistenci hodnocení jednotlivých hodnotitelů i jejich shodu napříč skupinou hodnotitelů. U holistických kritérií může být problematické to, jakou váhu připisuje posuzovatel jednotlivým komponentám popsáním v deskriptorech: zdá se mu zásadnější koherence textu, nebo přesnost použitých jazykových prostředků?

Vhodnost jednotlivých přístupů souvisí také s tím, zda je hodnocen mateřský jazyk, nebo jazyk cizí. Někteří autoři se domnívají, že u rodilých mluvčích jde obvykle ruku v ruce dovednost psát dobře strukturovaný text s používáním vhodných jazykových prostředků a využití holisticky orientovaných kritérií se jeví jako vhodné. Naproti tomu u testovaných v cizím jazyce tomu tak být nemusí, testovaní mají často velmi různorodý profil dovednosti psát v cizím jazyce, a vhodnější je tedy spíše využití analytického přístupu, který rozdíl v „jazykovém profilu“ testovaného často lépe pojmenuje.

Jedním z nejdůležitějších předpokladů pro validní hodnocení však není ani tak volba přístupu k hodnocení, jako spíše to, jak mnoho, jak často a jak dobře jsou hodnotitelé školeni a monitorováni a jak je spolehlivost jejich výkonu analyzována.

⁸⁶ Carr, N. T. (2011). *Designing and analyzing language tests*. Oxford University Press.