

- rozhovor s učitelem zaznamenávaný na video,
- rotace úkolů včetně vystoupení a práce ve skupině se záznamem odpovědí na počítači či papíře,
- úkoly typu „papír – tužka“.

Pro možnost srovnání výsledků 4. a 8. ročníků respektive vývoje v čase jsou do šetření zařazeny množiny propojujících testových položek. Konečně součástí šetření jsou také postojové otázky v rámci dotazníku. Otázky dotazníku se zaměřují mimo jiné na oblíbenost dílčích aktivit žáků, na typické chování žáků a jejich motivace či charakter domácího prostředí.

3. Administrace testu

S ohledem na skutečnost, že šetření NEMP a NMSSA nejsou tzv. *high-stake* šetření, je jejich administrace výrazně zjednodušena oproti národním testům. Takto jsou šetření NEMP a NMSSA realizována týmem proškolených učitelů, kteří jsou vybíráni z databáze zájemců, a to s důrazem na jejich kvalitu. Vyškolení učitelé postupně ve dvojicích navštěvují přidělené školy, kde realizují šetření s výběrovým souborem žáků.

Oprava testových položek je založena na speciálním postupu podle předem definovaných kritérií hodnocení. Úkoly s jednodušším vyhodnocením jsou známkovány studenty vysokých škol, úkoly s obtížnějším vyhodnocením pak profesionálními učiteli. Metodika šetření NEMP a NMSSA umožňuje měnit kritéria hodnocení v průběhu známkování v reakci na zjištěné poznatky. U úkolů, které vyžadují rozhodnutí známkujícího, jsou promíchány odpovědi 4. a 8. ročníků respektive dvou navazujících roků šetření tak, aby byla zajištěna možnost srovnání spolehlivosti hodnocení odpovědí.

4. Vyhodnocení testu

Vyhodnocení testu probíhá na základě následujícího metodického postupu. Pro jednotlivé úkoly je počítáno skóre žáků, jako procento správných odpovědí. Následně je počítáno celkové agregované průměrné skóre všech žáků pro jednotlivé úkoly. V dalším kroku jsou úkoly, tj. komponenty, zařazeny do tzv. klastrů znalostí¹¹⁶. Klastry znalostí jsou definovány experty projektového týmu v návaznosti na novozélandské kurikulum. Pro jednotlivé klastry znalostí je následně počítána hodnota agregovaného skóre pro všechny dílčí komponenty. Vztahy uvnitř klastrů znalostí a mezi klastry znalostí jsou hodnoceny korelačním počtem. Předpokládá se silná korelace komponent uvnitř klastrů a naopak slabší korelace komponent mezi klastry. Současně je aplikována faktorová analýza dosaženého skóre komponent.

Vyhodnocení vazeb je provedeno rovněž ve vztahu postojů žáků k jejich dosaženému skóre. Opětovně jsou aplikovány metody korelačního počtu a faktorové analýzy, a to pro postojové otázky konstruované na Likertově škále. Trendy jsou hodnoceny na bázi společných úkolů hodnocených v různých cyklech šetření, tedy ve čtyřletém intervalu. Hodnocení je provedeno na základě ukazatele *effect size* pro dosažené skóre, přičemž za rozdílově významnou je považována hodnota *effect size* vyšší než 0,30. U těchto rozdílů je následně provedena detailnější analýza odpovědí.

¹¹⁶ Pro oblast psaní jde například klastry základní znalosti (gramatika, spelling, interpunkce), hluboké znalosti (vhodné jazykové obraty, tón, slovník vzhledem k obecnosti) a pochopení (Gilmore et al. 2011).