

4.4.4 Úlohy se širokou odpovědí (esej test)

Úlohy se širokou odpovědí jsou vystaveny většímu riziku subjektivity při hodnocení. Voleny bývají zejména proto, že mohou být autentičtější ve smyslu větší podobnosti úkolům z reálného světa. Aby riziko subjektivity při hodnocení bylo co nejvíce eliminováno, bývá věnována zvýšená pozornost právě procesům hodnocení. Základem je mít dobře popsána kritéria hodnocení. Dalším důležitým faktorem je kvalitní výběr a zácvik hodnotitelů a procesy supervize nad vlastním hodnocením. Následující text je zaměřen převážně na úlohy z mateřského či cizího jazyka, i když popisované přístupy jsou přenositelné i na jiné obsahové oblasti. Částečně jsme se této problematice věnovali již v kapitole 2, kde jsme uváděli ukázkou kritérií pro hodnocení písemného projevu.

Většina institucí poskytujících zkoušky používá více či méně standardizované postupy hodnocení a interpretace výkonů pomocí sady deskriptorů popisujících výkon testovaného. V horizontálním pohledu se kritéria zaměřují na různé oblasti výkonu (splnění zadání, organizace textu, jazykové prostředky...), ve vertikálním pohledu pak na různé úrovně výkonu. Jednotlivé **deskriptory** (popisné, zde i hodnotící výroky) **musí být vztaženy ke konstrukt** (ověřované dovednosti) definovanému v testových specifikacích (konceptuálním rámci – viz podkapitola 4.1).

To, jaký přístup k hodnocení (analytický nebo holistický⁸⁴) bude použit a jak budou vypadat hodnotící kritéria (co a jak se hodnotí), souvisí velmi úzce např. s tím, jak je definována dovednost v konceptuálním rámci: zda zahrnuje splnění zadání, organizaci textu a koherenci či kohezi textu nebo jen aspekty používaného jazyka, tedy přesnost a rozsah použitých jazykových prostředků.

Volba typu hodnotící škály pak souvisí i s praktickými otázkami: např. kolik úloh nebo zadání a v jakém časovém limitu je třeba vyhodnotit, jak podrobné informace o výkonu testovaného chceme získat, kolik hodnotitelů je k dispozici, zda je proveditelné hodnocení více hodnotiteli.

Někdy také dochází k tomu, že pro hodnocení jedné zkoušky jsou používány oba typy přístupů – každý výkon je hodnocen nejméně dvěma hodnotiteli, jeden využívá holistická kritéria, druhý kritéria analytická, výsledek je potom kombinací obou hodnocení.

Holistický (globální) přístup k hodnocení

Používání holistických škál je založeno na konceptu jednotnosti a nedělitelnosti ověřované dovednosti. Uvažujme např. o jazykové kompetenci⁸⁵. Výkon testovaného je zde hodnocen z globálního pohledu, hodnotí se celková úroveň výkonu a textu je přiděleno označení pásma, do kterého byla práce zařazena. Každé z pásem hodnotící škály popisuje ideální úroveň výkonu v celé řadě posuzovaných aspektů očekávaného výkonu, práce testovaného je hodnocena jako celek a „přiřazena“ k tomu pásmu, jehož popisu nejlépe odpovídá. Přesto se v popisech nebo škálách často skrývá několik komponent, které jsou v jednotlivých pásmech hodnotící stupnice nebo škály hodnoceny či posuzovány. Příklad 4.44 obsahuje podrobné popisy stupňů škály bez vnitřního členění, příklad 4.45 zahrnuje v kritériích pro práci hodnotitelů čtyři strukturované komponenty jako klíčová vodítka pro celkové zhodnocení. V obou případech je výsledek jedince prezentován na jedné škále s popisem.

⁸⁴ Ukázky holistického a analytického přístupu jsme uváděli již v kapitole 2.

⁸⁵ Bachman, L. F. (2010). *Language assessment in practice*. Oxford University Press.