

NÁZOV: **SKÚŠANIE, TESTOVANIE A HODNOTENIE V EDUKAČNOM PROCESE**

OBSAH

Úvod

1 KONTROLA – OVEROVANIE VÝSLEDKOV EDUKÁCIE

- 1.1 Funkcie skúšania a hodnotenia**
- 1.2 Testovanie a skúšanie, test a skúška**
- 1.3 Typy skúšania a skúšok – testov**
- 1.4 Skúšanie a hodnotenie**
- 1.5 Kvantifikácia hodnotenia, známkovanie**
- 1.6 Hodnotiace normy**

2 TESTY

- 2.1 Vlastnosti kvalitných testov**
- 2.2 Špecifikácia testu – skúšky**
- 2.3 Postup prípravy testu**
- 2.4 Základné typy testových úloh**

3 ANALÝZA A INTERPRETÁCIA VÝSLEDKOV TESTU

- 3.1 Miery centrálnej tendencie**
- 3.2 Miery rozptylu**
- 3.3 Analýza testových položiek**

4 ANGLICKO-SLOVENSKÝ TERMINOLOGICKÝ SLOVNÍK

- 5.1 Zoznam najčastejšie používanej terminológie**

Zoznam použitej literatúry

.

Úvod

Skúša a hodnotí sa obyčajne vtedy, keď treba v krátkom časovom úseku a objektívne zistiť, či sa dosiahli ciele vzdelávacích podujatí, výcvikov, vyučovania alebo edukácie vôbec. Skúšaním a hodnotením sa overuje, či sa školení jednotlivci naučili to, čo je potrebné na získanie určitej kvalifikácie, či môžu postúpiť do vyššej fázy vzdelávania alebo získať určitý post v spoločenskej praxi.

Cieľové vedomosti a schopnosti sa overujú rôznymi skúškami a testami. Za skúšky môžeme považovať rôzne konkurzy, prijímacie skúšky do škôl, kvalifikačné skúšky a najmä rôzne typy skúšania vo vyučovaní. Na základe ich výsledkov prijímajú zodpovední pracovníci dôležité rozhodnutia, ktoré často zásadne ovplyvnia budúcnosť skúšaných jednotlivcov. Neobjektívne hodnotenie a neprímerané skúšanie sa bytostne dotkne každého človeka a tým citlivejším môže dokonca spôsobiť traumu.

Skúšanie možno považovať za meranie, ktoré by malo byť čo najpresnejšie a najobjektívnejšie, to znamená, že na všetkých skúšaných a hodnotených jedincov treba používať rovnaký „meter“. Meranie psychických vlastností je oveľa zložitejšie ako meranie fyzikálnych veličín, kedy je vzťah medzi veličinou a jednotkami merania jednoznačne jasný a postihnuteľný aj zmyslami. Skúšanie a hodnotenie vedomostí, zručností, schopností a vlastností osobnosti teda nevyhnutne vyžaduje odborný prístup, ktorého základom sú vedomosti o podstate týchto procesov a o možnostiach ich čo najobjektívnejšieho zmerania. Pri dodržaní určitých pravidiel možno dosiahnuť vysokú mieru objektívnosti aj pri meraní zložitých psychických vlastností.

Otázkou je, nakoľko sú tí, ktorí skúšajú a hodnotia iných, pripravení vykonávať tieto náročné činnosti tak, aby boli všetci zúčastnení spokojní a aby skúšanie splnilo svoj účel. Domnievame sa, že úroveň odbornej pripravenosti na procesy skúšania a hodnotenia je veľmi rôznorodá a mnohí skúšajúci pripúšťajú potrebu vedieť o problematike viac.

Predkladáme vám materiál, ktorý je určený najmä tým učiteľom a tvorcom testov, ktorí chcú poznať základné východiská a princípy skúšania a tvorby testov. Minimum nevyhnutnej teórie sme sa snažili podať jednoduchou a zrozumiteľnou formou s uvedením príkladov z praxe. Prvé tri kapitoly napísala Milena Štulrajterová, terminologický slovník vypracovala Ríta Rafajlovičová.

V texte sme vysvetlili základné pojmy v skúšaní a hodnotení, uvádzame definície skúšania a hodnotenia ako aj kategorizáciu testov.

Materiál ďalej objasňuje požiadavky na kvalitné testy a prináša stručný všeobecný návod na vypracovanie špecifikácie testu i všeobecný postup prípravy testov s príkladmi rôznych typov testových úloh.

Podstatná časť materiálu je venovaná hodnoteniu testových úloh a možnostiam interpretácie testových výsledkov. V tejto časti sú uvedené elementárne metódy kvantitatívneho spracovania a interpretácie výsledkov skúšania. Uvádzané štatistické operácie zvládne nepochybne každý čitateľ, ktorý má záujem vyťažiť zo skúšania maximum informácií.

Materiál môže byť úvodom do podrobnejšieho a hlbšieho štúdia problematiky pre tých, ktorí sa chcú zaoberať skúšaním a hodnotením na profesionálnej úrovni. Najmä im, ale i ostatným je určená posledná časť materiálu, ktorou je anglicko-slovenský slovník najpoužívanejších odborných termínov. Zdrojom slovenských ekvivalentov anglických termínov bol najmä Anglicko – slovenský pedagogický slovník (Gavora, Mareš, 1996) a ďalšie pramene uvedené v bibliografii (slovenské a český). V prípadoch, kedy sa nám nepodarilo nájsť zodpovedajúci slovenský ekvivalent, sme sa snažili vystihnúť význam termínu definíciou alebo opisom. Naše vysvetlenia vychádzajú okrem iného aj z materiálu Európskej rady, Spoločný európsky referenčný rámec.

V materiáli sa odráža skutočnosť, že autorky sú učiteľky jazyka, a preto uvádzajú väčšinou príklady zo skúšania a prípravy testov jazykových kompetencií. Základné východiská, princípy a poznatky však platia v skúšaní a príprave testov všeobecne, a preto môže slúžiť ako úvod do problematiky všetkým, ktorých daná oblasť zaujíma.

Autorky

Skúšanie, hodnotenie a príprava testov

1 Kontrola – overovanie výsledkov edukačného procesu

sa realizuje skúšaním a hodnotením. Predmetom kontroly je práca žiakov a učiteľov. Výsledky edukačného procesu sa posudzujú vzhľadom na krátkodobé a dlhodobé ciele stanovené v učebných osnovách, cieľových štandardoch, časovo-tematických plánoch a plánoch vyučovacích hodín.

1.1 Funkcie skúšania a hodnotenia

- overovanie efektívnosti vyučovania;
- zisťovanie celkovej úrovne dosahovania cieľov;
- zisťovanie postupu učenia a rozvoja jednotlivých žiakov;
- zhromažďovanie podkladov na klasifikáciu žiakov;
- získavanie informácií pre rozhodnutia o ďalšom vzdelávaní žiakov;
- informovanie žiakov o ich výkonoch, postupe, zameraní ďalšieho úsilia v zmysle možností rozvoja ich individuálneho potenciálu;
- informovanie rodičov a iných zainteresovaných inštitúcií;
- udržiavanie motivácie žiakov oceňovaním pozitívnych stránok ich výkonov;
- rozvíjanie sebareflexie a konštruktívnej sebakritiky žiakov aj učiteľov;
- zámerné a systematické poznávanie vedomostí, zručností, schopností, postojov a osobných vlastností žiakov.

1.2 Testovanie a skúšanie, test a skúška

Slová test a testovať boli prebrané z angličtiny a zdomácnili natoľko, že vytláčajú naše slová rovnakého významu, t. j. skúška a skúšať. V angličtine má slovo test aj ďalšie podobné významy. Označuje aj postup overovania kvality niečoho, spôsob zisťovania prítomnosti požadovaných vlastností niečoho. Napr. test trvanlivosti výrobku, test čistoty ovzdušia, test hladiny cukru v krvi, a pod. V slovenčine chápajú mnohí slová test a testovať zúžene, najčastejšie ako skúšku písomnú, ktorú možno objektívne ohodnotiť. V tomto zmysle je

zaužívané spojenie didaktický test, ktorý je obyčajne objektívnou písomnou skúškou vedomostí.

V pedagogickom slovníku je uvedené, že test je skúška, označenie prebraté z angličtiny. V pedagogike ide o nástroj merania výkonu (Prúcha, Walterová, Mareš, Pedagogický slovník, 1995).

Kompolt hovorí o edukačnom testovaní a diagnostikovaní, slovo skúšanie vôbec nepoužíva. Test je podľa neho špecifickým druhom skúšky, procedúrou na meranie vzorcov správania. Výsledky testovania sú vždy kvantifikované (Švec, 1998). Na inom mieste hovorí: „Test je súbor položiek reprezentujúcich vopred stanovené učebné ciele, ktorých stupeň splnenia chceme testom zisťovať“ (Švec 1998, s.279). Je toho názoru, že na Slovensku sa nepochopiteľne preferujú len testy písomné a ostatné testy sa ignorujú.

Britskí odborníci v oblasti skúšania cudzích jazykov používajú slovo test pre skúšky písomné aj ústne, hodnotené objektívne i subjektívne. Testom sa rozumie postup vytvorený na vyvolanie takého správania, z ktorého možno vyvodzovať závery o určitých charakteristikách jedinca (Bachman 1991, s.20). Táto definícia testu je veľmi široká. Ďalej však Bachman hovorí, že test musí dodržiavať štandardizovaný postup a výsledky sa kvantifikujú podľa explicitných pravidiel. V tomto zmysle je testovanie zadávaním starostlivo vybraných a na dostatočnej vzorke osôb odskúšaných úloh. Prebieha za štandardných podmienok, výkony jedincov sa vyhodnotia podľa predpísaných pravidiel, interpretujú sa podľa príslušných noriem.

Bachman sa domnieva, že slová testovať a skúšať (v angličtine *to test* a *to examine*) sú len štylistické varianty toho istého významu. V súvislosti so základným a stredným stupňom vzdelávania sa hovorí o testovaní – skúšaní, v súvislosti s vysokými školami sa používa slovo *exam* – skúška.

U nás sa skúškou obyčajne označuje ústne formálne skúšanie na vysokej škole, alebo záverečné skúšanie za účelom ukončenia štúdia či získania určitej kvalifikácie (maturitná skúška, štátna záverečná skúška) s formalizovaným priebehom. Skúška môže byť ústna, písomná, praktická, alebo môže pozostávať z viacerých častí (napr. vodičská skúška pozostáva z písomnej časti a časti praktickej). Zvládnutie jednej časti skúšky môže byť podmienkou pripustenia účasti na druhej časti skúšky.

Skúšanie a testovanie sa nerozlišuje podľa toho, či prebieha na základnej alebo vysokej škole, či sa skúša písomne alebo ústne, ale podľa toho, ako sa skúša a hodnotí. O testovaní

môžeme hovoriť vtedy, ak skúšanie priebeha v štandardizovaných podmienkach, ak sa výkon hodnotí podľa príslušných noriem a je kvantifikovaný podľa dopredu stanovených pravidiel. V tomto zmysle je každé testovanie skúšaním, no skúšanie nemusí byť testovaním ak nespĺňa uvedené podmienky.

Cieľom skúšania aj testovania je čo najobjektívnejšie a najpresnejšie zistenie úrovne kompetencií jedincov a skupín. Súhrnné štandardizované testy pre potreby škôl sú z viacerých dôvodov v nedohľadne. Príprava štandardizovaných testov je náročná činnosť, ktorou sa zaoberajú tímy odborníkov. Základné princípy testovania si však môže osvojiť a následne uplatňovať v skúšaní každý skúšajúci.

Žiaci, študenti a iní ľudia, ktorým je test zadávaný, resp. ktorí robia test alebo skúšku, sa nazývajú skúšaní, respondenti alebo probandi.

Osoby, ktoré realizujú, riadia a usmerňujú priebeh skúšania, sa nazývajú skúšajúci alebo examinátori (*test administrators*). Osoby, ktoré dohliadajú na to, aby respondenti pri testoch nepodvádzali a aby boli dodržané štandardné podmienky pre všetkých skúšaných jedincov, sa nazývajú dozor (*invigilators*).

Odborník, ktorý realizuje skúšanie ústneho prejavu v cudzom jazyku, vedie zahrievací rozhovor, prípadne participuje v rolovej aktivite alebo diskusii, sa nazýva examinátor (*interlocutor*).

1.3 Typy skúšania, testovania a testov

Skúšanie možno kategorizovať podľa rôznych hľadísk. Pokúsili sme sa o vyčerpávajúci prehľad, pričom sme využili kategorizáciu Hughesa a Türeka.

1. Podľa charakteru vyvolanej vzorky správania na:

- písomné, žiaci sa vyjadrujú písomne (*written verbal tests*);
- ústne, žiaci sa vyjadrujú ústne (*oral tests*);
- praktické, žiaci vykonávajú senzomotorickú činnosť, napr. cvičia, spievajú, kreslia (*performance tests*).

2. Podľa počtu súčasne skúšaných žiakov na:

- individuálne;
- skupinové;

- *frontálne*, otázky dostávajú postupne viacerí systematicky alebo náhodne vyberaní žiaci v celej skupine.

3. Podľa časového zaradenia a rozsahu skúšaného obsahu na:

- *priebežné* – skúšanie pokroku v učení, overuje sa plnenie cieľov za kratšie obdobie (*progress tests*);
- *súhrnné*, overuje zvládnutie cieľov stanovených na dlhšie časové úseky (*achievement tests*).

4. Podľa funkcie, účelu skúšania na:

- *záverečné*, napr. maturitné skúšanie (*final tests*);
- *opravné*, napr. na vysokej škole skúška, ktorú študent opakuje, ak ju prvýkrát nezvládne;
- *prijímacie* (*entrance* alebo *admission tests*);
- *rozmiestňovacie*, napr. zaradenie do skupín podľa počiatočnej úrovne jazykovej kompetencie pri vyučovaní cudzích jazykov (*placement tests*);
- *diagnostické*, cieľom je zistiť, či celá skupina zvládla učivo na požadovanej úrovni a podľa toho zamerať ďalšie vyučovanie (*diagnostic tests*);
- *kvalifikačné*, zisťuje sa spôsobilosť pre určité odborné činnosti alebo napr. celková úroveň ovládania cudzích jazykov (*proficiency tests*).

5. Podľa toho, či žiak vie, že je skúšaný, na:

- *formálne*, obyčajne priebežné alebo súhrnné klasifikované skúšanie (*final tests*);
- *neformálne*, žiak nevie, že učiteľ bude hodnotiť jeho výkon, napr. oznámkovanie niekoľkých žiakov, ktorí dobre obstáli v skupinovej aktivite alebo sústavne správne odpovedajú pri frontálnom overovaní vedomostí.

6. Podľa toho, akým spôsobom sa dajú hodnotiť odpovede, na:

- *objektívne*, kedy možno jednoznačne povedať, či je odpoveď správna alebo nesprávna (takto sú koncipované didaktické testy);
- *subjektívne*, odpovede majú taký charakter, že sa nedajú posudzovať ako správne a nesprávne, ale ako viac alebo menej kvalitné, podľa toho, nakoľko zodpovedajú

hodnotiacim kritériám. O hodnote odpovedí rozhoduje hodnotiaci subjekt. Pri tomto type hodnotenia viacerí hodnotiaci nemusia ohodnotiť jeden výkon úplne zhodne, t. j. rovnakým počtom bodov, známku, alebo stupňom na stupnici. Dôvodom môžu byť nejasné, nejednoznačné a neúplné kritériá, čo vedie k ich rozdielnemu pochopeniu hodnotiacimi. Príčinou nejednotného hodnotenia jedného výkonu môže byť aj nedostatočná odborná príprava a zácvik hodnotiacich. Takto ponímaná subjektívnosť teda neznamená, že hodnotenie je nespravodlivé alebo neobjektívne, ale že ide o kvalitatívne posudzovanie komplexných kompetencií vo výkone, kedy nie je možné dosiahnuť stopercentnú zhodu medzi niekoľkými hodnoteniami jedného výkonu. Čím jasnejšie sú definované hodnotiace kritériá, tým väčšiu zhodu v hodnotení možno dosiahnuť. Subjektívne testy sú potrebné a nevyhnutné, lebo merajú kompetencie vyššieho rádu, overujú dlhodobé ciele a majú pozitívny dopad na proces vyučovania a učenia.

7. Podľa toho, či sa výkon jednotlivca porovnáva s normou alebo kritériom na:

- *rozlišujúce*, zamerané na zisťovanie relatívneho výkonu (*norm referenced*), výkony jednotlivcov sa porovnávajú so sociálnou normou alebo populačným štandardom a sú hodnotené výkonovými stupňami – známkami;
- *overujúce*, zamerané na zisťovanie absolútneho výkonu (*criterion referenced*), zisťuje sa, či výkon spĺňa alebo nespĺňa dopredu stanovené kritérium, hodnotenie potom zahŕňa len dve možnosti, vyhovel alebo nevyhovel.

8. Podľa charakteru úloh na:

- *priame (direct)*, kedy sú skúšobné úlohy založené na komplexných kompetenciách, t.j. vyžadujú aplikáciu vedomostí, zručnosti, schopnosť riešiť teoretické a praktické problémy, tvorivé a hodnotiace myslenie. Dôraz sa kladie na to, aby to boli úlohy imitujúce situácie reálneho života, teda aby boli autentické. V cudzom jazyku sú to napr. rolové aktivity s otvoreným riešením, diskusie za účelom dosiahnutia názorovej zhody alebo riešenia problému, voľná konverzácia, napísanie formálneho listu. Skúšanie tohto typu je obyčajne náročné na čas, výber reprezentatívnej vzorky úloh ako aj hodnotenie, má však pozitívny vplyv na učenie a učenie sa. Inak sa takéto skúšanie označuje ako integrované (*integrated*), lebo každá úloha overuje niekoľko subkompetencií, napr.

prekladom textu skúšame slovnú zásobu, morfológiu, syntax, kultúrne znalosti, výstižnosť vyjadrovania;

- *nepriame (indirect)*, skúšobné úlohy predpokladajú využitie niektorých čiastkových kompetencií, na základe ktorých nie je možné s istotou usudzovať na schopnosť plnenia komplexných úloh. Ide o úlohy, ktoré sú neautentické, to znamená, že sa v reálnom živote v takej podobe ako pri skúšaní nevyskytujú vôbec alebo len výnimočne. Napríklad test slovnej zásoby, použitia slovesných časov, opravovanie chýb v texte, dopĺňanie chýbajúcich slov, test výslovnosti jednotlivých slov, čítanie nahlas. Testy využívajúce nepriamy prístup možno obyčajne rýchlo a objektívne opraviť, preto ich skúšajúci využívajú vo vysokej miere. Nepriamy prístup sa využíva v testoch so zatvorenými objektívne skórovateľnými úlohami (*discrete point item tests*), ktorými sa obyčajne overujú vedomosti a kompetencie založené na algoritmických postupoch riešenia úloh.

9. Podľa toho, či úloha vyžaduje od respondenta samostatnú produkciu celej dlhšej alebo kratšej odpovede alebo, či respondent iba určí správnu odpoveď z daných dvoch alebo viacerých možností, na:

- *skúšanie s produktívnymi – otvorenými úlohami (production tasks)*;
 - *skúšanie s úlohami, v ktorých treba rozpoznať správnu odpoveď (recognition tasks)*,
- kým produktívne úlohy vyžadujú dôkladné osvojenie vedomostí a zručností, úlohy na rozpoznanie správnej odpovede môžu byť menej náročné a ich splnenie môže byť pri alternatívnych odpovediach založené na hádaní, pričom pravdepodobnosť uhádnutia správnej odpovede je až 50 %.

1.4 Skúšanie a hodnotenie

Procesy skúšania a hodnotenia spolu úzko súvisia, ale nie sú totožné. Skúšanie je vyvolávaním vzoriek správania za účelom hodnotenia. Hodnotiť možno už v priebehu skúšania (napr. pri ústnom skúšaní učiteľ v duchu porovnáva odpovede s normou, zaznamenáva chyby, preruší žiaka a opravuje ho, oceňuje dobré odpovede) alebo až po skúšaní (oprava a vyhodnotenie písomného testu, oznámenie výsledkov žiakom). Cieľom skúšania je hodnotenie výkonu. Hodnotiť však možno aj bez formálneho skúšania na základe pozorovaní.

Hodnotením je podľa Kosovej v širšom slova zmysle každý verbálny alebo inak vyjadrený hodnotiaci súd, mienka alebo postoj subjektom hodnotenia objektu hodnotenia o jeho činoch alebo vlastnostiach (správne, veľmi dobre, pracuješ veľmi dobre, nie som s tebou celkom spokojná, zaujímavé, prikývnutie, úsmev, gesto). Hodnotením v užšom slova zmysle je porovnávanie objektu hodnotenia s reálnym alebo ideálnym vzorom, s prítomnosťou, absenciou alebo stupňom reálnej alebo ideálnej hodnoty vnímanej tým, kto hodnotí (Kosová 96/97, s.1).

Skúšanie možno chápať ako proces zámerného vyvolávania takých vzoriek správania, v ktorých sa prejaví množstvo a kvalita určitých vedomostí, zručností, schopností a vlastností osobnosti. Vzorky správania sa vyvolávajú zadávaním rozmanitých úloh, ktorých splnenie vyžaduje práve tie vedomosti, zručnosti a schopnosti, ktoré sú predmetom nášho záujmu, ktoré chceme ohodnotiť. Vyvolané vzorky správania sú výkony, ktoré sú predmetom hodnotenia. Výkony sa hodnotia kvalitatívne alebo kvantitatívne podľa rôznych noriem. Ak sa výkonom priradujú kvantitatívne hodnoty podľa dopredu stanovených pravidiel, hovoríme o meraní výkonov.

Merané vedomosti, zručnosti, schopnosti a iné cieľové vlastnosti sa inak nazývajú aj kompetenciami. Kompetencie majú svoje teoretické definície, ktoré by mali byť čo najpresnejšie, aby bolo možné rôzne kompetencie navzájom odlišiť. Teoreticky definovaná kompetencia sa nazýva konštrukt. Príkladmi kompetencií sú: praktické ovládanie cudzieho jazyka, vizuálna pamäť, tvorivé myslenie. Na rozdiel od priamo pozorovateľných fyzických vlastností žiakov, ako sú výška, hmotnosť, farba vlasov, učiteľ nevie, aké psychické kompetencie majú žiaci, kým sa tieto neprejavujú v správaní. Kompetencie teda nemeráme priamo, ako výšku postavy, hmotnosť, teplotu a pod., ale nepriamo na ne usudzujeme na základe ich pozorovateľných prejavov v správaní.

Prejavy správania sú rôzne činnosti alebo iným slovom povedané operácie. Aby sme vedeli posúdiť prítomnosť a mieru kompetencií, je potrebné ich operacionalizovať. „Operacionalizácia je prevedením požadovanej žiakovej vedomosti či zručnosti do činnosti (operácie). Táto procedúra umožňuje jednoznačné potvrdenie toho, či žiak danú činnosť zvládol alebo nezvládol, prípadne potvrdenie stupňa jej zvládania“. (Švec a kol. 1998, s.277) Vonkajšie prejavy cieľových kompetencií žiakov jednotlivých ročníkov by mali byť operacionalizované, t. j. konkrétne popísané v cieľových vzdelávacích štandardoch. Vzdelávacie štandardy špecifikujú, aké vedomosti, zručnosti, schopnosti a vlastnosti má žiak

nadobudnúť, a najmä, ako sa tieto prejavia v činnosti, t.j. aké konkrétne operácie - činnosti dokáže úspešne robiť žiak, ktorý dané kompetencie má. Napr. matematická kompetencia na určitej úrovni môže byť operacionalizovaná nasledovne: žiak vie sčítavať a odpočítavať čísla do desať; kompetencia v zemepise: žiak vie vymenovať hlavné mestá všetkých európskych štátov; kompetencia v cudzom jazyku: vie sa predstaviť vo formálnej aj neformálnej situácii, vie tvoriť prítomný čas slovesa „byť – to be“ v oznamovacom, opytovacom spôsobe, vie tvoriť druhý a tretí stupeň osvojených prídavných mien, dokáže gramatické štruktúry správne použiť v kontexte.

Relatívne najjednoduchšie a jednoznačne možno operacionalizovať faktické vedomosti, ktorých prejavom je ich presná reprodukcia, vymenovanie alebo presné definovanie. Rádovo vyššou kompetenciou je praktické využívanie vedomostí pri riešení rôznych úloh. Tu už je potrebné konkrétne opísať algoritmy úloh, ktoré žiaci majú vedieť riešiť.

Operacionalizovať zložité kompetencie nie je vôbec jednoduché. Odizolovanie jednotlivých kompetencií a ich presné definovanie je predmetom dlhodobého výskumu. Splnenie jednej úlohy obyčajne vyžaduje využitie viacerých kompetencií. Napr. riešenie slovnej matematickej úlohy môže vyžadovať uplatnenie vedomostí, použitie zapamätaných vzorcov ako aj logické uvažovanie. Na druhej strane sa jedna a tá istá kompetencia využíva pri plnení rôznych úloh (napr. pamäť sa uplatňuje v reprodukcii vedomostí, v zručnostiach i v tvorivých úlohách). Vzhľadom na súčinnosť a presah kompetencií je ich meranie veľmi náročné a nemôže byť dokonale presné.

Vzdelávacie štandardy u nás zatiaľ nemáme k dispozícii, preto sa pri skúšaní a hodnotení žiackych výkonov riadime učebnými osnovami a klasifikačným poriadkom. Tieto dokumenty neposkytujú v konkrétnych situáciách dostatok informácií potrebných na dôkladné hodnotenie výkonov.

Ak sa v našich podmienkach hovorí o tom, že učiteľ žiaka na hodine skúšal, predpokladáme, že žiakovi dával úlohy a že úroveň ich splnenia ohodnotil známku od 1 do 5. Znamka v tomto prípade predstavuje umiestnenie výkonu žiaka na poradovej stupnici, kde prvý stupeň znamená najlepší výkon a piaty stupeň výkon nedostatočný, ktorý nedosahuje ani klasifikačným poriadkom stanovené minimum, čiže kritérium. Skúšanie v škole teda obyčajne zahŕňa aj ohodnotenie výkonu známku – klasifikáciu.

Skúšanie a hodnotenie môžu prebiehať paralelne, keď skúšajúci sleduje výkon skúšaného, porovnáva ho s očakávaným výkonom, robí si poznámky v duchu alebo na papier.

Hodnotenie je proces porovnávanía výkonu v priebehu skúšania, ale aj konečné zhodnotenie výkonu a oznámenie výsledku skúšanému jedincovi. Oznámenie výsledku skúšania po jeho ukončení môže mať formu sumatívneho alebo formatívneho hodnotenia.

1.4.1 Sumatívne hodnotenie

Sumatívne hodnotenie sa podľa Rötlinga (Rötling 1977, s.5) chápe ako hodnotenie výsledkov po ukončení jednotiek učiva, tematických celkov ako aj celkových výstupov. Sumatívne hodnotenie sa inak označuje aj ako hodnotenie záverečné alebo bilančné. Má sa realizovať vtedy, keď sa naplnia podmienky na osvojenie učiva. Pri sumatívnom hodnotení sa výkon skúšaného porovnáva so sociálnou normou, populačnou normou alebo kritériom a je vyjadrené kvantitatívne známkou vyjadrenou číslom (u nás 1-5 v ZŠ a SŠ), slovom (výborne, chválitebne), písmenom (v Anglicku A, B, C, D, E) alebo počtom bodov (napr. 75 bodov zo 100). Sumatívne hodnotenie teda výkony klasifikuje, zaraďuje do výkonnostných skupín. V procese edukácie sumatívne hodnotenia rozhodujú o postupe do vyšších etáp vzdelávania (známky na konci školského roka, výsledky prijímacích skúšok), na ich základe sa udeľujú osvedčenia a certifikáty oprávňujúce ich držiteľov kvalifikovane vykonávať rôzne činnosti na trhu práce (napr. certifikáty o úrovni ovládania cudzích jazykov, vodičské oprávnenia, inštruktorské oprávnenia).

1.4.2 Formatívne hodnotenie

Formatívne hodnotenie sa v edukačnom procese považuje z hľadiska rozvoja osobnosti žiaka za mimoriadne dôležité a potrebné. „Formatívne hodnotenie je hodnotenie priebežné, ktorého cieľom je získať informácie o priebehu procesu, javu, o aktuálnom stave, momentálnej výkonnosti; diagnostikovať prípadné závady, odchýlky, typické chyby, tak, aby bolo možné včas poskytnúť spätnú väzbu všetkým zúčastneným, urobiť potrebné zmeny a odstrániť prípadné nedostatky.“ (Pedagogický slovník 1995) Spočíva v tom, že učiteľ verbálne hodnotí výkon žiaka, pokrok v učení za určitý čas, pričom jeho výkon neporovnáva len s normou, ale aj s možnosťami žiaka samotného vzhľadom na jeho potenciál alebo predchádzajúci výkon. Učiteľ v prvom rade vyzdvihne to, čo žiak dokázal, čo bolo v jeho výkone pozitívne. Konkrétne informuje žiaka o tom, čo a ako treba zlepšiť, akým smerom upriamiť ďalšiu

aktivitu. Takéto hodnotenie má byť stručné, zrozumiteľné a obsažné, aby bolo pre žiaka informatívne a pozitívne motivujúce. Formatívne hodnotenie by malo nasledovať bezprostredne po každom individuálnom i skupinovom formálnom skúšaní. Formatívne hodnotenie nevyklučuje klasifikáciu. Žiak môže dostať známku a zároveň formatívne hodnotenie, pričom tieto nemusia byť v zhode. Formatívne hodnotenie môže byť veľmi pozitívne, ak priemerný žiak pracuje usilovne a robí pokroky, hoci vzhľadom na normu dostáva obyčajne známku dobrý. Naopak, priebežné hodnotenie žiaka, ktorý podáva priemerný výkon, hoci má vysoký potenciál, môže byť negatívne, lebo tento nepracuje primerane svojim možnostiam. Z formatívneho hodnotenia sa žiak dozvedá o kritériách hodnotenia, o možnostiach nápravy, o tom, na čo sa má v budúcnosti zamerať, ako sa môže zlepšiť. Formatívne hodnotenie umožňuje prežívať pocit úspešnosti a vlastnej hodnoty aj tým žiakom, ktorí možno nikdy nebudú v sumatívnom hodnotení zaradení do najvyšších výkonových skupín. Zážitok úspechu stimuluje aktivitu žiaka, ovplyvňuje jeho vzťah k predmetu ako aj vzťah k sebe samému. Kým sumatívne hodnotenie klasifikuje výkony a umožňuje ich porovnávanie, formatívne hodnotenie kladie dôraz na proces učenia a vedie k jeho následnej optimalizácii.

Výlučné používanie sumatívneho hodnotenia v škole spôsobuje porovnávanie žiakov medzi sebou, môže vyvolať nezdravé súperenie a komplexy menejcennosti u priemerných žiakov. Sumatívne hodnotenie je pre učiteľa výhodné v tom, že je stručné, no vo výchovnom procese je nedostatočné a nemôže byť náhradou za hodnotenie formatívne.

Projekty, písomné práce a iné produkty praktických skúšok (kresby, artefakty, jednoduché výrobky) by tiež nemali byť hodnotené len známku, známka by mala byť zdôvodnená ústne alebo v stručnom písomnom komentári.

1.5 Kvantifikácia hodnotenia a známkovanie

1.5.1 Objektívne hodnotenie

Pri testoch, ktoré možno hodnotiť objektívne, je prisudzovanie bodových hodnôt za jednotlivé položky, resp. odpovede jednoduchšie ako pri testoch subjektívnych. Vzhľadom na to, že odpoveď môže byť iba správna, alebo nesprávna, prisúdime jej určitý vopred stanovený počet bodov alebo žiadne body. V praxi sa zvyčajne udeľuje vyšší počet bodov za úlohy, ktoré vyžadujú produkciu odpovedí, ako za tie, ktoré vyžadujú iba receptívne zručnosti, t. j. úlohy,

pri ktorých respondent iba rozpozná správnu odpoveď. Niekedy sa presne stanovia kritériá pre čiastočné splnenie úloh, kedy sa prideli len časť z maximálneho možného počtu bodov, napr. dva body z piatich možných za správny postup riešenia príkladu s nesprávnym výsledkom, štyri body zo šiestich za dve správne preložené časti vety pozostávajúcej z troch častí a pod.

1.5.2 Subjektívne hodnotenie

Pri hodnotení komplexných otvorených úloh sa posudzuje ich celková kvalitatívna úroveň podľa stanovených kritérií. Hodnotenie otvorených produktívnych úloh vyžaduje dôkladnú a detailnú prípravu kritérií hodnotenia, lebo inak je pravdepodobné, že rôzni známkujúci prisúdia body za nepodstatné charakteristiky výkonu ako napr. za neestetický rukopis a nedokonalú úpravu, t. j. za irelevantné kompetencie, ktoré nie sú predmetom skúšania. Navyiac, dvaja známkujúci sa môžu značne líšiť v počte pridelených bodov, jeden z nich môže byť prísnejší ako druhý, najmä ak nerešpektuje kritériá hodnotenia, ktoré si dôkladne neosvojil. Vychádzajúc z týchto faktov, je veľmi dôležité, aby pre hodnotenie subjektívnych testov bol vypracovaný záväzný systém, prinajmenšom stručný popis toho, čo má respondent splniť, resp. dosiahnuť v tej-ktorej oblasti, aby mu bola pridelená určitá známka, alebo počet bodov. Ideálna je možnosť použitia hotových štandardizovaných **analytických a globálnych hodnotiacich stupníc**.

Hodnotiace stupnice pozostávajú z popisov kvalitatívnych požiadaviek pre jednotlivé úrovne výkonu. Ak použijeme analytickú hodnotiacu stupnicu, hodnotíme každé kritérium zvlášť. Napr. pri hodnotení ústneho prejavu budeme mať zvlášť stupnicu, podľa ktorej budeme hodnotiť plynulosť, gramatickú presnosť, rozsah slovnej zásoby, a pod. Výsledkom hodnotenia bude profil stupňov - hodnôt pre každé kritérium. Ak použijeme globálnu hodnotiacu škálu, priraďujeme výkon k jednej úrovni na stupnici. Globálna stupnica môže mať napríklad päť kvalitatívnych stupňov. V každom stupni sú opísané všetky dôležité kvality požadovaného výkonu .

1.5.3 Hodnotenie metódou celkového dojmu a analytické hodnotenie

Subjektívne hodnotenie možno realizovať dvomi spôsobmi: **metódou celkového dojmu** (*holistic method*) alebo **analyticky** (*analytic method*). Prvá z metód je založená na hodnotení podľa celkového dojmu. Pri tejto metóde úlohu hodnotí niekoľko hodnotiteľov, každý z nich ju oznámkuje, známky sa porovnávajú, zdôvodnia a dohodne sa výsledná známka. I keď ide o

hodnotenie podľa dojmu, každý hodnotiaci vychádza z viac alebo menej uvedomelých kritérií, ktoré on sám považuje za najdôležitejšie, čo znamená, že viacerí hodnotiaci nemajú úplne totožné kritériá. V prospech globálneho hodnotenia na základe celkového dojmu hovorí skutočnosť, že celok a výsledný efekt diela (napr. eseje, prednesu básne, namaľovaného obrázku) nie je len súčtom kvalít jednotlivých častí. Na celkovom výsledku sa časti nepodieľajú rovnakým dielom, navyše sú v zložitej interakcii a ovplyvňujú sa navzájom. Keďže väčšina učiteľov nemá možnosť využiť pri známkovaní služby niekoľkých kolegov, vhodnejšou sa javí analytická metóda. Táto metóda je založená na hodnotení výkonov podľa starostlivo pripravených hodnotiacich schém alebo stupníc, ktoré spravidla zahŕňajú hodnotenie všetkých podstatných zložiek výkonu (napr. pri slohovej práci sa zvlášť hodnotí gramatická a štylistická správnosť, zrozumiteľnosť, relevantnosť obsahu, atď.). Táto procedúra hodnotenia je prijateľná v kontexte vyučovacieho procesu, pretože žiaci majú jasnú predstavu, na základe čoho získali danú známku, a pri hodnotení nie je opomenuté žiadne podstatné kritérium.

Hodnotenie na základe celkového dojmu je výhodné v tom, že je rýchlejšie ako hodnotenie analytické. Aj napriek vyššej objektívnosti analytického hodnotenia odporúčame kombinovať obidve metódy. Po dlhodobom používaní analytickej metódy sa môžu hodnotiace kritériá zvnútorniť do takej miery, že hodnotiaci prichádzajú k rovnakému výsledku aj pri použití metódy dojmu. Vždy je však potrebné, aby obsah kritéria bol jednoznačne rovnako chápaný všetkými hodnotiacimi a aby si kritériá hodnotenia priebežne osvojovali všetci tí, ktorých výkony sa podľa týchto kritérií budú hodnotiť. (Výskumom bolo napr. zistené, že kritérium plynulosti v ústnom prejave chápali známkujúci až tromi rozličnými spôsobmi, čím sa znížila reliabilita hodnotenia.)

1.6 Hodnotiace normy

Výsledok respondenta v teste vyjadrený počtom bodov, t. j. hrubé skóre, samo osebe neposkytuje dostatočnú informáciu o úrovni výkonu respondenta. Aby sme mohli posúdiť úroveň výkonu individuálneho respondenta, porovnáваме jeho výsledok s normou.

1.6.1 Sociálna norma

V situácii, keď sú štandardizované testy zriedkavosťou, je sociálna norma najčastejšie používanou normou. Podstatou sociálnej normy je porovnanie výsledku jedného respondenta s výsledkami ostatných respondentov, ktorí sú členmi danej skupiny a robili test. Učiteľ obyčajne porovnáva výkon konkrétneho žiaka s výkonom najlepších a najslabších žiakov v triede. Normu si vytvára postupne poznávaním výkonov žiakov, ktoré sú sústavne lepšie alebo slabšie. Ako sociálnu normu môže učiteľ v lepšom prípade používať uvedené miery centrálnej tendencie – priemer, modus a medián. Keďže sa sociálna norma odvodzuje od výsledkov určitej skupiny, bude v skupinách s rozdielnou úrovňou schopností rôzna. Sociálnu normu ovplyvňuje v oboch smeroch aj náročnosť učiteľa a jeho osobná intuitívna predstava o optimálnom výkone. Učiteľ sám určí, či kritériom zvládnutia testu je 40, 60 alebo 80 percent správne vyriešených úloh. Žiak, ktorý má u veľmi náročného učiteľa z testov sústavne dvojku, môže byť po prestupe k málo náročnému učiteľovi hodnotený známkou výborne. Používanie sociálnej normy je vynútené tým, že neexistujú špecifické hodnotiace stupnice na hodnotenie výkonov jednotlivých predmetoch ani vzdelávacie štandardy.

1.6.2 Populačná norma – štandard

Hodnotenie podľa populačnej normy je možné pri štandardizovaných testoch. Štandardizované testy sú zostavované na základe dôkladnej špecifikácie a sú najprv použité v reprezentatívnej vzorke populácie, pre ktorú je test určený. Výsledky získané v reprezentatívnej vzorke sú štatisticky spracované. Zisťuje sa variabilita výsledkov, frekvenčné zastúpenie výsledkov, rozdelenie výsledkov, priemer a smerodajná odchýlka od priemeru. Výsledky celých skupín a jednotlivcov sa potom hodnotia vzhľadom na parametre získané v celej populácii respondentov. Tieto parametre predstavujú normu alebo štandard.

Na základe súčasného trendu zvyšovania profesionality a odbornosti skúšania možno očakávať vznik štandardizovanej maturitnej skúšky, čo bude znamenať zvýšenie reliability hodnotenia v rámci celého Slovenska a možnosť použitia výsledkov maturitných skúšok pri prijímaní na vysoké školy. Ak bude obsah maturitných skúšok dobre reprezentovať vzdelávacie ciele, bude zabezpečený aj pozitívny spätný vplyv maturitných skúšok na proces vyučovania a samoštúdia žiakov.

Populačným štandardom posudzovania výkonov žiakov budú aj cieľové vzdelávacie štandardy, ktoré podrobne špecifikujú, aké kompetencie (vedomosti, zručnosti, schopnosti,

postoje) má v jednotlivých predmetoch dosiahnuť väčšina populácie žiakov na konci určitej etapy vzdelávania, t. j. napríklad po skončení niektorého z ročníkov alebo 1. a 2. stupňa základnej školy pri relatívne rovnakých podmienkach vzdelávania (počty hodín, počty žiakov, kvalifikácia učiteľov, jednotné ciele, rovnaké učebnice).

2 Testy

2.1 Vlastnosti kvalitných testov

Všetky testy a skúšky by mali byť pripravené a realizované – administrované na primeranej odbornej úrovni, čo znamená, že by mali byť **objektívne, validné a reliabilné**. Objektivita spočíva v tom, že testy by mali merať schopnosti čo najpresnejšie a hodnotiace kritériá by mali byť také, aby umožnili rovnaké hodnotenie toho istého výkonu rôznymi hodnotiacimi. Požiadavky validity a reliability by mali spĺňať najmä tie testy, výsledky ktorých sú pre respondentov a iných zainteresovaných závažné alebo vedú k dôležitým rozhodnutiam.

Testy by ďalej mali spĺňať požiadavku **praktickosti**, to znamená, že ich zadávanie a vyhodnotenie by mala byť jednoduché, v daných podmienkach časovo a materiálne zvládnuteľné, inými slovami, test by mal byť z hľadiska času (príprava, písanie, oprava) a výdavkov (materiál, rozmnožovanie, ai.) čo najekonomickejší.

V poslednej dobe sa pripisuje významná úloha aj **spätnému vplyvu** testov na proces učenia a učenia sa, tzv. *washback/backwash*. Žiaci a študenti sa v učení sústredia na to, čo sa bude skúšať. Ak vedia, že budú hodnotení len na základe gramatických testov, sústredia sa na gramatiku, hoci oficiálnym cieľom vyučovania jazyka je schopnosť komunikovať. Niekedy sa komunikácia neskúša preto, lebo jej skúšanie je časovo náročné. Skúša sa to, čo sa dá jednoducho, rýchlo a objektívne hodnotiť, nie to, čo je najdôležitejšie. Potom možno hovoriť o negatívnom spätnom vplyve skúšania na učenie.

2.1.1 Validita testu – platnosť testu (*test validity*)

Test je platnou skúškou stanovených kompetencií vtedy, ak meria to, čoho je deklarovaným meradlom. Ten, kto test zostavuje, musí vedieť, čo chce daným testom merať, aké vedomosti a schopnosti, t. j., ktoré kompetencie a subkompetencie. Potom sa validita zabezpečuje v prvom rade tým, že merané kompetencie – cieľové vlastnosti sú nielen jasne definované ale aj operacionalizované. Operacionalizovať kompetenciu znamená, definovať – presne opísať činnosti, ktoré dokáže urobiť kandidát, ktorý cieľovú kompetenciu má. Napr. operacionalizácia schopnosti čítania s porozumením v cudzom jazyku môže byť nasledovná: kandidát dokáže identifikovať v danom texte podstatné body, nájde požadované informácie, vyberie z daných možností názov, ktorý najlepšie vystihuje obsah textu, odhadne význam

neznámych slov z kontextu, vie odpovedať na otázky vychádzajúce z textu, a pod. Operacionalizácia ovládania vybraných slov po r môže znieť napr. vybrané slová vie vymenovať a správne ich napíše v diktáte a v slohových prácach, identifikuje chybné napísané slová; naučené pravidlá vie použiť v danej triede úloh. Zvládnutie testových úloh teda musí od kandidáta vyžadovať použitie tých schopností, ktoré chceme merať.

2.1.1.1 Obsahová validita (*content validity*)

Obsahovú validitu vedia posúdiť kompetentné osoby porovnaním obsahu testu – testových úloh s kompetenciami, ktoré má test merať. Kompetencie sú ako usudzované konštrukty definované. Merané vedomosti a schopnosti - kompetencie sú stanovené v špecifikácii testu. Test by mal zahŕňať reprezentatívne vzorky úloh, ktoré sú prejavom meraných vedomostí, schopností, či vlastností. Najčastejšie ide o kompetencie obsiahnuté v osnovách, dlhodobých plánoch a cieľových štandardoch. Reprezentatívnosť znamená, že všetky merané kompetencie sú v teste zastúpené primeraným množstvom úloh a zároveň sú v teste zastúpené všetky dôležité oblasti, ktoré chceme overiť. Napr. test, ktorého obsahom sú len gramatické úlohy rôznych typov môže byť validným testom schopnosti použitia gramatických vedomostí, nie však validným testom schopnosti použiť gramatické vedomosti v komunikácii. Skúška schopnosti kandidáta foneticky správne prepísať počuté slová v cudzom jazyku nie je validnou skúškou jeho správnej výslovnosti, pretože fonetický prepis odráža schopnosť zvuky správne identifikovať ale nie vyslovovať, čiže produkovať.

Ak ide o test pokroku v učení, testové úlohy by mali reprezentovať obsah a ciele učenia za určité obdobie. Skúšajúcemu alebo tvorcovi testu musí byť jasný účel testu, musí vedieť, čo chce overovať, a aké úlohy tomuto cieľu zodpovedajú.

2.1.1.2 Kritériová validita (*criterion – related validity*)

Validitu testu možno posúdiť aj tak, že výsledky v teste porovnáme s výsledkami v inom teste, ktorý meria rovnaké kompetencie. Iný test, s ktorým výsledky nášho testu porovnáme bude teda kritériom validity nášho testu. Výsledky by sa mali zhodovať resp. pozitívne korelovať. Na posúdenie validity môžeme použiť aj externé kritérium. Napríklad výsledky testu ústneho prejavu v cudzom jazyku porovnáme s hodnotením učiteľa, ktorý dobre

pozná ústny prejav žiaka z dlhodobých pozorovaní. Kritériom môžu byť aj výsledky v rozsiahlejšej verzii toho istého testu, ktoré porovnáme s výsledkami kratšej verzie. Ak je medzi nimi vysoká miera zhody, môžeme aj kratšiu verziu považovať za validnú.

2.1.1.3 Prediktívna validita (*predictive validity*)

Výsledky písomnej časti maturitnej skúšky v cudzom jazyku môžu mať vysokú prediktívnu validitu, ak sa ukáže, že kandidáti s dobrými výsledkami rovnako dobre uspejú aj na prijímacích skúškach z toho istého jazyka na vysokú školu. Ak sa takáto súvislosť potvrdí, tak potom bude možné na základe výsledku v maturitnej skúšky predpovedať aj výsledok prijímacej skúšky. Predikčná validita testu je schopnosť testu byť dobrým prediktorom budúceho výkonu v určitej oblasti.

2.1.1.4 Validita dojmu (*face validity*)

Test sa už na prvý pohľad, zbežné prezretie, môže javiť ako validný alebo nevalidný vzhľadom na to, čo má merať. Validitu dojmu môžu posúdiť samotní kandidáti, alebo kompetentní kolegovia tvorcu testu, ktorí môžu vyjadriť svoj názor na to, čo sa podľa nich testom meria, či meria to, čo merať má, aké sú jeho dobré stránky a nedostatky už na prvý pohľad. Takéto úsudky môžu byť povrchné, lebo sa nezakladajú na dôkladnej analýze testu, ale môžu byť aj vecné, správne a poučné. Validita zjavu nie je vedeckým pojmom, no jej posúdenie môže byť pri priebežných testoch menšieho rozsahu žiaduce a veľmi užitočné. Napríklad napísanie eseje v cudzom jazyku nie je validnou úlohou na posúdenie plynulého ústneho vyjadrovania už na prvý pohľad, úloha nájsť a opraviť chyby v písanom texte nie je validnou skúškou ovládania písomného prejavu.

2.1.2 Reliabilita testu – Spôľahlivosť testu (*test reliability*)

Výsledky testov by mali byť spoľahlivé, to znamená, že vplyv rôznych nežiadúcich faktorov by mal byť čo najmenší. Výsledky testov by mali čo najpresnejšie odzrkadľovať úroveň overovaných kompetencií. Ak by bol rovnaký test robili tí istí respondenti dvakrát,

výsledky by mali byť veľmi podobné, i keď nie úplne rovnaké. Zhoda vo výsledkoch je daná tým, že test meria pri oboch administráciách rovnaké kompetencie. Rozdiely vo výkonoch sú spôsobené rôznymi náhodnými faktormi, ktoré nie je možné eliminovať. Napr. únava kandidáta po dlhej ceste, zlá nálada, bolesti hlavy, nepríjemný zážitok bezprostredne pred testom, a pod.. Snahou tvorcov testu by malo byť, aby rozdiely vo výsledkoch testov odrážali rozdiely v úrovni meraných kompetencií, a aby výsledky neboli skreslené takými formálnymi a obsahovými nedostatkami testu, ktorým sa možno vyhnúť. K takýmto nedostatkom patrí: nedostatočné množstvo času na splnenie úloh, poskytnutie dlhšieho času niektorým respondentom, nejednoznačné inštrukcie, neprehľadná forma testu, možnosť uhádnutia odpovedí, administrácia testu v rôznych skupinách v rôznom čase, možnosť odpisovania, pomoc dozoru niektorým respondentom, nízky počet položiek testu overujúcich závažnú kompetenciu, nevhodné radenie úloh, neznáme slová v inštrukciách, neprítomnosť ilustračných príkladov úloh, a pod.. Spoľahlivosť subjektívnych testov môže byť ohrozená vtedy, ak testy hodnotia viacerí hodnotiaci, ktorí nechápu hodnotiace kritériá úplne rovnako. Toto sa môže stať vtedy, ak sú požadované kvality opísané nejednoznačne a nepresne. Hodnotiaci členovia skupiny by mali niekoľko úloh vyhodnotiť spoločne a pritom si ujasniť sporné formulácie kritérií.

Spoľahlivosť závažného testu možno dopredu overiť pilotným použitím testu na vzorke jedincov s rovnakými charakteristikami, aké má cieľová populácia, pre ktorú sa test pripravuje. Jednou z možností je metóda test - retest, kedy kandidáti robia ten istý test dvakrát v určitom časovom odstupe a porovnávajú sa výsledky. Časový odstup musí byť primeraný, aby nedošlo k výraznému zabúdaniu, a zároveň dostatočne dlhý, aby si kandidáti nemohli zapamätať odpovede mechanicky.

Ďalšou možnosťou je porovnanie výsledkov dvoch rovnocenných foriem testu. Príprava dvoch rovnocenných foriem je veľmi náročná, pretože môže byť ťažké nájsť zodpovedajúci počet úloh, respektíve úlohy úplne rovnakej obtiažnosti.

Ako relatívne najjednoduchší sa javí spôsob porovnania výsledkov polovic toho istého testu. Jednou polovicou budú párne a druhou polovicou nepárne položky v teste. To však znamená, že dve za sebou nasledujúce položky by mali byť rovnocenné, t. j. mali by skúšať to isté a byť rovnako náročné.

Reliabilita testu je vyjadrená koeficientom reliability, ktorého najvyššia, ideálna hodnota je 1. Čím bližšie je koeficient reliability k hodnote 1, tým vyššia je reliabilita testu.

Na výpočet *reliability* sa používa nasledovný vzorec:

$$r = 1 - \frac{\overline{X}(n - \overline{X})}{n\sigma^2},$$

kde \overline{X} znamená aritmetický priemer, n je počet položiek v teste a σ je smerodajná odchýlka.

Čím je výsledok bližšie k 1.00, tým je test spoľahlivejší.

Pri štandardizovaných testoch sa koeficient *reliability* ako aj spôsob validizácie testu uvádza v testovej príručke – manuáli testu. V dnešnej dobe existujú rôzne štatistické programy, ktoré po zadaní dát vypočítajú všetky potrebné ukazovatele testu vrátane koeficientu *reliability*.

2.2 Špecifikácia testu – skúšky

Na základe výsledkov niektorých testov a skúšok sa robia dôležité rozhodnutia týkajúce sa respondentov. Môže ísť napr. o určenie výslednej známky, udelenie kvalifikácie, doporučenia na štúdium či získanie zamestnania alebo štipendia. Testy a skúšky, ktoré môžu závažným spôsobom ovplyvniť životy respondentov, by mali byť pripravené kvalitne a zodpovedne.

Prvým krokom k vytvoreniu kvalitnej skúšky alebo testu je vypracovanie ich špecifikácie. Špecifikácia testu obsahuje všetky dôležité charakteristiky testu, z ktorých sa pri tvorbe testu vychádza. Tieto charakteristiky navzájom spolu súvisia a podmieňujú sa.

Špecifikácia testu slúži okrem tvorcov testu aj tým, ktorí respondentov na test pripravujú, ako aj samotným respondentom.

Špecifikácia by mala obsahovať nasledujúce položky:

- 1) **Aký je účel testu.** Z účelu testu vyplynie, o aký typ testu pôjde, či pôjde o test overujúci alebo rozlišujúci a aký bude náročný. Účel testu určí potrebný rozsah úloh v teste, typy úloh v teste a spôsob spracovania výsledkov. Z účelu testu je jasné, kto a ako využije výsledky testu a v akej forme majú tieto byť vyjadrené. Účelom testu môže byť napr. stanovenie úrovne praktického ovládania cudzieho jazyka s vydaním úradného osvedčenia.
- 2) **Komu bude test určený,** aká je predpokladaná úroveň kompetencií respondentov.
- 3) **Čo má test merať:** ak meria vedomosti, ktoré oblasti majú byť zahrnuté, ak meria zručnosti a schopnosti, budú uvedené ich operačné definície.

- 4) Pri testoch cudzích jazykov, ktoré merajú schopnosť komunikácie, budú uvedené **komunikatívne funkcie – operácie**, situačný kontext a tematické okruhy v rámci ktorých sa operácie môžu realizovať, typy interakcií s adresátmi. Pri skúšaní čítania a počúvania typy a rozsah textov, ktoré sa použijú.
- 5) **Aký formát bude test mať**. Koľko bude mať častí, čo budú jednotlivé časti overovať a aký budú mať rozsah, koľko času bude ktorá časť trvať, akú váhu budú mať časti testu v celkovom hodnotení.
- 6) **Koľko a aké úlohy-položky budú obsahovať časti testu**, ako budú úlohy skórované (kritériá hodnotenia a kľúč), aké budú inštrukcie, či budú uvedené príklady položiek so správnym riešením, či sa budú odpovede zapisovať do testov alebo do odpovedových hárkov.

Okrem uvedených charakteristík môže špecifikácia testu zahŕňať aj technicko-administratívne údaje, napr. koľko testov bude treba pripraviť, informácie o tom, kto a dokedy bude testy opravovať, kto štatisticky spracuje výsledky.

2.3 Postup prípravy testu

Dôležitý test, ktorý bude hromadne použitý, vždy zostavuje tím zostavovateľov. Môžu to byť skúsení učitelia, metodici a iní kompetentní odborníci, ktorí sa zaoberajú skúšaním v teórii i praxi.

Tvorcovia testov sa najprv dôkladne oboznámia so špecifikáciou testu. Ďalším krokom je vytvorenie banky úloh-položiek podľa požiadaviek špecifikácie. Nasleduje výber potrebného počtu úloh a ich posúdenie podľa určitých kritérií. Každú položku posudzuje celá skupina, keď sú výhrady voči položke uznané, položka je upravená alebo vyradená (*vetting*). Potom sa môže pristúpiť k zostaveniu testu vrátane inštrukcií a príkladov, ku grafickej úprave testu a k jeho vytlačeniu.

Hotový test sa **predskúša** (pilotuje) na reprezentatívnej vzorke respondentov, ktorá je v relevantných charakteristikách (úroveň schopností, vek, absolvovaná príprava) podobná budúcim respondentom. Predtestovaním sa identifikujú nedostatky v teste ako napr. nejasné inštrukcie-zadania, veľmi ľahké, resp. ťažké položky, primeranosť času na splnenie úloh, atď. Na základe analýzy priebehu administrácie a opravy testov budú vykonané potrebné zmeny

C) Úlohy na dopĺňanie textu (multiple gap-filling)

A: Excuse me, can you tell me ____1____ the station?

B: Yes, certainly. ____2____ and then go ____3____ the church.

- | | | |
|----------------------|------------------------|-----------|
| 1. A the way to | 2. A Turn to the right | 3. A past |
| B how I must go to | B Turn right | B along |
| C what is the way to | C Turn the right | C at |

D) Úlohy na priradovanie (matching)

Príklad: Ku každému slovu v prvom stĺpci priradte podstatné meno z druhého stĺpca.

- | | | |
|--------------|-------|------------|
| 1 to hijack | _____ | a/ a child |
| 2 to kidnap | _____ | b/ a woman |
| 3 to hold up | _____ | c/ a plane |
| 4 to raid | _____ | d/ a coach |
| 5 to rape | _____ | e/ a bank |

E) Úlohy na usporiadanie nesprávne zoradených slov, z ktorých treba utvoriť vetu alebo odsek (reordering)

Dané slová napíšte v takom poradí, aby tvorili zmysluplnú vetu.

at / poor / look / that / woman / old _____

F) Identifikácia chýb vo vete alebo texte (error recognition)

The famous actress told the story of her life in two book.

A B C D

Výhodou zatvorených úloh je rýchla oprava, objektívne hodnotenie a možnosť zaradenia dostatočného počtu úloh, pretože respondenti sa nezdržiavajú dlhým písaním. Nevýhodou je možnosť uhádnutia správnej odpovede ako aj skutočnosť, že ide len o reprodukciu vedomostí. Navyše je pri tomto type úloh obmedzený rozsah skúšaných kompetencií, čo má nepriaznivý dopad na vyučovanie.

Úlohy produktívne – otvorené úlohy (*production tasks, candidate supplied response items*)

Tieto úlohy vyžadujú od respondentov stručné, ale aj dlhšie odpovede.

A) Úlohy s krátkou odpoveďou (*short response items*)

V ktorom roku bola založená Oxfordská univerzita? _____

Kde sa narodil Shakespeare? _____

Otázky k textu: What did Mary miss on her holiday? _____

What did Mary appreciate most on the local people? _____

B) Doplňanie chýbajúcich slov vo vetách (napr. po počúvaní textu) alebo v súvislých textoch.

1. *Voľná odpoveď* (*free response*)

I hope you _____ if I smoke.

2. *Výber z možností* (*choice from options*)

Doplňte „a“ alebo „the“

_____ dinosaur is extinct

3. *Doplňanie správneho tvaru daného slovesa* (*supply correct form of specified verb*)

He _____ (*sing*) at yesterday's concert.

4. *Doplňanie vynechaných vybraných položiek* (*selected deletions*)

_____ the United States, 30 million people have successfully kicked the habit _____ 1964,
and only one _____ three Americans now smoke....

5. *Úlohy typu „cloze“, t.j. doplňanie systematicky vynechovaných položiek* (*systematic deletion*)

People in developing countries are, it ___1___, being persuaded by advertising that
affluent ___2___ normally smoke, through non-smokers outnumber smokers ___3___
Britain, the United States, Sweden and other ___4___ countries.

C) Transformácia štruktúr alebo vyjadrenie rovnakej funkcie iným spôsobom. (*sentence transformation*)

I don't think it's necessary for you to stay any longer.

I don't think you _____

Whose watch is this?

BELONG

D) Kombinácia viet (sentence combination)

Ann did her homework. Then she went swimming.

After _____

E) Prepisovanie (rewriting)

Prepíšte vetu v nepriamej reči:

“I think Tom is at home,” she said. _____

F) Slovotvorné úlohy (word-formation)

Podčiarknuté slová dajte do správneho tvaru.

Animal species share our urge to survive and remain free from injure.

They do not receive this benefit. They are bred and captured, sold, forced to endure harm procedures and slaughtered.

G) Oprava chýb vo vete alebo texte (error correction)

Is that man which lives next door to you?

H) Produkcia ústnych a písomných prejavov s rôznou mierou kontroly a usmernenia.
(open-ended and guided oral and writing tasks)

Napr. písanie krátkych listov, odkazov na základe poskytnutých stimulov, písanie esejí, riadené a voľné rozhovory, hranie rolí, diskusie a riešenie problémov (*problem solving, coming to consensus*).

Otvorené úlohy typu A-G je možné objektívne skórovať. Úlohy typu H sa hodnotia podľa dopredu stanovených špecifických kritérií, pričom je žiadúce kombinovať analytické hodnotenie s globálnym hodnotením. Kvalitu výkonov možno určovať podľa analytických a globálnych hodnotiacich škál. Jeden výkon môže hodnotiť viac posudzovateľov.

Výhodou otvorených úloh je možnosť skúšania kompetencií vyššieho rádu, tvorba odpovedí vyžaduje hlbšie a presnejšie osvojené poznatky, otvorené úlohy sú autentickejšie a ľahšie sa tvoria, možnosť uhádnutia odpovede je minimálna. Aj pri hodnotení otvorených úloh je možné dosiahnuť vysokú mieru objektivity hodnotenia. Keďže týmito úlohami sa skúša aj schopnosť praktického použitia jazyka v situáciách imitujúcich reálny život, majú pozitívny dopad na vyučovanie.

Nevýhodou otvorených úloh je časová náročnosť opravovania a hodnotenia, vyššie nároky na odbornosť hodnotiacich, náročnosť vytvárania takých položiek, aby bolo respondentovi jednoznačne jasné, aká odpoveď sa vyžaduje, nutnosť rozhodovania o tom, nakoľko brať do úvahy formálne chyby, keď nie sú v danej úlohe podstatné.

Produktívne otvorené úlohy môžu vyžadovať dlhšie súvislé verbálne odpovede ako aj neverbálnu produkciu. Takéto úlohy vyžadujú uplatnenie rádovo vyšších schopností, ako sú: aplikácia vedomostí v súlade s určeným cieľom, rôzne zručnosti, logické, tvorivé a hodnotiace myslenie. K úlohám tohto typu patria napr. eseje, rozbor umeleckých diel, riešenie modelových situácií a problémov, tvorivé úlohy.

Otvorené produktívne úlohy v skúšaní cudzích jazykov sú napr. rolové aktivity, diskusie za účelom dosiahnutia zhody alebo vyriešenia úlohy (problému), voľná konverzácia. Na tomto mieste si dovoľíme upozorniť, že vymýšľanie príbehov na základe stimulu vo forme obrázku je úloha produktívna, ale vyžaduje nielen jazykové kompetencie, ale aj tvorivosť, nie je teda len skúškou jazykovej schopnosti a znevýhodňuje menej tvorivých, hoci jazykovo zdatných jedincov. Navyše ide o úlohu neautentickú.

3 Analýza a interpretácia výsledkov testu

Príprava kvalitného testu vyžaduje množstvo času a námahy, preto je prirodzené, že z výsledkov testu sa snažíme získať čo najviac informácií. Analýza výsledkov nám umožní nielen ich interpretáciu, ale naznačí aj možnosť opakovaného použitia testu v budúcnosti.

Základné informácie, ktoré možno získať prostredníctvom analýzy výsledkov testu, sú:

- *aká bola úroveň obtiažnosti testu* (či bol ťažký, ľahký alebo primeraný pre respondentov, ktorí ho absolvovali);
- *aké náročné boli pre respondentov jednotlivé položky testu;*
- *ktoré položky rozlíšili úspešných a menej úspešných respondentov;*
- *či boli v skupine respondentov veľké rozdiely vo výkonoch*, t. j. či sa skupina javí ako homogénna alebo heterogénna;
- *ako obstáli v teste rôzne skupiny respondentov.*

3.1 Miery centrálnej tendencie

Ukazovateľmi náročnosti testu sú miery centrálnej tendencie, ktorými sú **priemer**, **modus** a **medián**. Priemerné skóre, t. j. **aritmetický priemer** (*mean*) skóre vypočítame, keď spočítame všetky skóre – jednotlivé výsledky a ich súčet vydáme celkovým počtom skóre. Vzorec na výpočet priemeru je

$$\bar{X} = \frac{\sum X}{N}$$

kde \bar{X} je priemer, x je skóre, N je počet výsledkov – skóre, \sum je suma alebo spočítať.

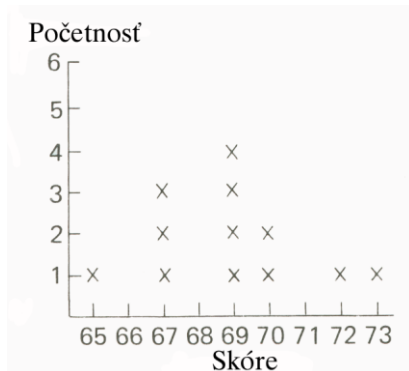
Aritmetický priemer je mierou centrálnej tendencie a umožňuje porovnať výsledok jednotlivca s priemerom, ktorý dosiahla skupina ako celok alebo porovnať výsledky viacerých skupín, ktoré robili rovnaký test. Čím je priemer vyšší vzhľadom na maximálny možný počet bodov, tým bol test pre skupinu ľahší, a naopak, čím je priemer nižší, tým bol test pre skupinu ťažší.

Výsledok, ktorý dosiahlo najviac respondentov sa nazýva **modus** (*mode*). Ak sú výsledky napr. 55, 56, 57, 57, 57, 58, 59, 60, 60, 61, 62, tak modus je 57, pretože tento výsledok dosiahlo najviac respondentov. V skupine výsledkov sa môže vyskytovať aj viac modusov, vtedy hovoríme o bimodálnom alebo trimodálnom rozdelení výsledkov.

Ďalšou mierou centrálnej tendencie je **medián** (*median*). Je definovaný ako stredný bod v rozdelení výsledkov alebo bod, pod ktorým je 50 % výsledkov a nad ktorým je tiež 50 % výsledkov. Medián pre nasledovné výsledky 10, 25, 30, 43, 57, 65, 68 je **43**, rovnaký počet skóre je vyšší aj nižší ako medián. Ak však máme párny počet výsledkov, medián bude medzi dvomi strednými výsledkami, napr. pre skóre 25, 27, 32, 37, 41, 42, 43 je medián **34.5**. Ak sú všetky miery centrálnej tendencie v jednej skupine respondentov vyššie ako v druhej skupine, môžeme povedať, že prvá skupina obstála v teste lepšie ako druhá skupina.

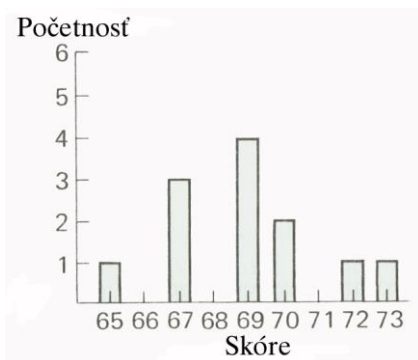
3.2 Miery rozptylu

Názorne a prehľadne môžeme výsledky testov prezentovať zostrojením *histogramu* alebo ešte lepšie *stĺpcového diagramu*. Dosiahnuté hodnoty skóre zoradíme podľa veľkosti za sebou od najnižšieho po najvyššie, zároveň zistíme a poznačíme si, koľkokrát sa jednotlivé výsledky opakujú, t. j., aké sú ich početnosti. Takto pripravené výsledky testu potom môžeme graficky znázorniť tak, že na vodorovnú os **x** nanesieme všetky vyskytujúce sa hodnoty výsledkov od najnižšieho po najvyšší a na vertikálnu os **y** ich početnosti. Dostaneme graf, ktorý sa volá histogram, vid' obrázok 1.



Obrázok 1 *Histogram rozdelenia skóre (Brown, 1993)*

Podiely jednotlivých výsledkov v skupine ako celku lepšie znázorňuje stĺpcový diagram, kedy podiely početností vzhľadom na celkový počet znázorníme zodpovedajúcimi plochami. Vid' obrázok 2.



Obrázok 2 Stĺpcový diagram rozdelenia skóre (Brown, 1993)

Pri analýze výsledkov testu si nepochybne všimneme aj rozdiel medzi najlepším–najvyšším a najhorším–najnižším dosiahnutým výsledkom. Keď k tomuto rozdielu pripočítame hodnotu jedna, dostaneme rozpätie – **variačnú šírku** (*range*) rozptylu výsledkov. Ak bol najlepší výsledok 90 bodov a najslabší výsledok 40 bodov, tak je rozpätie rozptylu 51. Variačná šírka je ovplyvnená najvyššou a najnižšou hodnotou dosiahnutých výsledkov. Čím väčší je rozdiel medzi nimi, tým väčší je **rozptyl výsledkov** (*spread of scores*). Lepšou mierou posudzovania rozptylu výsledkov od priemeru je **smerodajná odchýlka** (*standard deviation*). Nižšie hodnoty smerodajnej odchýlky znamenajú väčšiu homogenitu skupiny v meranej vlastnosti, čo sa prejaví sústredením výsledkov bližšie k priemeru, vyššie hodnoty naznačujú vyššiu heterogénnosť skupiny, ktorá sa prejaví tým, že výsledky budú rozptýlené ďalej od priemeru.

Smerodajná odchýlka sa označuje symbolom sigma a vypočítame ju podľa vzorca

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N}}$$

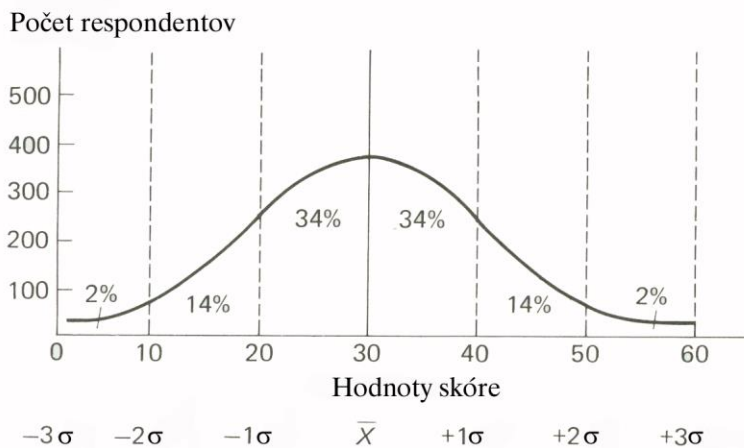
Na výpočet smerodajnej odchýlky si pripravíme údaje podľa tabuľky 1.

Študenti	Skóre X	Priemer \bar{X}	Rozdiel $(X - \bar{X})$	2. mocnina rozdielu $(X - \bar{X})^2$
Ivan	45	63	- 18	321
Juraj	64	63	1	1
Martin	62	63	- 1	1
Igor	40	63	- 23	529
Vlasta	80	63	17	289
Ingrid	76	63	13	169
Jana	70	63	7	49
Karol	67	63	4	16
—		$\sum (X - \bar{X})^2$	=	1375

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = \sqrt{\frac{1375}{8}} = \sqrt{171,875} = 13,11$$

Tabuľka 1 Výpočet smerodajnej odchýlky, hypotetický príklad

Pri normálnom rozdelení výsledkov v základných súboroch respondentov (napr. všetci žiaci 5. ročníka ZŠ na Slovensku alebo všetci maturanti gymnázií, resp. ich reprezentatívne početné zastúpenia) sa nachádza 68 % výsledkov v rozpätí plus – mínus jedna sigma a až 96 % výsledkov v rozpätí plus – mínus 2 sigmy. Len o niečo viac ako 2 % výsledkov spadá medzi plus 2 až 3 sigmy a rovnaký počet výsledkov spadá medzi mínus 2 až 3 sigmy. Vid' obrázok 3.



Hodnoty skóre: $\bar{X} = 30$; $\sigma = 10$

Obrázok 3 Normálne rozdelenie výsledkov v hypotetickom základnom súbore (Brown, 1993)

Smerodajná odchýlka je veľmi dobrou mierou posúdenia celkovej úrovne vedomostí alebo schopností, ktoré daný test meria, ako aj mierou porovnania úrovne niekoľkých veľkých skupín, napr. porovnanie výsledkov získaných v určitej škole s výsledkami získanými v rámci celého Slovenska alebo porovnanie s inými vybranými školami, ktorých výsledky nás z nejakých dôvodov zaujímajú.

3.3 Analýza testových položiek

Podstatnými údajmi, ktoré môžeme získať o jednotlivých položkách v teste, je informácia o úrovni ich obtiažnosti ako aj miera schopnosti oddeliť lepších (úspešnejších) respondentov od slabších (menej úspešných). Každú položku v teste môžeme teda skúmať z hľadiska jej obtiažnosti a diskriminačnej sily. Tieto charakteristiky môžeme zistiť pomocou **položkovej analýzy** (*item analysis*) buď počítačom (program ITEMAN) alebo vypočítat manuálne.

O obtiažnosti jednotlivých položiek vypovedá **index náročnosti** (*facility index*), ktorý je daný percentom študentov, ktorí správne odpovedali na otázku, resp. vyriešili úlohu. Ideálnou hodnotou náročnosti položiek v teste je rozmedzie od 40 % do 60 %, hoci v mnohých testoch sú tiež prijateľné hodnoty medzi 30 až 70 %. Čím je index bližšie k 100 %, tým je položka ľahšia.

Diskriminačný index (*discrimination index*) je ukazovateľ, ktorý poukazuje na mieru schopnosti položiek rozlíšiť úspešných a menej úspešných respondentov. Tento index oddeľuje schopnejších (úspešnejších) respondentov od menej schopných (úspešných). Ak študenti, ktorí dosiahli vysoké celkové skóre, úspešne zodpovedali danú otázku a menej úspešní študenti ju nezodpovedali, vtedy môžeme povedať, že otázka (položka) je dobrá, lebo odlišuje „dobrých“ študentov od „slabých“ takým istým spôsobom ako celkové výsledky testu. Diskriminačný index vypočítame tak, že zistíme počet správnych odpovedí na danú položku u jednej tretiny respondentov, ktorí mali najvyššie skóre (*top group*). Od tohto počtu odpočítame počet správnych odpovedí získaných na danú položku v „dolnej skupine“ (*bottom group*), ktorú tvorí jedna tretina celkového počtu testantov, ktorí získali najnižšie skóre. Rozdiel vydelíme tretinou celkového počtu respondentov.

$$D = \frac{\text{Počet správnych odpovedí hornej tretiny} - \text{počet správnych odpovedí dolnej tretiny}}{1/3 N}$$

Príklad v skupine 48 respondentov: $D = \frac{12 \text{ správnych} - 8 \text{ správnych}}{16} = 0.25$

Pri efektívnom rozlišovaní (diskriminácii) by mal diskriminačný index dosiahnuť aspoň hodnotu 0.3, ale vyššie hodnoty sú lepšie. Ak je hodnota veľmi nízka, to znamená, že obe skupiny skúšaných, horná aj dolná (lepší a slabší študenti), správne, resp. nesprávne odpovedali na tú istú otázku, t. j. niet rozdielu v odpovediach dobrých a slabších študentov.

Príklad interpretácie výsledkov položkovej analýzy:

Aké závery môžeme vyvodit' z nasledovných údajov?

IN – index náročnosti; DI – diskriminačný index

Položka	IN	DI	Interpretácia
1.	0.95	0.08	veľmi ľahká, slabá diskriminácia
2.	0.23	0.91	dobrá diskriminácia, ľahká pre hornú tretinu?
3.	0.60	0.62	priemerná položka
4.	0.08	-0.02	príliš ťažká pre všetkých, odpovede uhádnuté?
5.	0.75	0.68	dost' ľahká pre väčšinu, ale nie pre dolnú tretinu

4 Anglicko – slovenský terminologický slovník

Vzhľadom na skutočnosť, že väčšina dostupnej literatúry v oblasti testovania jazykov je písaná v anglickom jazyku, tento slovník má pomôcť učiteľom pri čítaní literatúry zaoberajúcej sa touto problematikou písanej v angličtine.

Slovník obsahuje približne 150 hesiel a pokrýva oblasť skúšania, hodnotenia a testovania jazyka, analýzy a štatistického vyhodnotenia testov. Nachádza sa tu predovšetkým terminológia vyššie spomínaná v texte, ako aj množstvo najbežnejších základných termínov používaných v tejto oblasti jazykovedy. Anglické výrazy sú preložené do slovenčiny a vysvetlené tak, aby boli zrozumiteľné aj učiteľom, ktorí sa s touto terminológiou ešte nestretli.

Heslá v slovníku sú zoradené podľa abecedy. Heslové záhlavie je hrubo vytlačené, za ním nasleduje slovenský preklad (ekvivalent) a vysvetlenie hesla písané kurzívou.

V poslednej časti 5.1 je zoznam hesiel usporiadaných hniezdovým spôsobom. V každej skupine (hniezde) sú heslá v abecednom poradí kvôli ľahšej orientácii a uľahčeniu jeho vyhľadania v slovníku.

achievement test – výkonový test: *test zisťujúci úroveň vedomostí a zručností obyčajne po istom období vzdelávania alebo výcviku*

administration (of a test) – administrovanie: *zadávanie testu, zabezpečenie a dohliadanie nad písaním testu*

anchor item – kotvová položka: *položka, ktorá sa nachádza v dvoch alebo viacerých testoch. Tieto položky majú známu charakteristiku a vytvárajú jednu časť novej verzie testu kvôli tomu, aby poskytli informácie o teste a o kandidátoch, ktorí ho písali.*

answer key – kľúč k odpovediam: *všetky (možné) správne riešenia a odpovede na otázky v teste s bodovým hodnotením jednotlivých položiek*

answer sheet – odpoveďový hárok: *hárok, do ktorého respondent zapisuje správne odpovede na otázky v teste, resp. riešenia úloh*

aptitude test – test schopností: *test zisťujúci potenciálnu spôsobilosť, schopnosť človeka učiť sa daný predmet, resp. disciplínu. Používa sa pred začiatkom výučby (kurzu) za účelom výberu najvhodnejších kandidátov, napr. test pamäti, schopnosti imitovať zvuky a slová, a pod.*

assessment – skúšanie a hodnotenie

assessor – (tiež **examiner, rater**) skúšajúci: *osoba, ktorá prideluje body za výkon kandidáta v teste na základe subjektívneho hodnotenia. Skúšajúci sú zvyčajne kvalifikovaní vo svojom odbore a vyžaduje sa, aby absolvovali špeciálne školenie v štandardizácii hodnotenia odpovedí*

attainment test - súhrnný, cieľový test: *test používaný na zistenie množstva naučenej látky po dlhšom časovom úseku*

authenticity – autenticnosť: *miera prirodzenosti testových úloh a materiálov (úlohy imitujúce reálne životné situácie, texty vybrané z novín a časopisov, nahrávky prirodzenej konverzácie, bežné komunikatívne situácie atď.)*

backwash-effect – spätné pôsobenie: *vplyv testu na proces vyučovania a učenia sa*

banked cloze – test typu „cloze“ (pozri cloze test) s možnosťou výberu z ponúknutých odpovedí: *typ úlohy na dopĺňanie textu, v ktorom sú vynechané celé slová, odpoveď respondent vyberá zo série daných slov*

candidate – kandidát, respondent, uchádzač: *študent, alebo osoba, ktorá sa podrobí skúšaniam a hodnoteniu*

calibration – kalibrácia: *proces určovania úrovne obtiažnosti testu, alebo testov. Kalibrácia môže zahŕňať aj zakotvenie položiek z rôznych testov pre stanovenie spoločnej úrovne obtiažnosti. Ak je test zostavený z kalibrovaných položiek, potom sú výsledky testu ukazovateľom úrovne "schopnosti" uchádzačov*

central tendency – centrálna tendencia: *štatistické miery, ktoré vyjadrujú stredné umiestnenie alebo hodnotu, typickými mierami centrálnej tendencie sú priemer, medián a modus*

cloze test – cloze test: *typ úlohy na dopĺňanie textu, v ktorom sú vynechané celé slová. V tradičnom type tohto testu je vynechané každé n-té slovo (piate, ôsme), ktoré má žiak doplniť na základe pochopenia kontextu.*

communicative task – komunikatívna úloha: *cieľom úlohy je realizovať komunikatívny cieľ*

component – komponent, zložka, súčasť: *časť skúšky, prezentovaná ako samostatne hodnotená časť s inštrukciami a časom splnenia, komponenty sú často zamerané na zručnosti a majú vlastné pomenovania ako napr. počúvanie s porozumením, kompozícia, atď.*

competence – spôsobilosť, schopnosť, kompetencia: *zahŕňa vedomosti, zručnosti, schopnosti, vlastnosti, napr. jazyková kompetencia je schopnosť jedinca vytvárať a chápať vety vrátane takých, ktoré ešte predtým nepočul*

communicative competence – komunikatívna kompetencia: *schopnosť nielen aplikovať gramatické pravidlá daného jazyka za účelom vytvárania gramaticky bezchybných viet, ale aj spôsobilosť vhodne používať tieto vety za účelom splnenia komunikatívnych potrieb v danom kontexte, komunikatívna kompetencia zahŕňa: vedomosti o gramatike a lexike, vedomosti o pravidlách používania jazyka, schopnosť používať a reagovať na rôzne rečové prejavy (ako napr. žiadosť, ospravedlnenie, pozvanie, ponuka...) a spôsobilosť vhodne používať daný jazyk*

content analysis – obsahová analýza: *spôsob popisu a analýzy obsahu testových materiálov. Táto analýza je potrebná pre potvrdenie toho, že obsah testu sa zhoduje so špecifikáciou testu, obsahová analýza je nevyhnutná pre zabezpečenie obsahovej a konštruktivej validity*

contextualised item – kontextualizovaná položka: *položka, ktorú testujeme, nie je izolovaná, vyňatá z kontextu, ale vsadená do väčšieho súvislého celku, ktorý napomáha k odhadnutiu*

funkcie, ktorú položka v kontexte plní v interakcii s ostatnými faktormi, integrované položky si vyžadujú kontext, aby boli schopné plniť svoj účel, kontextualizácia je nevyhnutnou črtou výkonu

correlation – korelácia: *štatistická závislosť medzi dvoma alebo viacerými premennými, vyjadruje sa korelačným koeficientom*

correlation coefficient – korelačný koeficient: *číselná hodnota v rozmedzí od 0,00 do +1 až -1,00, ktorá vyjadruje tesnosť vzťahu (štatistickú závislosť) medzi dvoma alebo viacerými premennými, označuje sa symbolom r*

criterion-referenced testing – overovanie, skúšanie absolútneho výkonu: *výkon jednotlivca sa pri ňom porovnáva s určitým kritériom, so zvládnutím učiva, cieľom, externou normou a nie s výkonom iných jednotlivcov*

critical values – kritické hodnoty; *hodnoty, ktoré limitujú kritickú oblasť do rámca ktorej sa musia zmestiť štatistické hodnoty v prípade, že pozorované rozdiely majú byť považované za štatisticky významné (signifikantné) na danej úrovni*

C-test – *test, v ktorom respondenti dopĺňajú chýbajúce hlásky v slovách, vždy sa dopĺňa rovnaký počet hlások alebo o jednu hlásku viac, ako je naznačených*

decontextualised (disembodied) language – jazyk bez kontextu: *nazýva sa tak jazyk v minimálnom alebo v žiadnom kontexte, týmto spôsobom sú formulované tzv. „discrete-point items“(pozri nižšie), ktoré sú od seba oddelené a navzájom nesúvisia*

descriptor – deskriptor: *stručný popis stupňa hodnotiacej škály, ktorá sumarizuje stupeň zručnosti alebo typ výkonnosti očakávaný od uchádzača na dosiahnutie daného stupňa*

development, construction (of a test) – príprava, zostavenie, resp. výroba testu

diagnostic test – diagnostický test: *test určený na zistenie nedostatkov (aj silných stránok), vychádzajúc z výsledkov testu ďalšie vyučovanie sa sústreďuje na odstránenie zistených nedostatkov*

dichotomous – dichotomický: *dvojčlenný*

direct assessment (testing) – priame hodnotenie (testovanie): *hodnotenie, ktoré uznáva výkonnosť ako adekvátny a hodnoverný ukazovateľ komunikatívnej kompetencie, z toho hľadiska sú priame testy zostavené z takých úloh, ktoré uchádzač môže očakávať v každodenných komunikatívnych situáciách a ktoré hodnotia schopnosť integrovať vedomosti o jazyku s jeho používaním pri komunikácii*

directed writing task – riadená písomná úloha: *písomná úloha písaná na základe daných inštrukcií (pozri **guided writing task**)*

discrete item – oddelená, separátne položka: *položka, ktorá je izolovaná, nezávislá od ostatných položiek, alebo akéhokoľvek doplňujúceho materiálu, príkladom takto používaného*

typu položky je „multiple-choice“ (výber z viacerých možností)

discrete-point test – jazykový test, ktorý hodnotí vedomosti iba z jedného z komponentov jazyka, napr. gramatiku alebo slovnú zásobu. Tento druh testov vychádza z teórie, že jazyk pozostáva z rôznych komponentov (napr. gramatika, slovná zásoba) a rôznych zručností (počúvanie, čítanie, písanie, rozprávanie) a tieto sú zostavené z elementov, ktoré môžu byť testované oddelene. Testy „oddelených položiek“ zvyčajne pozostávajú z množstva položiek, ktoré sa sústreďujú na konkrétnu oblasť vedomosti, a môžeme ich dať do protikladu s integrovanými testami.

discrimination – rozlišovanie obtiažnosti: stupeň, do akého test alebo položka v teste rozlišuje medzi lepšími a slabšími respondentmi, miera diskriminácie sa vyjadruje diskriminačným indexom

discrimination index – diskriminačný index: miera rozlíšenia medzi dobrými a slabšími študentmi

discursive composition – slohová úloha, polemická esej: písomná úloha, v ktorej úlohou kandidáta je diskutovať o téme, na ktorú je veľa názorov, alebo argumentovať v prospech svojho postoja

dispersion – rozptyl: miera rozptylu skóre v skupine

distribution – distribúcia, rozdelenie: vzor alebo model skóre alebo mier v skupine

double marking – dvojnásobné známkovanie: metóda hodnotenia výkonnosti, pri ktorom uchádzačov výkon hodnotia dvaja známkujúci nezávisle od seba

distractor – distraktor: ponúknutá nesprávna možnosť odpovede na testovú položku

editing – redigovanie, vydávanie: proces v ktorom sú materiály na skúšku modifikované a upravené do konečnej podoby, v ktorej sa objavia na skúške

error – chyba: použitie jazykovej položky (v prehovore alebo v písomnom prejave používateľa cudzieho jazyka) takým spôsobom, ktorý rodený užívateľ jazyka, alebo jedinec plynulo hovoriaci daným jazykom považuje za nesprávne alebo nedostatočne naučené

error-recognition – identifikácia chýb: test, v ktorom kandidát má za úlohu rozpoznať rôzne druhy chýb, ako napríklad nesprávne napísane slová, gramatické chyby vo vetách, atď.

examination, exam – skúška

examination board – skúšobná komisia

examinee – skúšaný, osoba skladajúca skúšku

examiner – skúšajúci

evaluation – evaluácia, hodnotenie, posúdenie

facility value – item difficulty – miera obťažnosti: *miera obťažnosti položky v teste, percento respondentov, ktorí správne odpovedali, čím je index vyšší, tým je položka ľahšia*

frequency distribution – rozdelenie početností: *frekvencia výskytu toho istého výsledku*

formative evaluation – formatívne hodnotenie: *hodnotenie, ktorého funkciou je zlepšiť priebeh a výsledky vyučovania a učenia sa, ide o priebežné, diagnostické, spätno-väzbové hodnotenie*

gap-filling – doplňovanie: *testová úloha, v ktorej respondent dopĺňa odpovede (písmená, čísllice, slová, vety, slovné spojenia, vety alebo celé odseky) do textu na vyznačené miesto*

grade – známka

grading – známkovanie, klasifikácia

guided writing task – usmernená písomná úloha: *úloha, ktorá vyžaduje od kandidáta produkciu písaného textu na základe danej textovej alebo grafickej informácie, osnovy (obrázok, pohľadnica, tabuľka, graf, inštrukcie, atď.), ktorá sa použije na overenie správnosti a štandardizovania predpokladanej odpovede*

histogram – grafické znázornenie frekvencie výsledkov testu

indirect assessment (testing) – nepriame hodnotenie (testovanie): *skúšanie a hodnotenie čiastkových kompetencií, na základe ktorých sa usudzuje na celkové kompetencie, napr. schopnosť gramaticky správne hovoriť sa posudzuje na základe správneho riešenia gramatických úloh*

integrative test - *test, ktorý vyžaduje od učiaceho sa používať niekoľko zručností naraz, ako napr. diktát, pri ktorom študent využíva svoje gramatické vedomosti, slovnú zásobu a demonštruje svoju zručnosť počúvania s porozumením*

interval scale – intervalová stupnica: *škála, ktorej stupne znamenajú rovnaké úseky (intervaly)*

item - položka testu, úloha, otázka: *každá položka v teste, ktorá sa hodnotí známkou (bodmi), napr. jedno doplnenie v teste s vynechanými položkami, jedna otázka s výberom odpovede z daných možností, jedna veta pre transformáciu, jedna otázka, na ktorú je potrebné odpovedať, aby bola úloha vyriešená, resp. splnená*

item analysis – položková analýza: *popis výkonu skupiny v jednotlivých položkách, zvyčajne zahrňujúci také štatistické indexy ako ľahkosť a rozlišovanie (facility, discrimination), bežne sa na túto analýzu používa software MicroCAT IteMan*

item banking - zhromažďovanie položiek: *prístup k spravovaniu testových položiek, zhromažďovanie a uskladňovanie informácií o položkách za účelom možnosti skonštruovania*

testov s určitým obsahom a úrovňou obtiažnosti, v praxi tento prístup využíva počítačovú databázu, kde sú položky zoskupené podľa úrovne obtiažnosti

item writer – tvorca testových položiek

language usage – vedomosti o jazyku: *funkcia lingvistického prvku, ako elementu lingvistického systému, manifestácia vedomostí o jazykovom systéme*

language use – schopnosť používať jazyk: *funkcia jazyka ako súčasť systému komunikácie, realizácia jazykového systému, ako zmysluplné komunikatívne správanie sa*

mark - známka: *výsledok skúšky, často percentuálne vyhodnotený, "mark" nie vždy zodpovedá konečnému súčtu bodov, keďže položky v teste môžu mať odlišnú váhu*

marker - ten, kto známkuje: *učiteľ, ktorý opravuje, známkuje a vyhodnocuje, resp. bodovo hodnotí test*

marking - známkovanie: *vyhodnocovanie testu pridelovaním známok, resp. bodov, percentuálne vyhodnotenie úspešnosti testu*

marking scheme - hodnotiacia schéma, schéma známkovania: *zoznam všetkých prijateľných odpovedí na otázky v teste s prisúdeným počtom pre každú položku. Takáto schéma umožňuje tomu, kto známkuje, presne pridelovať body v teste.*

matching – priradovanie: *testová úloha, v ktorej sa vyžaduje priradovanie elementov z dvoch separátnych zoznamov. Respondent napr. vyberá z daného zoznamu správne slovné spojenie na doplnenie vety alebo priraduje význam slovám, atď.*

mean – priemer (aritmetický): *priemer súboru skóre. Je to súčet všetkých skóre vydelený celkovým počtom položiek. Priemer je najčastejšie aplikovaná miera centrálnej tendencie.*

measurement - meranie: *vo všeobecnosti je to proces hľadania množstva niečoho porovnaním s ustálenou jednotkou, napr. používanie pravítka na meranie dĺžky. V sociálnych vedách sa meranie často vzťahuje na kvantifikáciu charakteristík (napr. jazyková zručnosť) osôb.*

measures of central tendency – miery centrálnej tendencie: pozri **central tendency**

measuring instrument - merací nástroj: *v praxi je to napr. nejaká forma skúšania, alebo testu.*

median – medián: *jeden z ukazovateľov strednej miery (centrálnej tendencie). Je to hodnota, ktorá rozdeľuje údaje zoradené podľa veľkosti do dvoch rovnakých polovic – jedna je nad a druhá pod mediánom.*

mistake - omyl: *nedostatok pri písomnom alebo ústnom prejave, ktorý je zapríčinený nedostatočnou pozornosťou, napr. únavou, nedbalosťou alebo iným aspektom výkonnosti.*

mode – modus: *skóre alebo hodnota, ktorá sa najčastejšie vyskytuje v distribúcii. Je to jedna z mier centrálnej tendencie.*

multiple-choice item – test s voliteľnými odpoveďami: *pri každej úlohe sa uvádza viacero možností, z ktorých zvyčajne iba jedna je správna*

multiple-matching – viacnásobné priradovanie: *testová úloha na rozpoznávanie správnej odpovede, v ktorej je daný určitý počet otázok alebo položiek na doplnenie vety, zvyčajne založených na prečítanom texte. Odpovede sú poskytnuté vo forme vopred danej skupiny slov alebo slovných spojení, z ktorých každé môže byť použité niekoľkokrát*

normal distribution – normálne rozdelenie: *rozdelenie nameraných hodnôt, ktoré pri grafickom znázornení má tvar symetrického zvonu, pričom priemer, medián a modus sú v jednom bode*

norm-referenced testing - testovanie relatívneho výkonu: *rozlišovacie testovanie (skúšanie), pri ktorom sa výkon jednotlivca porovnáva v rámci populácie, pre ktorú je test určený*

objective test - objektívny test: *test, v ktorom odpovede môžeme spoľahlivo a jednoznačne vyhodnotiť na základe hodnotiacej schémy, bez subjektívneho hodnotenia úlohy*

open-ended question (task) – otvorená otázka (úloha): *testová úloha v písomnom teste vyžadujúca od respondenta doplnenie odpovede. Cieľom tohto druhu otázky je získať nevynútenú odpoveď.*

oral interview – rozhovor: *ústna skúška hodnotiacia rečové zručnosti, ktorá prebieha formou dialógu medzi skúšaným a skúšajúcim*

percentile – percentil: *údaj ukazujúci, koľko percent skóre skupiny je pod výkonom daného jednotlivca. Napr. percentilové skóre 68 znamená, že 68% výkonov v skupine bolo pod uvedeným skóre.*

percentile rank - percentilové poradie: *údaj, ktorý charakterizuje umiestnenie osoby na škále od 0 do 100%. Číselná hodnota udáva, ako sa daná osoba umiestnila v sledovanej charakteristike vzhľadom k skúmanej populácii.*

performance - výkon, výkonnosť, uskutočnená práca: *aktuálne využívanie jazyka v konkrétnych situáciách*

placement test - test zaradenia, umiestnenia: *test používaný na umiestnenie žiakov do skupín s približne rovnakou úrovňou vedomostí a zručností*

practicality - praktickosť: *administrácia testu by mala byť čo najefektívnejšia, nemala by byť náročná na energiu, čas a financie*

pretesting = piloting a test - predskúšanie, pilotovanie testu, predtestovanie: *predskúšanie položiek testu na reprezentatívnej vzorke cieľovej skupiny, za účelom stanovenia ich obtiažnosti. Na základe následnej štatistickej analýzy sa použijú v teste len tie položky, ktoré splnili kritéria obtiažnosti, a tzv. „zlé“ položky sa z testu vynechajú, resp. nahradia inými.*

probability – pravdepodobnosť: *pomer vyjadrujúci pravdepodobnosť výskytu určitého javu*

productive skills - produktívne jazykové zručnosti: *aktívny spôsob využívania jazyka, sú to zručnosti, ktoré si vyžadujú aktívnu účasť jedinca na plnení úloh. Za aktívne jazykové zručnosti sa považujú rozprávanie (ústny prejav) a písomný prejav.*

proficiency test - test zručnosti, zbehlosti, výkonu: *test používaný na zistenie, či človek má potrebnú úroveň a typ vedomostí a zručností na zvládnutie danej úlohy alebo práce v budúcnosti. Tento typ testov nevychádza z predchádzajúcich učebných osnov, ale sa vzťahuje na špecifické akademické, resp. pracovné oblasti, v ktorých sa daná zručnosť vyžaduje.*

progress test - test pokroku v učení: *test používaný v jednotlivých etapách vzdelávacieho programu. Týmto typom testov sa zisťuje, či žiaci zvládli učivo posledných vyučovacích hodín.*

prompt - nápoď, ponuka: *napr. v písomných, alebo ústnych testoch sú to grafické materiály, alebo texty, ktoré majú pomôcť respondentovi vyriešiť úlohu, resp. zodpovedať otázku*

proof-reading task – korekčná úloha: *testovacia úloha, ktorá vyžaduje opravu špecifických chýb v texte, napr. vo hláskovaní alebo gramatických štruktúrach*

question - otázka: *niekedy sa používa ako pomenovanie úlohy, alebo položky v teste*

rank order - zoradenie, poradie podľa hodnotenia: *termín používaný pri lingvistickej analýze, označujúci poradie zostavené na základe určitých kritérií hodnotenia*

raw score - hrubé skóre: *výsledok testu vo forme pôvodnej numerickej hodnoty, nekonvertovanej na inú hodnotu. Je to prvý kvantitatívny výsledok testu, ktorý nie je štatisticky upravený a interpretovaný.*

rearrangement – usporiadanie: *úlohou kandidáta je napr. usporiadanie poprehadzovaných slov vo vete*

receptive skills - receptívne jazykové zručnosti: *pasívna jazyková zručnosť, ako napr. čítanie a poslech, pri ktorých sa neprodukuje jazykový akt*

reliability - reliabilita, spoľahlivosť: *termín zahrňujúci opakovateľnosť a vnútornú konzistentnosť, je to miera, v ktorej zisťovací nástroj umožní poskytnúť podobné výsledky v totožnej situácii*

rephrasing - parafrázovanie: *typ testovej úlohy, v ktorej respondent má vyjadriť danú myšlienku svojimi slovami, prípadne s použitím inej gramatickej štruktúry*

rubric - zadanie úlohy: *inštrukcie pre uchádzačov k vyplneniu úloh v teste*

sample – vzorka: *časť celku, ktorá je reprezentantom toho, ako celok vyzerá*

scale - stupnica, škála: *súbor čísiel, alebo kategórií na meranie niečoho. Rozlišujeme štyri druhy meracích stupníc - nominálnu, ordinálnu (poradovú), intervalovú a pomerovú.*

nominal scale - nominálna škála: *škála, pri ktorej čísla predstavujú kategórie (nie veľkosť alebo poradie, ale druh)*

ordinal scale - poradová škála: *typ meracej škály umožňujúcej stanoviť poradie, veľkosť a pod. Na rozdiel od intervalovej škály sú rozdiely medzi údajmi len relatívne.*

interval scale - intervalová škála: *škála, na ktorej úseky znamenajú rovnaké jednotky (intervaly)*

ratio scale - pomerová škála: *typ meracej škály umožňujúcej stanoviť okrem poradia a veľkosti rozdielov aj veľkosť násobkov a podielov. Pomerová škála musí mať pevný začiatok a hodnoty na škále majú len kladné znamienka.*

scanning – vyhľadávacie čítanie: *rýchle čítanie zamerané na vyhľadávanie špecifických informácií v texte*

scattergam, scatterplot – bodkový diagram: *graf, na ktorom osi predstavujú hodnoty dvoch premenných a každý člen skúmanej vzorky je reprezentovaný bodkou, pozícia ktorej je daná hodnotami dvoch veličín pre daný člen*

score - skóre, bodové ohodnotenie: *počet bodov pridelený za správne vyplnenie úlohy, resp. za správnu odpoveď*

scoring - hodnotenie: *prisudzovanie numerických hodnôt alebo skóre odpovediam v teste.*

script - rukopis: *papier obsahujúci uchádzačove odpovede na zadania v teste, používaný hlavne pri úlohách s voľnou odpoveďou*

sentence completion – dopĺňanie viet: *typ testovej úlohy, v ktorej má kandidát za úlohu doplniť chýbajúce slovo alebo slová vo vete*

sentence transformation – transformácia vety: *typ položky s vopred danou vetou slúžiacou ako náповeď, za ktorou nasleduje jedno alebo viac slov nasledujúcej vety, ktorá má vyjadrovať obsah predchádzajúcej. Úlohou respondenta je doplniť druhú vetu tak, aby použil inú gramatickú štruktúru, ale nemenil obsah predchádzajúcej.*

semi-authentic text - polo autentický text: *text vybraný zo skutočného zdroja, ktorého obtiažnosť (gramatika a slovná zásoba) je prispôsobená účelom testovania a úrovni respondentov*

setting - prostredie, situácia: *proces vypracúvania materiálov a zostavovania testov*

short answer questions – úlohy vyžadujúce stručnú odpoveď: *úlohy (otázky) pri testovaní porozumenia vyžadujúce krátke, zvyčajne jednoslovné odpovede*

significance level – úroveň významnosti: *pravdepodobná úroveň, pod ktorou sme ochotní považovať skúmané rozdiely za významné*

skewedness – skewed distribution – šikmé, kosé rozdelenie: *odchýlenie od normálneho rozdelenia. V takomto prípade krivka rozdelenia nie je v tvare symetrického zvonu, ale rozdelenie je posunuté buď smerom k vyšším, alebo nižším hodnotám (pozitívne/negatívne)*

skimming - rýchle čítanie: *čítanie s cieľom získať všeobecnú informáciu o obsahu textu*

specification - špecifikácia, spresnenie: *detailný popis charakteru skúšky, t. j. čo sa skúša, ako sa skúša, presné vymedzenie počtu a dĺžky zadání, úloh a položiek v teste, typy úloh a položiek, časová dotácia atď.*

standard error of measurement – smerodajná chyba merania: *štatistické meranie na určenie stupňa možnosti odchýlky predpokladaného parametra od vypočítaného*

standard deviation – smerodajná odchýlka, miera variability: *numerický index, ktorý vyjadruje rozptyl hodnôt (skóre) okolo priemeru*

standardized tests - štandardizovaný test: *profesionálne pripravený test, ktorý prešiel odborným posúdením, bol overený na dostatočnej vzorke osôb, má dostatočnú reliabilitu a validitu, sú k dispozícii jeho normy a podrobný manuál*

structural competence - štruktúrna kompetencia: *vedomosti jedinca o gramatických štruktúrach daného jazyka*

subjective testing - subjektívne skúšanie: *skúšanie, ktoré je hodnotené podľa osobného posúdenia skúšajúceho (známkujúceho), sú to písomné úlohy, napr. kompozícia, esej a i.*

summative evaluation - sumatívne hodnotenie: *hodnotenie, ktoré je globálne, súhrnné, záverečné, výstupné a má obyčajne podobu binárnu: prosperel-neprosperel, postúpil-nepostúpil*

table-top marking – metóda známkovania testov, ktorá prebieha spoločne na jednom mieste v určitom časovom úseku

task - úloha: *kombinácia zadania, spracovania a odpovede. Je to činnosť, ktorá sa vykonáva ako výsledok spracovania alebo porozumenia jazyka (napr. odpoveď). Úlohy môžu, ale nemusia vyústiť do produkcie jazyka. Od učiteľa sa vyžaduje, aby špecifikoval, čo bude považovať za úspešné splnenie úlohy. Používanie rôznych druhov úloh vo vyučovaní jazykov je v súlade s komunikatívnym prístupom k učeniu sa cudzích jazykov.*

Threshold level – úroveň „threshold“: *určitá jazyková úroveň angličtiny stanovená Radou Európy, ktorá predpokladá, že začiatočník potrebuje asi 375 vyučovacích hodín na jej dosiahnutie*

test - test, skúška: *prostriedok na meranie množstva a kvality vedomostí*

testee – kandidát, respondent, testant: *jedinec, ktorý je skúšaný / hodnotený*

test battery – testová batéria: *súbor testov, ktorý sa naraz dáva respondentom*

test item – testová položka: *slovo, otázka, úloha alebo iný prvok ktorú musí kandidát doplniť, resp. zodpovedať v teste. V jazykových testoch sa používajú rôzne typy položiek.*

test format – typ testu, resp. typy položiek v teste: *v jednom teste môžu byť viaceré druhy položiek v rámci jednotlivých úloh*

test-maker – tvorca, zostavovateľ testu

test-taker = **testee** – kandidát, respondent, testant

trialog – diagnostikovanie: *štádium v príprave, zostavovaní testových úloh, s cieľom zistiť, či test funguje tak, ako sa predpokladalo. Väčšinou sa to robí so subjektívne hodnotenými úlohami, ktoré sú preskúšané na limitovanej vzorke populácie.*

true/false items– úloha s alternatívnou odpoveďou (pravdivá/nepravdivá, áno/nie)

value – hodnota: *hodnota (vyjadrená v bodoch, resp. percentách) priradená jednotlivým položkám*

validity - validita, platnosť, adekvátnosť: *vyjadruje nakoľko výskumný nástroj meria to, čo chceme merať a nie niečo iné. Poznáme niekoľko druhov validít:*

concurrent validity - súbežná validita: *miera zhody výsledkov výskumného nástroja s iným výskumným nástrojom, o ktorom je známe, že zisťuje tú istú vec*

construct validity - konštruktová validita: *miera zhody výskumného nástroja s konštruktom. Konštrukt je pojem označujúci jav, ktorý je priamo pozorovateľný (napr. zručnosť). Je to forma validity založená na miere do akej položky v teste odzrkadľujú základné aspekty teórie na ktorej je test založený*

content validity - obsahová validita: *miera, ktorou výskumný nástroj pokrýva celú tematickú oblasť, ktorú má skúmať. Obsahová validita je dôležitou kvalitou testov absolútneho výkonu., v ktorých obsah testu musí reprezentovať obsah učiva prebratom počas vyučovacieho procesu.*

face validity - validita na prvý pohľad: *zistenie, či výskumný nástroj zisťuje, čo zisťovať má. Ide o subjektívny odhad učiteľov a študentov, nie o štatistické zistenie.*

variable – premenná veličina, hodnota: *vlastnosť, na základe ktorej sa členovia skupiny od seba odlišujú*

vetting – posudzovanie a skúmanie položiek v teste: *fáza v cykle prípravy testu, v ktorej sa hodnotia materiály pripravené tvorcami testových položiek a rozhoduje sa o tom, ktoré testové položky budú vylúčené pretože nespĺňajú kritériá špecifikácie, a ktoré postúpia ďalej do fázy prípravy testu*

Waystage level – úroveň „Waystage“: *špecifikáca základnej úrovne cudzojazyčnej kompetencie vydané Radou Európy v r. 1977 a revidovanej pre anglický jazyk v r. 1990, ktorá vyžaduje nižšiu jazykovú úroveň, približne polovicu učiva úrovne „Threshold“*

weighing – váženie: *prisudzovanie rôzneho maximálneho počtu bodov jednotlivým položkám, úlohám alebo komponentom v teste za účelom zmeny ich relatívneho prínosu vo vzťahu*

k ostatným častiam testu. Napr., ak niektorej úlohe alebo časti testu prisúdime vyššie skóre, tak bude mať v teste väčšie pomerné zastúpenie, väčšiu váhu v rámci celkového skóre ako ostatné.

word formation – slovotvorba: typ úlohy, v ktorej kandidát má vytvoriť tvar slova odvodeného z iného tvaru toho istého slova, ktoré je vopred dané. Novovytvorené slovo musí formálne aj funkčne zapadať do daného textu.

4.1 Zoznam najčastejšie používanej terminológie

DRUHY JAZYKOVÝCH TESTOV podľa ich účelu

achievement test.

aptitude test

attainment test

diagnostic test

placement test

proficiency test

progress test

VLASTNOSTI KVALITNÉHO TESTU

acceptability

backwash-effect

practicability

reliability

validity: **concurrent**

content

construct

face

predictive

TYPY TESTOVÝCH ÚLOH A POLOŽIEK (TEST FORMATS)

banked cloze

cloze

communicative task

C- test

error correction

error-recognition

dichotomous

directed writing task

discursive composition

gap-filling

guided writing task

information transfer

matching

multiple-choice item

multiple-matching

open-ended question
open-ended task
oral interview
proof-reading task
rearrangement
rephrasing
rewriting
scanning
sentence completion
sentence transformation
short answer questions
skimming
true/false
word formation

ANALÝZA A INTERPRETÁCIA VÝSLEDKOV TESTU

correlation
correlation coefficient
critical values
discrimination
discrimination index
dispersion
facility index
frequency distribution
frequency polygon
histogram
interval scale
item analysis
mean
measures of central tendency
median
mode
normal distribution
percentile rank
probability
sample
scale
scatterplot
significance
skewedness
standard deviation
standard error of measurement
utility
value
variable
weighing

Zoznam použitej literatúry

ALDERSON, J. Charles – CLAPHAM, C. – WALL, D. *Language Test Construction and Evaluation*. Cambridge : Cambridge University Press, 1995.

BACHMAN, L. F. *Fundamental Considerations in Language Testing*. Oxford : Oxford University Press, 1991.

BROWN, J. D. *Understanding Research in Second Language Learning*. Cambridge : Cambridge University Press, 1988.

GAVORA, J. – MAREŠ, J. *Anglicko–slovenský pedagogický slovník*. Bratislava : Iris, 1996.

HARRISON, A. *A Language Testing Handbook*. London : MacMillan Press, 1983.

HUGHES, A. *Testing for Language Teachers*. Cambridge : Cambridge University Press, 1994.

KOSOVÁ, B. Humanizácia procesu hodnotenia žiakov. In: *Pedagogické rozhľady*, 96/97, č. 4, s.1 – 8. Metodické centrá Slovenska.

Modern Languages: Learning, Teaching, Assessment. A Common European Framework of Reference. Strasbourg: Council of Europe, 1996.

PRÚCHA J., WALTEROVÁ E, MAREŠ J. *Pedagogický slovník*. Praha: Portál, 1995.

SLAVÍK, J. *Hodnocení v současné škole*. Praha : Portál, 1999.

RICHARDS, J. – PLATT, J. – WEBER, H. *Longman Dictionary of Applied Linguistics*. Harlow : Longman, 1989.

RÖTLING, G. Didaktický test ako nástroj rozvoja. In: *Pedagogické rozhľady žiaka*, 2000, č. 5, s. 11-14.

TUREK, I. *Zvyšovanie efektívnosti vyučovania*. Bratislava : Metodické centrum v Bratislave, 1997.

2PhDr. Rita RAFAJLOVIČOVÁ, PhD. pôsobí ako odborná asistentka na Katedre anglického jazyka a literatúry na Fakulte humanitných a prírodných vied (FHPV) Prešovskej univerzity od roku 1990. Tu vedie prednášky a semináre z anglickej syntaxe a vybrané kapitoly z aplikovanej lingvistiky. Testovaním anglického jazyka sa zaoberá už niekoľko rokov a v r. 2000 obhájila svoju dizertačnú prácu na tému: *Testing Grammar and its Assessment from the Linguistic Viewpoint*. Je autorkou niekoľkých publikácií a vystúpení na konferenciách s tematikou testovania anglického jazyka.

Mgr. Milena ŠTULRAJTEROVÁ je odbornou asistentkou na Katedre anglistiky a amerikanistiky Fakulty humanitných vied Univerzity Mateja Bela v Banskej Bystrici. Vedie kurzy Didaktika vyučovania anglického jazyka, Metodológia výskumu a Americké štúdie. Predmetom jej výskumu je skúšanie a hodnotenie ústneho prejavu v cudzom jazyku.