

Koncepcia zavádzania kriteriálneho testovania na Slovensku

Východiská, podmienky a smerovanie

STRATEGICKÝ MATERIÁL PRE IMPLEMENTÁCIU
KURIKULÁRNEJ REFORMY

Autorka
PhDr. Romana Kanovská

Bratislava 2026

OBSAH

ÚVOD	4
1. MEDZINÁRODNÉ TRENDY V KRITERIÁLNOH TESTOVANÍ	7
1.1 KOMPARATÍVNY RÁMEC A VÝCHODISKÁ ANALÝZY VYBRANÝCH VZDELÁVACÍCH SYSTÉMOV	7
1.2 NEMECKO – SYSTÉM VERA A DÔSLEDKY PISA ŠOKU	12
<i>Nemecký vzdelávací systém</i>	13
<i>Primárne vzdelávanie</i>	13
<i>Nižšie sekundárne vzdelávanie</i>	14
<i>Vyššie sekundárne vzdelávanie</i>	14
<i>Externé hodnotenie žiakov</i>	14
<i>Inštitucionálne zabezpečenie testovania</i>	16
<i>Vývoj kritériálnych testov</i>	17
<i>Vývoj modelov úrovni spôsobilostí</i>	18
<i>Národné hodnotiace štúdie a sledovanie trendov</i>	19
<i>Dotazníkové prieskumy</i>	20
<i>Testovanie s využitím počítačov – výhody a nevýhody</i>	20
<i>Testovanie s využitím počítača a akademické výsledky</i>	21
<i>Limity prenositeľnosti a implikácie pre Slovensko</i>	24
1.3 ESTÓNSKO – MODEL KOMBINÁCIE EXTERNEJ EVALVÁCIE A FORMATÍVNEHO HODNOTENIA	25
<i>Estónsky vzdelávací systém</i>	26
<i>Základné vzdelávanie</i>	26
<i>Všeobecné stredoškolské vzdelávanie</i>	27
<i>Odborné školstvo</i>	27
<i>Externé hodnotenie žiakov v Estónsku</i>	27
<i>Testy s nízkymi dôsledkami (low-stakes tests)</i>	28
<i>Spätná väzba</i>	29
<i>Testy na hodnotenie kompetencií</i>	30
<i>Zameranie na budúcnosť</i>	30
<i>Nástroje prieskumu</i>	30
<i>Žiacke prieskumy</i>	30
<i>Prieskumy medzi učiteľmi</i>	31
<i>Digitálne technológie vo vzdelávaní</i>	32
<i>Limity prenositeľnosti estónskeho modelu</i>	33
1.4 PORTUGALSKO – REFORMY PO ROKU 2000 A ICH VÝVOJ	34
<i>Portugalský vzdelávací systém</i>	34
<i>Základné vzdelávanie</i>	35
<i>Vyššie sekundárne vzdelávanie</i>	36
<i>Externé hodnotenie žiakov v Portugalsku</i>	36
<i>Vplyv štúdie PISA na zmeny vo vzdelávacom systéme Portugalska</i>	37
<i>Pokrok Portugalska vo vzdelávaní</i>	38
<i>Zavedenie národných externých testovaní od roku 2005</i>	38
<i>Zmeny vzdelávacích politík po roku 2015</i>	40
<i>Limity prenositeľnosti portugalského modelu</i>	41
2. METODOLOGICKÁ PRÍPRAVA A VÝVOJ KRITERIÁLNYCH TESTOV NA SLOVENSKU	43
2.1 TEORETICKÉ A METODOLOGICKÉ VÝCHODISKÁ KRITERIÁLNEHO TESTOVANIA	43
2.2 METODOLOGICKÁ PRÍPRAVA CR-TESTOVANÍ NA SLOVENSKU	44
<i>Medzinárodná spolupráca</i>	44
<i>Porovnanie kritériálnych a normatívnych testov z hľadiska účelu</i>	45
<i>Ilustračný príklad rozdielov medzi kritériálnym a normatívnym testovaním</i>	46

3. METODIKA TVORBY KRITERIÁLNYCH TESTOV	49
3.1 ZÁKLADNÉ KROKY PRI VÝVOJI KRITERIÁLNYCH TESTOV	49
3.2 TVORBA CR-TESTOV Z HĽADISKA ROLÍ EXPERTOV	54
3.3 VÝVOJ KRITERIÁLNYCH TESTOV NA SLOVENSKU	55
<i>Teoretický rámec</i>	55
<i>Dizajn testovacieho nástroja</i>	61
<i>Psychometrické vyhodnocovanie testov</i>	61
4. SPÄTNÁ VÄZBA A VYUŽITIE VÝSLEDKOV KRITERIÁLNYCH TESTOV	65
4.1 POSKYTOVANIE SPÄTNEJ VÄZBY NA ZÁKLADE VÝSLEDKOV KRITERIÁLNYCH TESTOV	67
4.2 VYUŽITIE VÝSLEDKOV KRITERIÁLNYCH TESTOV NA FORMATÍVNE HODNOTENIE	69
4.3 LIMITY INTERPRETÁCIE VÝSLEDKOV KRITERIÁLNEHO TESTOVANIA	71
5. IMPLEMENTÁCIA SYSTÉMU KRITERIÁLNYCH TESTOVANÍ NA SLOVENSKU	73
5.1 BANKA ÚLOH A PILOTNÉ TESTOVANIA	73
5.2 HLAVNÉ MERANIA	74
5.3 DOTAZNÍKOVÉ PRIESKUMY	75
5.4 ELEKTRONICKÉ TESTOVANIA	76
5.5 UDRŽATEĽNOSŤ MERANÍ, ĎALŠÍ VÝSKUM A VÝVOJ V PREPOJENÍ S KURIKULÁRNymi ZMENAMI	77
ZÁVER: PODMIENKY A SMEROVANIE KRITERIÁLNEHO TESTOVANIA NA SLOVENSKU	78
ZOZNAM POUŽITEJ LITERATÚRY	81

ÚVOD

Každé hodnotenie výsledkov vzdelávania odpovedá na určitú otázku. Môže ukázať, ako si žiak alebo škola vedú v porovnaní s ostatnými, ale môže tiež zisťovať, do akej miery žiaci dosiahli konkrétne vzdelávacie ciele. Tieto dva pohľady nie sú totožné. Kým prvý poskytuje informáciu o relatívnom postavení, druhý umožňuje obsahovo interpretovať, čo žiaci podľa výsledkov pravdepodobne ovládajú, v ktorých oblastiach sa im darí, a kde potrebujú ďalšiu podporu. Práve na druhom z týchto pohľadov je založené kritériálne orientované hodnotenie.

Potreba rozvoja kritériálneho testovania na Slovensku úzko súvisí s prebiehajúcou kurikulárnou reformou. Nový štátny vzdelávací program vymedzuje vzdelávacie ciele, očakávané výkony a kompetencie, ktoré majú žiaci postupne rozvíjať a dosahovať v jednotlivých vzdelávacích cykloch. Ak má systém hodnotenia poskytovať informácie využiteľné pri monitorovaní implementácie reformy, nemôže sa obmedzovať iba na určenie relatívneho postavenia žiakov v porovnávaní populácii. Potrebne sú aj údaje (ďalej aj „dátá“) umožňujúce obsahovo interpretovať, do akej miery sa dosahujú ciele a výkonové štandardy vymedzené v kurikule. Kritériálne testovanie preto môže predstavovať dôležitý nástroj monitorovania implementácie reformy. Po ukončení jednotlivých vzdelávacích cyklov môže poskytovať informácie o dosahovaní vzdelávacích štandardov, poukazovať na oblasti vyžadujúce ďalšiu podporu a vytvárať podklady na odborné rozhodovanie na úrovni škôl aj vzdelávacej politiky. V spojení s ďalšími kvantitatívnymi a kvalitatívnymi zdrojmi údajov môže pomáhať sledovať, ako sa kurikulárne zámery premietajú do výsledkov učenia žiakov, a podporovať priebežné vyhodnocovanie a zdokonaľovanie implementácie reformy na základe dôkazov.

Box 1: Prečo Slovensko potrebuje kritériálne testovanie?

Kurikulárna reforma mení spôsob, akým sa vymedzujú očakávané výsledky vzdelávania. Ak majú školy rozvíjať vedomosti, zručnosti a kompetencie opísané v novom štátnom vzdelávacom programe, vzdelávací systém potrebuje nástroje, ktoré poskytujú informácie o tom, do akej miery žiaci tieto očakávania dosahujú.

Kritériálne testovanie odpovedá na otázky, ktoré nemožno dostatočne zodpovedať iba porovnávaním výkonu žiakov:

Do akej miery žiaci dosahujú vzdelávacie ciele a výkonové štandardy vymedzené v novom kurikule?
Ktoré vedomosti, zručnosti a kompetencie žiaci pravdepodobne zvládajú a v ktorých oblastiach sa prejavuje potreba ďalšej podpory?

Aký obraz o dosahovaní vzdelávacích cieľov poskytujú výsledky žiakov po ukončení jednotlivých vzdelávacích cyklov?

V ktorých oblastiach sa výsledky žiakov približujú očakávaným štandardom a ktoré oblasti si vyžadujú podrobnejšiu analýzu?

Aké podklady môžu výsledky poskytnúť školám, tvorcom kurikula a vzdelávacej politike pri plánovaní ďalších krokov?

Kritériálne testovanie preto nepredstavuje iba odlišný spôsob interpretácie výkonu žiakov. Môže byť aj jedným z nástrojov monitorovania implementácie kurikulárnej reformy: poskytovať informácie o dosahovaní vzdelávacích štandardov, podporovať spätnú väzbu pre školy a vytvárať podklady na ďalšie odborné rozhodovanie. V spojení s pozorovaním učiteľov, žiackymi prácami, údajmi o podmienkach vzdelávania a ďalšími kvantitatívnymi a kvalitatívnymi zdrojmi môže prispievať k priebežnému vyhodnocovaniu a zdokonaľovaniu implementácie reformy.

Ak kurikulárna reforma vymedzuje, čo sa majú žiaci naučiť, kritériálne testovanie pomáha získavať informácie o tom, do akej miery sa tieto očakávania naplňujú.

Samotné zavedenie nového prístupu k testovaniu a interpretácii výsledkov však kvalitu vzdelávania automaticky nezvýši. Hodnota externého hodnotenia závisí od jasného vymedzenia vzdelávacích cieľov, primeraného zastúpenia hodnoteného obsahu v testových úlohách, presnosti merania a správnej interpretácie výsledkov. Rovnako dôležité je, na aké účely sa výsledky používajú. Testovanie môže podporovať zlepšovanie vzdelávania, ale jeho neprimerané využívanie môže viesť k prispôbovaniu vyučovania obsahu a formátu testov na úkor ďalších vzdelávacích cieľov alebo k nepodloženým záverom o žiakoch, učiteľoch a školách.

Kriteriálne testovanie preto nemožno chápať iba ako technickú zmenu spôsobu výpočtu alebo prezentácie skóre. Predstavuje prístup, pri ktorom sa výsledky interpretujú vo vzťahu k obsahovo vymedzeným vzdelávacím cieľom, výkonovým štandardom alebo výkonovým úrovniám. Jeho primárnou otázkou nie je, či žiak dosiahol lepší výsledok než ostatní, ale čo jeho výkon vypovedá o miere dosiahnutia očakávaných vzdelávacích cieľov. Normatívna a kritériálna interpretácia pritom odpovedajú na rozdielne otázky a môžu sa vzájomne dopĺňať.

Navrhovaný systém kritériálneho testovania na Slovensku má byť koncipovaný predovšetkým ako externé hodnotenie s nízkymi dôsledkami. Jeho výsledky nemajú slúžiť na vytváranie verejných rebríčkov, sankcionovanie škôl alebo učiteľov ani ako jediný podklad na rozhodovanie o ďalšej vzdelávacej dráhe jednotlivých žiakov. Majú predstavovať jeden zo zdrojov informácií o dosahovaní vzdelávacích cieľov, podporovať sebahodnotenie škôl a poskytovať podklady na odborné rozhodovanie o ďalšom rozvoji kurikula a vzdelávacej politiky. Formatívny prínos výsledkov pritom nevzniká automaticky ich získaním, ale až vtedy, keď sa pretavia do zrozumiteľnej spätnej väzby a využijú na plánovanie ďalších krokov vo vyučovaní a učení.

Zahraničné skúsenosti ukazujú, že kvalitný systém externého hodnotenia nemožno vybudovať jednorazovým zavedením testov ani mechanickým prevzatím modelu inej krajiny. Nemecko poukazuje na význam jasných vzdelávacích štandardov, odborného konsenzu a transparentného prepájania kurikula s testovými úlohami. Estónsko zdôrazňuje kultúru práce s dátami, dôveru v profesionalitu učiteľov a prepojenie externého hodnotenia so sebahodnotením škôl. Portugalsko upozorňuje, že zlepšenie vzdelávacích výsledkov nemožno pripísať len samotnému testovaniu a že testy s vysokými dôsledkami môžu popri zvýšenej motivácii k výkonu vytvárať tlak na žiakov a školy a podporovať prispôbovanie vyučovania testom. Tieto skúsenosti nepredstavujú návody na kopírovanie, ale zdroj poučení pre vytvorenie systému zodpovedajúceho slovenským podmienkam.

Predkladaná koncepcia prepája tri roviny. Prvou sú medzinárodné skúsenosti s kritériálne orientovaným a externým hodnotením. Druhou sú teoretické a metodologické východiská pedagogického merania a vývoja testovacích nástrojov. Treťou je návrh postupnej implementácie systému na Slovensku vrátane pilotného overovania, poskytovania spätnej väzby, elektronického testovania, budovania bánk úloh a rozvoja odborných kapacít. Osobitná pozornosť sa venuje vzťahu medzi kritériálnou interpretáciou výsledkov a psychometrickými modelmi vrátane teórie odpovede na položku (IRT), ktorá môže podporiť prepájanie testových foriem, tvorbu spoločných škál a v dlhodobšej perspektíve aj rozvoj adaptívneho testovania.

Cieľom dokumentu nie je prezentovať kritériálne testovanie ako bezchybnú ani jedinou cestu hodnotenia výsledkov vzdelávania. Jeho cieľom je vymedziť obsahové, metodologické, psychometrické a organizačné podmienky, za ktorých môže tento prístup poskytovať dostatočne presné a využiteľné výsledky a podporovať ich odborne podloženú interpretáciu. Dokument zároveň pomenúva limity a riziká tohto prístupu, aby sa už pri zavádzaní systému vytvorili podmienky na predchádzanie neprimeraným interpretáciám a používaniu výsledkov na účely, pre ktoré testy neboli navrhnuté.

Úspech navrhovanej zmeny nebude závisieť iba od kvality jednotlivých testov. Rozhodujúca bude schopnosť dlhodobo prepájať dostatočne jasné a obsahovo stabilné kurikulum, odbornú kvalitu merania, transparentnú interpretáciu výsledkov, metodickú podporu škôl a zodpovedné využívanie dát. Iba tak sa externé hodnotenie môže stať nie cieľom reformy, ale jedným z nástrojov jej odborného sledovania a podpory.

1. Medzinárodné trendy v kriteriálnom testovaní

Táto kapitola predstavuje vybrané medzinárodné prístupy ku kriteriálnemu testovaniu a externému hodnoteniu žiakov na príklade vzdelávacích systémov Nemecka, Estónska a Portugalska. Analýza ukázala, že účinné systémy hodnotenia sú založené na prepojení jasne formulovaných vzdelávacích štandardov, pravidelného externého testovania a zmysluplnej spätnej väzby pre školy. Analyzované krajiny rôznym spôsobom prepájajú externé hodnotenie so vzdelávacími štandardmi, spätnou väzbou a využívaním výsledkov na úrovni škôl alebo vzdelávacieho systému. Miera formatívneho využitia a dôsledky testovania sa však medzi jednotlivými krajinami líšia. Identifikované medzinárodné skúsenosti a dobré praxe tvoria východisko pre návrh a postupné zavádzanie systému kriteriálneho testovania v podmienkach Slovenskej republiky.

1.1 Komparatívny rámec a východiská analýzy vybraných vzdelávacích systémov

Pri posudzovaní účelu a možných spôsobov využitia kriteriálneho testovania je potrebné rozlišovať medzi testovaním s nízkymi a vysokými dôsledkami, keďže toto rozlíšenie ovplyvňuje požiadavky na dizajn testu, interpretáciu výsledkov aj nastavenie ochranných mechanizmov.

Box 2: Testovanie s nízkymi a vysokými dôsledkami

V pedagogickom hodnotení sa rozlišuje medzi **testovaním s nízkymi dôsledkami (*low-stakes testing*)** a **testovaním s vysokými dôsledkami (*high-stakes testing*)**. Rozhodujúci pritom nie je samotný typ testu, ale spôsob využitia jeho výsledkov a závažnosť rozhodnutí, ktoré sa na ich základe prijímajú.

Pri **testovaní s nízkymi dôsledkami** sa výsledky nepoužívajú ako rozhodujúci podklad na prijímanie závažných rozhodnutí o ďalšej vzdelávacej dráhe žiaka, hodnotení učiteľov, financovaní škôl alebo uplatňovaní sankcií. Slúžia najmä na získavanie informácií o vzdelávacích výsledkoch, poskytovaní spätnej väzby, podporu pedagogického rozhodovania, plánovanie ďalšieho učenia, sebahodnotenie škôl alebo monitorovanie kvality vzdelávania.

Pri **testovaní s vysokými dôsledkami** sú výsledky priamo spojené so závažnými rozhodnutiami. Môže ísť napríklad o prijatie na ďalší stupeň vzdelávania, ukončenie štúdia, získanie kvalifikácie alebo o rozhodnutia týkajúce sa hodnotenia, financovania či ďalšieho fungovania škôl a hodnotenia učiteľov.

Rozlíšenie medzi nízkymi a vysokými dôsledkami nie je vždy úplne ostré. Miera dôsledkov sa môže líšiť podľa toho, koho sa výsledky týkajú. Testovanie môže mať nízke dôsledky pre jednotlivého žiaka, ale zároveň môže byť spojené s významnejšími dôsledkami pre školu alebo jej zriaďovateľa. Dôsledky môžu vzniknúť aj dodatočným použitím výsledkov na účely, na ktoré test pôvodne nebol navrhnutý.

Ak sú s výsledkami spojené významné dôsledky, môže sa zvyšovať tlak na školy a učiteľov, aby prispôbovali vyučovanie obsahu a formátu testov. To môže viesť k nadmernému precvičovaniu testových úloh a k obmedzovaniu netestovaných, komplexnejších alebo ťažšie merateľných vzdelávacích cieľov. Takéto riziko nevzniká automaticky pri každom testovaní, ale rastie so závažnosťou dôsledkov spojených s jeho výsledkami.

Označenia ***low-stakes*** a ***high-stakes*** preto nevypovedajú o kvalite ani o psychometrických vlastnostiach testu. Charakterizujú spôsob využitia jeho výsledkov a závažnosť nadväzujúcich rozhodnutí. Kriteriálne aj normatívne interpretované výsledky možno využívať v režime nízkych aj vysokých dôsledkov.

Model **kriteriálneho testovania** navrhovaný v tomto materiáli je koncipovaný predovšetkým ako **testovanie s nízkymi dôsledkami**. Jeho cieľom je poskytovať školám a učiteľom využiteľné informácie o miere dosahovania vzdelávacích cieľov a výkonových štandardov. Výsledky môžu v spojení s ďalšími dôkazmi o učení podporovať aj spätnú väzbu žiakom. Nemajú však slúžiť na prijímanie závažných rozhodnutí o jednotlivých žiakoch, učiteľoch alebo školách.

Koncepcia zavádzania kritériálneho testovania do škôl na Slovensku vychádza z komparatívnej analýzy medzinárodných skúseností s externým hodnotením žiakov. Táto časť predstavuje analytický rámec a východiská pre skúmanie vybraných vzdelávacích systémov, ktorých prístupy k hodnoteniu (ďalej aj „evalvácia“) sú následne interpretované vo vzťahu k možnostiam ich využitia v podmienkach Slovenska. Dôraz sa pritom kladie na identifikáciu systémových prvkov, ktoré prispievajú k efektívnemu prepojeniu externého testovania, spätnej väzby a zlepšovania kvality vzdelávania.

V tejto kapitole sú analyzované systémy národného hodnotenia žiakov vo vybraných krajinách, ktoré dlhodobo dosahujú nadpriemerné alebo stabilné výsledky v medzinárodnom meraní OECD PISA (Programme for International Student Assessment). Výber krajín bol realizovaný na základe vopred stanovených kritérií, ktoré zahŕňajú: (1) dlhodobé výsledky žiakov v PISA, (2) existenciu rozvinutého systému externého hodnotenia a (3) preukázateľné využívanie výsledkov testovania na podporu zlepšovania vzdelávacieho procesu. Takto definovaný výber umožňuje identifikovať relevantné prístupy k externému hodnoteniu a analyzovať ich potenciál pre aplikáciu v podmienkach slovenského vzdelávacieho systému.

Zároveň je dôležité uviesť, že úspech týchto krajín nemožno pripisovať jednému faktoru, ale kombinácii systémových opatrení. Významnú úlohu však zohráva prepojenie externého testovania s poskytovaním spätnej väzby školám, učiteľom a žiakom. V analyzovaných krajinách sa postupne etablovalo **kritériálne testovanie ako nástroj externého formatívneho hodnotenia**, pričom miera jeho využitia a dôsledky výsledkov sa medzi jednotlivými systémami výrazne líšia. Vo väčšine krajín výsledky tohto testovania slúžia pre školy ako zdroj informácií na skvalitňovanie vyučovacích postupov a podporu učenia žiakov. Kritériálne testovanie (criterion-referenced assessment) je založené na porovnávaní výkonu žiaka s vopred definovanými štandardmi, na rozdiel od normatívneho hodnotenia (norm-referenced assessment), ktoré porovnáva výkon jednotlivca s referenčnou skupinou (Glaser, 1963).¹

Kritériálne orientované hodnotenie môže plniť aj významnú formatívnu funkciu, pretože poskytuje informáciu nielen o celkovom výsledku žiaka, ale najmä o tom, ktoré vopred vymedzené vedomosti, zručnosti alebo úrovne výkonu už žiak dosiahol a v ktorých oblastiach potrebuje ďalšiu podporu. Umožňuje tak porovnať aktuálny výkon žiaka s očakávaným štandardom, pomenovať rozdiel medzi nimi a určiť ďalšie kroky učenia (Sadler, 1989)². Formatívny účinok však nevzniká samotným zadaním testu alebo oznámením výsledku. Je podmienený tým, že výsledky sú pretavené do zrozumiteľnej, konkrétnej a využiteľnej spätnej väzby a následne použité na úpravu učenia žiaka alebo vyučovacieho postupu učiteľa (Black a Wiliam, 1998, Hattie a Timperley, 2007)³. Takto koncipovaná spätná väzba pomáha žiakovi porozumieť cieľom učenia, sledovať vlastný pokrok a postupne preberať väčšiu zodpovednosť za riadenie svojho učenia sa.

Prvou z predstavovaných krajín v tomto dokumente je *Nemecko*. Po tom, čo sa nemecký vzdelávací systém podľa medzinárodného merania PISA v roku 2000 ocitol v porovnaní výsledkov žiakov pod priemerom krajín

¹ GLASER, R. 1963. *Instructional technology and the measurement of learning outcomes: Some questions*. Washington, D.C.: American Psychological Association.

² SADLER, D. R. 1989. Formative Assessment and the Design of Instructional Systems. *Instructional Science*, roč. 18, č. 2, s. 121. Dostupné na: <https://doi.org/10.1007/BF00117714>.

³ BLACK, P. a WILIAM, D. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*. 1998, roč. 5, č. 1, s. 7–10. Dostupné na: <https://doi.org/10.1080/0969595980050102>; HATTIE, J. – TIMPERLEY, H. 2007. The Power of Feedback. *Review of Educational Research*. 2007, roč. 77, č. 1, s. 86–90. Dostupné na: <https://doi.org/10.3102/003465430298487>.

OECD, naštartovali sa v Nemecku významné systémové reformy vo vzdelávaní. Nemecko redefinovalo ciele kurikula, obsahy aj vzdelávacie štandardy. Táto krajina zároveň investovala veľa odborných personálnych kapacít a finančných zdrojov do tvorby národných diagnostických nástrojov a testov s cieľom čo najpresnejšie monitorovať úroveň vzdelávania na školách v jednotlivých spolkových krajinách. Silnou inšpiráciou pri vývoji testovacieho rámca a testov, ktoré sa nazývajú VERA, bola práve metodológia štúdie PISA. Po prijatí systémových reforiem je nemecký vzdelávací systém v medzinárodnom meraní PISA dlhodobo nad úrovňou alebo na úrovni priemeru krajín OECD (pozri obr. 1 – 3), čo naznačuje možnú súvislosť medzi implementovanými reformami a vývojom výsledkov, pričom vplyv jednotlivých opatrení nie je možné jednoznačne izolovať. Nemecké odborné tímy podrobujú výskumom a analýzám aj postupné zavádzanie elektronického testovania do škôl, napr. detailne skúmajú **vplyv počítačového prostredia na úspešnosť žiakov**. Zistenia vyvodzované z analýz porovnávajúcich výsledky papierových testov a testov administrovaných na počítačoch sú mimoriadne dôležité aj pre slovenský vzdelávací systém, ktorý prechádza podobným procesom smerujúcim k čoraz významnejšiemu účelnému využívaniu digitálneho testovania na školách.

Druhý vzdelávací systém, ktorý je v tejto štúdii predstavený z pohľadu celoštátneho monitorovania úrovne vedomostí a zručností žiakov, je estónsky vzdelávací systém. *Estónsko* je krajina s podobným historickým vývojom ako Slovensko, ktorá si ako postsocialistická krajina začala v deväťdesiatych rokoch minulého storočia budovať a spravovať vlastný štát a veľmi skoro zaviedla reformy vo vzdelávaní. Dnes patrí estónsky vzdelávací systém k jedným z najúspešnejších na svete a v medzinárodnom meraní PISA sa dlhodobo pohybuje nad priemerom krajín OECD (pozri obr. 1 – 3). Estónsko zaviedlo v druhej polovici 90. rokov celoštátne testy ako systém externej evalvácie základných a stredných škôl, a to normatívne maturitné testy a testovanie žiakov pri ukončovaní sekundárneho stupňa vzdelávania v 9. ročníku. Výsledky týchto testovaní sú akceptované v prijímacom konaní na vyšší stupeň vzdelávania.

Pre slovenský vzdelávací systém je veľmi inšpiratívny vývoj formatívneho hodnotenia (*low-stakes tests*) v Estónsku, na ktoré estónski odborníci upriamujú v poslednom období veľkú pozornosť. Postupným zavádzaním takýchto testovaní, ktoré sú na školách administrované dobrovoľne a sú na vzorke základných škôl centrálné vyhodnocované, buduje Estónsko kultúru hodnotenia, ktoré prispieva k rozvoju kultúry autoevalvácie škôl. V estónskom vzdelávacom systéme je kladený silný dôraz na efektívnu spätnú väzbu vychádzajúcu z výsledkov externého hodnotenia zameranú na to, aby estónski učitelia hľadali prístupy a metódy, ako zlepšovať tie obsahové oblasti a zručnosti, v ktorých mali žiaci pri externom hodnotení nedostatky a ako zároveň čo najefektívnejšie naďalej rozvíjať ich silné stránky. Na to slúžia hodnotiace správy z testovaní určené ako spätná väzba pre žiakov a učiteľov, kde sú okrem výsledkov z nameraných dát opísané aj oblasti, v ktorých by sa mali jednotliví žiaci zlepšiť, resp. ktoré témy alebo časti učiva by si mali zopakovať a utvrdiť. Pre učiteľov sú tieto externé správy významným pomocným nástrojom, ktorý využívajú na cielenú podporu žiakov a na skvalitnenie vyučovania v triede. Do celkového systému hodnotenia zručností a kompetencií žiakov Estónci úspešne zaraďujú aj pravidelné dotazníkové prieskumy medzi učiteľmi a žiakmi na rôzne témy, čím sa im darí komplexnejšie monitorovať širší kontext vzdelávacieho procesu a na základe spätnej väzby ho zlepšovať. Dlhodobo špičkové výsledky estónskych žiakov môžu naznačovať, že namerané dáta môžu prispievať k úprave vyučovacích metód a nastavovať postupy v zmysle tvrdenia „**čo je zviditeľnené, to môžeme meniť**“. Estónsky vzdelávací systém zostáva ako úspešný príklad dobrej praxe veľkou inšpiráciou pre Slovensko v kurikulárnej reforme vrátane nastavovania metodiky a obsahu kritériálneho testovania žiakov aj zavádzania testovania na školách s využitím počítačov.

Tretia krajina, *Portugalsko*, je zaujímavá tým, že výrazne podpriemerný výsledok v čitateľskej gramotnosti v PISA 2000 naštartoval veľmi účinné reformy vo vzdelávaní a krajina zaviedla vo vzdelávaní politiku založenú na dôkazoch a dátach, ktorá viedla k celkovému zlepšeniu vzdelávacieho systému. Pozoruhodné je najmä obdobie najvýraznejšieho zlepšenia výsledkov portugalských žiakov v čitateľskej a matematickej gramotnosti (PISA 2006 – 2009), čo odborníci dávajú do časového súvisu so zavedením povinných národných testov z matematiky a portugalského jazyka (od roku 2005), ktorých výsledky mali významný vplyv na rozhodovanie o vzdelávacej dráhe žiakov 9. ročníka (*high-stakes tests*).

Napriek hospodárskej kríze v rokoch 2008 – 2013, keď sa národný HDP Portugalska znížil o 8 % a súčasne dochádzalo k poklesu verejných výdavkov na vzdelávanie, sa Portugalsku podarilo zlepšiť svoje výsledky v PISA 2015. Národné testovanie spolu s ďalšími systémovými opatreniami je v literatúre uvádzané ako jeden z faktorov, ktoré mohli prispieť k zlepšeniu výsledkov žiakov, pričom jeho samostatný vplyv nie je možné jednoznačne oddeliť.

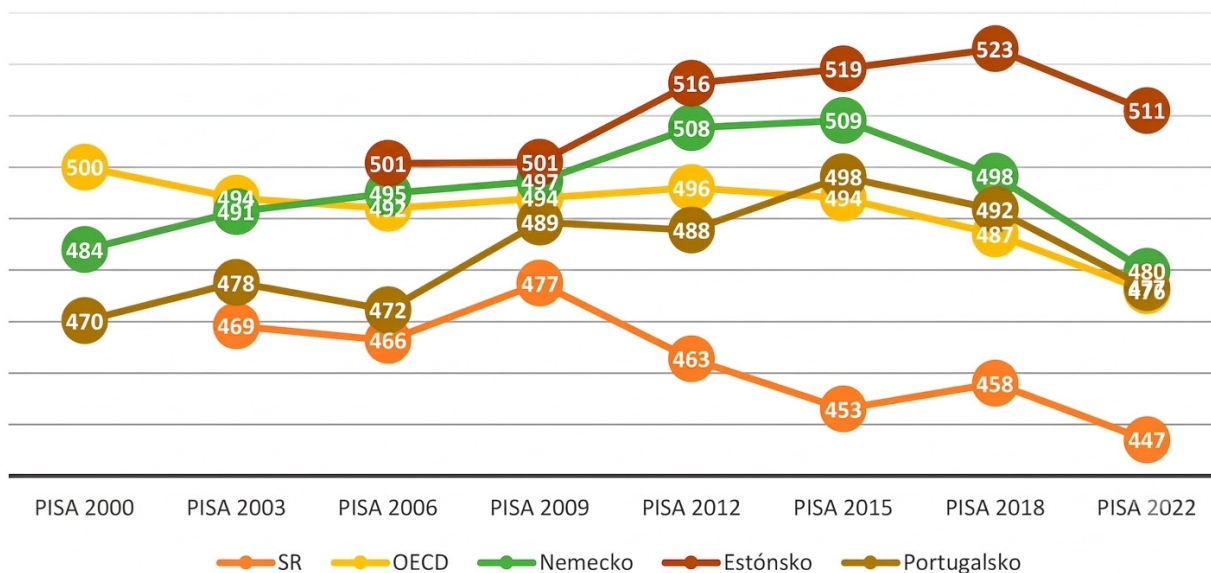
Z dlhodobejšieho hľadiska teda Portugalsko po výraznom zlepšení medzi cyklami PISA 2006 a PISA 2009 pokračovalo v pozitívnom trende aj v nasledujúcej dekáde. Podľa analýzy dlhodobých výsledkov PISA dosiahla krajina v roku 2015 historické maximum naprieč všetkými tromi testovanými doménami a prvýkrát sa dostala na priemer OECD, pričom bola považovaná za jeden z európskych úspešných príkladov zlepšovania kvality vzdelávania⁴.

Po roku 2016 však Portugalsko zrušilo niektoré zo zavedených *high-stakes* testov na základných školách a nahradilo ich diagnostickými testami, ktorých výsledky nemajú významný vplyv na rozhodovanie o ďalšom pokračovaní v štúdiu žiakov (*low-stakes tests*). Treba tiež zdôrazniť, že Portugalci popri národných testovaniach inšpirovaných PISA zaviedli veľa účinných reforiem týkajúcich sa obsahu vzdelávania, hodnotenia a vzdelávania učiteľov a učebnicovej politiky. Tieto reformy spolu s veľkým dôrazom na **zlepšovanie prístupu všetkých žiakov k vzdelávaniu** sú cenným zdrojom inšpirácie pre reformné zmeny na Slovensku.

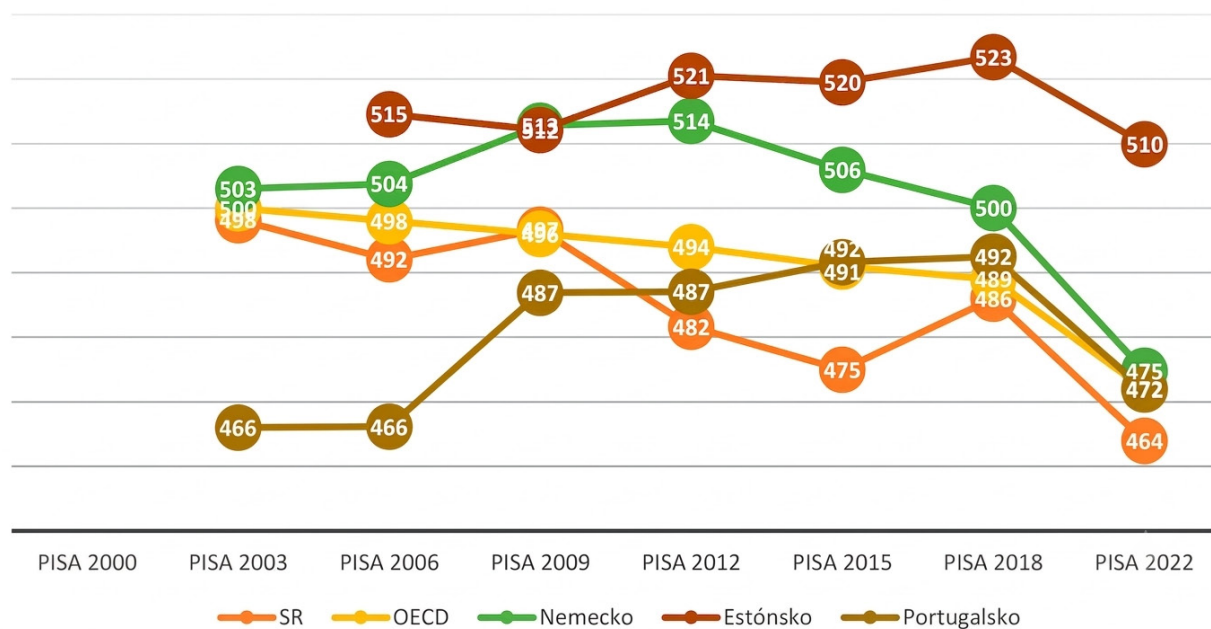
Posledné výsledky PISA 2022 však ukazujú určitý pokles výkonu portugalských žiakov oproti roku 2018, čo zodpovedá celosvetovému trendu ovplyvnenému pandémie, pričom výsledky Portugalska sa pohybujú približne na úrovni priemeru OECD (pozri obr. 1 – 3). Hoci Portugalsko už nedosahuje rekordné výsledky, jeho aktuálne skóre zostáva v európskom kontexte stredné a stabilné a krajina si udržiava výrazne lepšie výsledky než na začiatku svojich reforiem v roku 2000⁵.

⁴ NUNES, L. C., FREITAS, P. a COLAÇO, R. Trends in Student Achievement in Portugal: What Does PISA Tell Us? Porto: Fundação Belmiro de Azevedo – EDULOG, 2024. ISBN 978-989-35739-0-7. Dostupné na: <https://www.edulog.pt/storage/app/uploads/public/666/875/9da/6668759daa0d6352676543.pdf>.

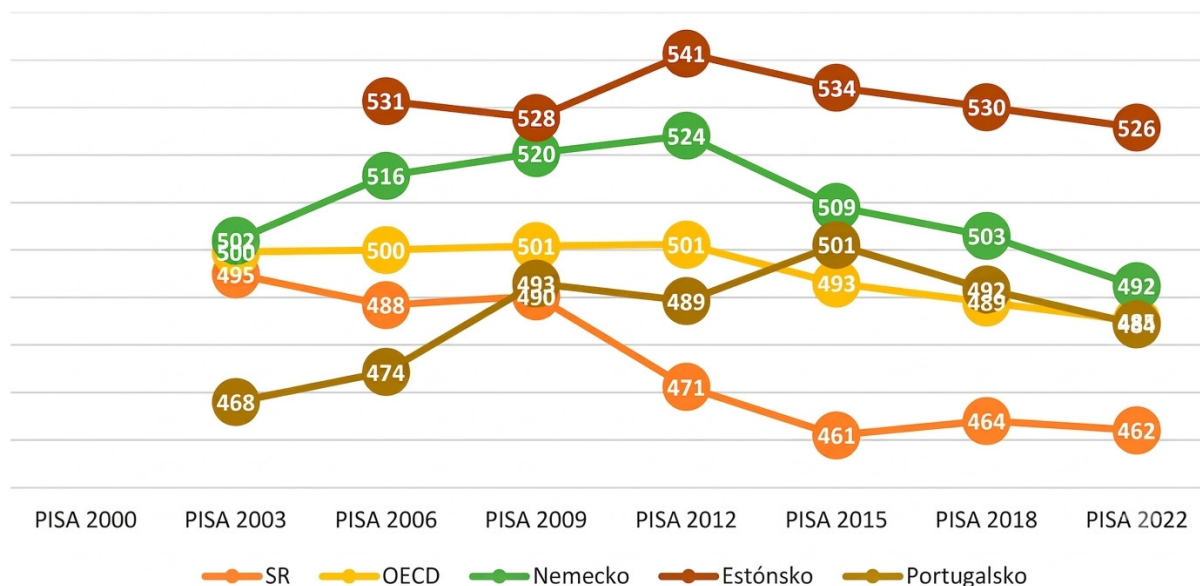
⁵ THE GLOBAL ECONOMY. Portugal: PISA reading scores [online]. Dostupné na: https://www.theglobaleconomy.com/portugal/pisa_reading_scores/



Obr. 1 Porovnanie priemerného skóre žiakov vybraných krajín a priemeru krajín OECD v štúdii PISA – čitateľská gramotnosť



Obr. 2 Porovnanie priemerného skóre žiakov vybraných krajín a priemeru krajín OECD v štúdii PISA – matematická gramotnosť



Obr. 3 Porovnanie priemerného skóre žiakov vybraných krajín a priemeru krajín OECD v štúdii PISA – prírodovedná gramotnosť

Práve inšpirácia systémami externého hodnotenia v *Nemecku* a *Estónsku* bola v predchádzajúcom období dôvodom nadviazania intenzívnej spolupráce Slovenska s odborníkmi z oboch uvedených krajín v oblasti vývoja kritériálnych testov. Nemeckí a estónski výskumníci a odborníci na štatistiku a evalváciu vzdelávania boli partnermi vtedajšieho Národného ústavu certifikovaných meraní vzdelávania (NÚCEM) pri zavádzaní metodiky kritériálneho testovania na Slovensku. NÚCEM začal v spolupráci so zahraničnými expertmi a partnermi v roku 2019 na Slovensku vyvíjať tzv. kritériálne (overovacie) testy ako vzorový model pre ich budúce zavedenie v širšom meradle na základných školách. Zahraničný projektový tím tvorili odborníci na štatistiku a metodiku meraní z Universität Stuttgart (Nemecko) a didaktickú evalváciu z inštitútu Innove v Tallinne (Estónsko). Projekt s názvom *Vyvíjanie kritériálnych testov pre primárne a nižšie sekundárne vzdelávanie na Slovensku* bol realizovaný v rámci Programu na podporu štrukturálnych reforiem EÚ v rokoch 2019 – 2021⁶. Vďaka tejto spolupráci bol zabezpečený prenos skúseností a know-how odborníkov z Nemecka a Estónska, na základe ktorého môže Národný inštitút vzdelávania a mládeže (NIVaM) v súčasnosti ďalej pokračovať v implementácii metodiky prípravy a vyhodnocovania kritériálnych testov. V roku 2022 navštívili odborníci z NIVaM *Portugalsko* so zámerom inšpirovať sa jeho skúsenosťami pri zavádzaní kurikulárnej reformy, ako aj hodnotení kvality vzdelávania.

V nasledujúcich podkapitolách budú všetky vzdelávacie systémy týchto troch krajín – Nemecka, Estónska a Portugalska – detailne analyzované s akcentom na zavedené systémy hodnotenia, ktoré sú inšpiráciou aj pre Slovensko.

1.2 NEMECKO – systém VERA a dôsledky PISA šoku

V Spolkovej republike Nemecko je zodpovednosť za vzdelávací systém rozdelená medzi federálnu úroveň a spolkové krajiny⁷. Rozsah zodpovedností spolkovej vlády v oblasti vzdelávania je definovaný v základnom zákone (*Grundgesetz*). Pokiaľ základný zákon neudeľuje zákonodarné právomoci federácii, majú spolkové

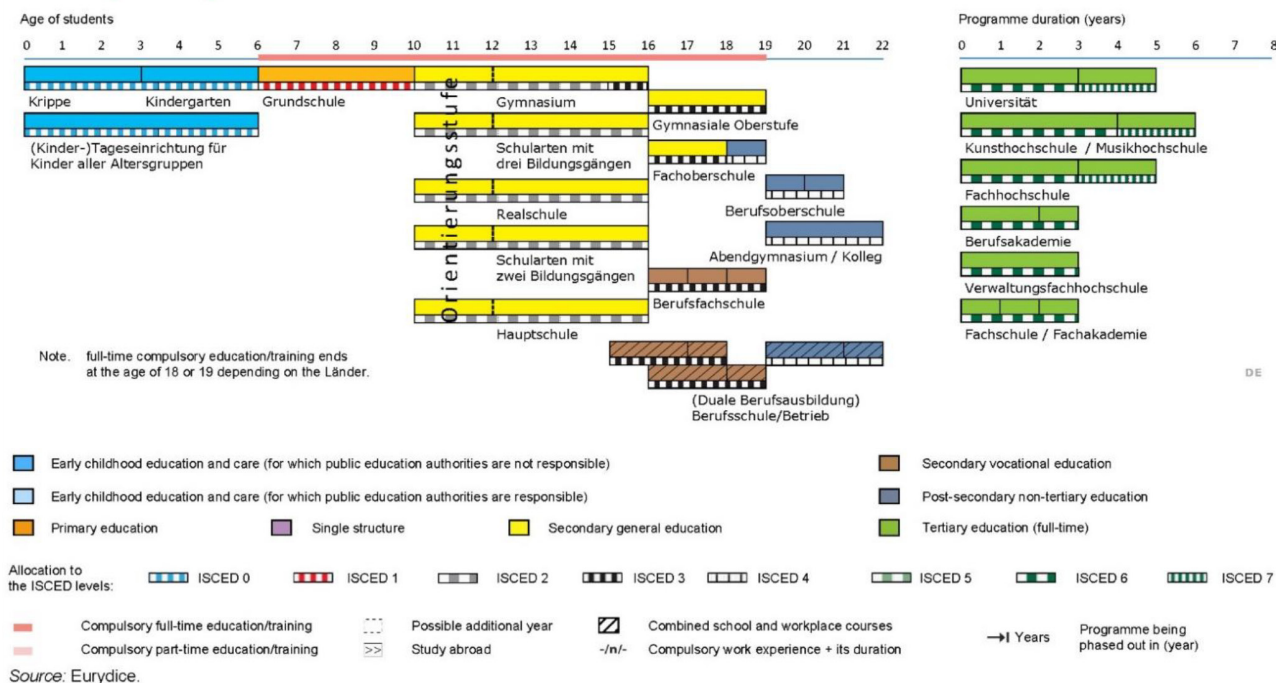
⁶ NÁRODNÝ ÚSTAV CERTIFIKOVANÝCH MERANÍ VZDELÁVANIA (NÚCEM). Kritériálne testy (Program SRSP): Vyvíjanie kritériálnych testov pre primárne a nižšie sekundárne vzdelávanie na Slovensku [online]. Bratislava: NÚCEM. Dostupné na: <https://www2.nucem.sk/sk/projekty-a-linky/kriterialne-testy>.

⁷ EURYDICE. Germany: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/euryperia/germany/overview>.

krajiny právo prijímať vlastné zákony. Spravovanie vzdelávacieho systému je takmer výlučne záležitosťou spolkových krajín. Jednotnými pre všetky spolkové krajiny sú dĺžka povinnej školskej dochádzky, zabezpečovanie nadväznosti vzdelávacích stupňov, označovanie vzdelávacích zariadení, vzájomné uznávanie vysvedčení (certifikátov), ako aj systém známkovania a hodnotenia žiakov.

Nemecký vzdelávací systém

Germany – 2023/2024



Obr. 4 Štruktúra vzdelávacieho systému v Nemecku⁸

Primárne vzdelávanie

Povinnosť nastúpiť do základnej školy v Nemecku spravidla vzniká v kalendárnom roku, v ktorom dieťa dovŕši šiesty rok veku. Základná škola nadväzuje na starostlivosť a vzdelávanie v materských školách poskytované deťom od troch rokov veku. Všetci žiaci v Nemecku nastupujú na *Grundschule* (základnú školu), ktorá takmer vo všetkých spolkových krajinách pozostáva z 1. až 4. ročníka (v Berlíne a Brandenbursku z 1. až 6. ročníka).

Vzdelávanie na základnej škole je orientované na rozvoj a nadobúdanie kompetencií a vyžaduje si zodpovedajúce formy hodnotenia žiakov. Základom hodnotenia je spätná väzba od učiteľov orientovaná na úroveň nadobudnutých kompetencií získaných v priebehu procesu učenia sa.

Žiakom aj ich rodičom učitelia pravidelne poskytujú správy o pokroku, ktoré obsahujú informácie o napredovaní žiakov v učení ako aj o následných vzdelávacích krokoch. Táto spätná väzba sa uskutočňuje podľa transparentných kritérií, pričom zaznamenáva individuálny pokrok žiakov a úroveň dosahovaných kompetencií podľa vzdelávacích štandardov. Učitelia oboznamujú žiakov s nástrojmi na sebahodnotenie a povzbudzujú ich, aby uvažovali o svojich vzdelávacích dráhach a výsledkoch spôsobom primeraným ich veku a vývoju. Postupne ich posilňujú v kompetencii sebahodnotenia umožňujú im stanovovať si vlastné ciele a vnímať externé hodnotenie ako príležitosť na vlastné učenie sa. Hodnotenie

⁸ EURYDICE. Germany: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurylopedia/germany/overview>.

v nemeckých školách sa zameriava na individuálne **povzbudenie každého žiaka, aby dosiahol všetko, čoho je schopný**. Na to je potrebné, aby učitelia neustále sledovali individuálny vývoj a výkon každého žiaka, ako aj jeho pracovné a sociálne správanie a tieto faktory komplexne posudzovali.

Pokrok vo vzdelávaní sa zvyčajne skúma neustálym monitorovaním procesov učenia a používaním ústnej a písomnej kontroly. V 1. a 2. ročníku sa pozornosť sústreďuje na priame pozorovanie žiakov. Najneskôr v 3. ročníku sa žiaci začínajú stretávať aj s písomnými testami z niektorých predmetov (najmä z nemeckého jazyka a matematiky). Spravidla od 3. ročníka je každý žiak zaradený do vhodného ročníka v závislosti od svojho prospechu a výsledkov, a to buď postupom do vyššieho ročníka alebo opakovaním ročníka. Vek 10 rokov sa stáva pre žiakov kľúčovým medzníkom pre voľbu ich ďalšieho vzdelávania.

Nižšie sekundárne vzdelávanie

V Nemecku sú žiaci približne vo veku 10 rokov na základe svojich študijných výsledkov a odporúčania základnej školy zaraďovaní do rôznych vetiev nižšieho sekundárneho vzdelávania, predovšetkým do *Hauptschule*, *Realschule* alebo *Gymnasia*. Tieto vzdelávacie dráhy sa odlišujú dĺžkou štúdia aj typom ukončovacích kvalifikácií. *Hauptschule* sa spravidla končí v 9. ročníku (v niektorých spolkových krajinách v 10. ročníku), pričom žiaci získavajú kvalifikáciu *Hauptschulabschluss*. *Realschule* trvá do 10. ročníka a jej absolvovaním žiaci získavajú *Realschulabschluss (Mittlere Reife)*. *Gymnasium* pokračuje až do 12. alebo 13. ročníka v závislosti od spolkovej krajiny a ukončené je skúškou *Abitur*, ktorá umožňuje vstup na vysokoškolské štúdium.⁹

Prvý ukončovací certifikát môžu žiaci získať po absolvovaní 9. ročníka na *Hauptschule*, pričom v niektorých spolkových krajinách je jeho udelenie podmienené aj záverečnou skúškou. Na rozdiel od toho žiaci v *Realschule* získavajú prvý formálny kvalifikačný stupeň až po 10. ročníku.

Vyššie sekundárne vzdelávanie

Keď žiaci skončia povinnú školskú dochádzku – spravidla pri dosiahnutí veku 15 rokov – prechádzajú na vyššie sekundárne vzdelávanie. Typ školy závisí od dosiahnutých výsledkov žiakov na konci nižšieho sekundárneho vzdelávania. Ponuka kurzov zahŕňa denné všeobecné vzdelávanie a odborné školy, ako aj odborné vzdelávanie a prípravu v rámci duálneho systému. Po úspešnom ukončení vyššieho sekundárneho vzdelávania môžu žiaci pokračovať vo vysokoškolskom vzdelávaní.

Externé hodnotenie žiakov

V nasledujúcej časti budú podrobnejšie analyzované nástroje centrálneho externého hodnotenia žiakov na primárnom a nižšom sekundárnom stupni vzdelávania na celoštátnej úrovni. Vedomostné testy boli v Nemecku zavedené v roku 2006 ako súčasť systému externého hodnotenia škôl¹⁰. Nemecké spolkové krajiny používajú testy označované ako VERA, ktoré sú v súčasnosti zavedené v 3. a 8. ročníku základnej školy. Cieľom zavedenia testov VERA je podporiť rozvoj škôl prostredníctvom poskytovania spätnej väzby, ktorá môže prispieť k zlepšovaniu vzdelávacích výsledkov. Vďaka vyhodnoteniu výsledkov testov VERA môžu učitelia získať lepší prehľad o nedostatkoch vo vedomostiach žiakov vo svojich triedach. Na základe týchto poznatkov

⁹ GOETHE-INSTITUT. School system – Mein Weg nach Deutschland [online]. 2026. Dostupné na: <https://www.goethe.de/prj/mwd/en/idl/ssy.html>.

¹⁰ KULTUSMINISTERKONFERENZ. Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring [Strategic Approach of the Standing Conference of the Ministers of Education and Cultural Affairs on Educational Monitoring] [online]. Berlin: Wolters Kluwer Deutschland GmbH, 2016. ISBN 9783556071069. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf.

môžu následne upravovať vlastné metódy vyučovania, napr. s väčším zameraním na rozvoj tých kompetencií a dôrazom na tie obsahové oblasti, ktoré ich žiaci v testoch VERA uplatnili, resp. riešili nedostatočne. Potrebne je poznamenať, že prípadné neuspokojivé výsledky v testoch nemajú negatívny dopad na hodnotenie nemeckých učiteľov.

VERA je skrátenou podobou názvu *Vergleichsarbeiten*, čo označuje štandardizované písomné porovnávacie testy, ktoré žiaci v Nemecku absolvujú v 3. a 8. ročníku (VERA-3 a VERA-8). Tieto testy sú realizované koordinovane vo všetkých spolkových krajinách na základe spoločného rámca a ich cieľom je zistiť, akú úroveň kompetencií žiaci dosiahli v určitých obdobiach ich vzdelávacej dráhy v súlade s federálne záväznými vzdelávacími štandardmi. VERA-3 a VERA-8 sa v súčasnosti povinne realizujú vo všetkých 16 spolkových krajinách, pričom testovanie prebieha každoročne vo všetkých triedach 3. a 8. ročníka škôl poskytujúcich všeobecné vzdelávanie.

Nemecké spolkové krajiny sa dohodli, že VERA-3 (realizovaná v 3. ročníku) bude zameraná na povinné testovanie z nemeckého jazyka alebo z matematiky. Ak je v danom roku vybraný v rámci celoštátneho testovania nemecký jazyk, tento test musí povinne overovať aspoň čitateľské kompetencie žiakov. Ďalšie oblasti, ktoré môže test VERA v nemeckom jazyku obsahovať, sú počúvanie, pravopis, jazyk a používanie jazyka. V budúcnosti je plánované zahrnúť do testovania aj písanie ako voliteľnú oblasť. V prípade matematiky sa testy VERA-3 zameriavajú výberovo vždy na dve z piatich obsahových oblastí (napr. čísla a operácie a údaje, frekvencia/početnosť a pravdepodobnosť). Spolkové krajiny a inštitút zodpovedný za testovanie vyberajú tieto dve oblasti pred každým kolom testovania.

Všetky spolkové krajiny zapojené do projektu VERA musia testovať aspoň z jedného predmetu aj žiakov 8. ročníka. Ak je vybraným predmetom nemecký jazyk, žiaci musia byť testovaní minimálne z čitateľských kompetencií. Zavedená je aj voliteľná možnosť testovať počúvanie, pravopis alebo jazyk a používanie jazyka.

Podobne ako v prípade testovania na prvom stupni vzdelávania, majú v budúcnosti testy VERA-8 obsahovať aj test z písania. Testy VERA-8 z matematiky zahŕňajú všetkých päť obsahových oblastí (so zameraním na kľúčové kompetencie). Pokiaľ ide o prvý cudzí jazyk (anglický alebo francúzsky), žiaci sú testovaní minimálne v oblastiach čítania a počúvania. Krajiny môžu školám umožniť, aby v testoch VERA-3 aj VERA-8 testovali viac ako jeden predmet alebo obsahovú oblasť.

Realizácia porovnávacích testov vo všetkých spolkových krajinách Nemeckej republiky je súčasťou Komplexnej stratégie Stálej konferencie pre monitorovanie vzdelávania (KMK)¹¹, ktorá bola prijatá v roku 2006. Cieľom monitorovania vzdelávania bolo dosiahnuť, aby sa vzdelávací systém viac zameriaval na kompetencie žiakov. Namiesto toho, aby systém požadoval, aký obsah by sa mal v danom predmete vyučovať, požaduje kompetencie, ktoré by mali žiaci v danom predmete dosiahnuť v určitom období ich vzdelávacej dráhy.

Z pohľadu žiakov predstavuje testovanie VERA nástroj včasnej diagnostiky, ktorým sa identifikujú medzery v úrovni kompetencií vo vzťahu k definovanému štandardu. Výsledky VERA-8 poskytujú žiakovi spätnú väzbu dva roky pred dosiahnutím vzdelávacích štandardov v 10. ročníku.

Pre učiteľov sú výsledky testov zasa spätnou väzbou o účinnosti ich pedagogickej práce. Sú aj jedným z nástrojov podporujúcich ich diagnostické pedagogické zručnosti. Kľúčovou je však podpora výučby, ktorá je zameraná na rozvoj a nadobúdanie kompetencií žiakov a následná dodatočná a dostatočná podpora žiakov v prípade, že dosahujú slabšie výsledky.

¹¹ KULTUSMINISTERKONFERENZ. 2006. *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. [online]. Dostupné na: <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring.html>.

Stála konferencia pre monitorovanie vzdelávania v dohode o ďalšom vývoji VERA testov¹² zdôraznila, že hlavnou úlohou porovnávacích testov je podpora vyučovania a rozvoj jednotlivých škôl. VERA má zároveň uľahčovať implementáciu predmetových a didaktických koncepcií vyjadrených vo vzdelávacích štandardoch. V dohode sa zdôrazňuje, že testy VERA nie sú vhodné na známkovanie žiakov ani na hodnotenie úspešnosti škôl a že výsledky jednotlivých škôl sa nesmú zverejňovať s cieľom vzájomného porovnávania.

Hoci sú porovnávacie testy VERA súčasťou systému externého hodnotenia v tom zmysle, že predstavujú *externý, štandardizovaný nástroj monitorovania výsledkov vzdelávania*, ich funkcia je striktne diagnostická. Výsledky poskytujú školám spätnú väzbu na účely pedagogického plánovania a zlepšovania výučby, no neslúžia ako podklad na hodnotenie alebo klasifikáciu škôl zvonka. Sú sprístupňované výhradne škole a jej pedagogickému kolektívu a nepoužívajú sa na porovnávanie škôl ani na zverejňovanie rebríčkov.

Stála konferencia pre monitorovanie vzdelávania predstavila v roku 2010 koncepciu využívania vzdelávacích štandardov ako nástroja systematického rozvoja vyučovania¹³. Dokument zdôrazňuje, že **spätná väzba z porovnávacích testov má predstavovať jeden z kľúčových prvkov cyklu zlepšovania škôl založeného na dátach**. Účinná kultúra spätnej väzby má pritom fungovať ako prepojenie medzi samotným zberom a vykazovaním údajov a ich následným pedagogickým využitím, ktoré umožňuje školám cielene upravovať výučbu a zefektívňovať proces učenia sa žiakov.

Inštitucionálne zabezpečenie testovania

Inštitút pre rozvoj kvality vo vzdelávaní (IQB)¹⁴ je vedecká inštitúcia pridružená k Humboldtovej univerzite v Berlíne a poverená úlohami KMK. Inštitút bol založený v roku 2004 a je financovaný z rozpočtov 16 nemeckých spolkových krajín. Zamestnáva interdisciplinárny tím odborníkov, ktorí úzko spolupracujú. Tím tvoria psychológovia, výskumníci vo vzdelávaní, učitelia a psychometrici. Odborníci IQB pripravujú predmetovo-didaktické expertízy a zabezpečujú vývoj testov VERA-3 a VERA-8.

Zodpovednosť za testy VERA je jasne rozdelená medzi IQB na jednej strane a štátne ústavy, agentúry pre kvalitu a príslušné ministerské úrady na strane druhej (pozri tab. 1 nižšie).

Tab. 1 Rozdelenie zodpovednosti za testovanie VERA v Nemecku

Rozdelenie zodpovednosti	
IQB	Spolkové krajiny
• Tvorba a vývoj položiek • Pilotovanie položiek • Určenie obťažnosti položiek (škálovanie) • Zostavovanie testových zošitov • Vypracovanie sprievodných výučbových materiálov • Návrh spätnej väzby	• Tlač a distribúcia testových zošitov • Administrácia testov • Hodnotenie testov, spracovanie dát • Štatistická analýza • Spätná väzba • Vykonávanie ďalších opatrení na podporu škôl po získaní spätnej väzby o výsledkoch

¹² KULTUSMINISTERKONFERENZ. Vereinbarung zur Weiterentwicklung der Vergleichsarbeiten (VERA): Beschluss der Kultusministerkonferenz vom 08.03.2012 i. d. F. vom 15.03.2018 [online]. Berlin: Kultusministerkonferenz, 2018. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_03_08_Weiterentwicklung-VERA.pdf.

¹³ SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND. Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung [online]. Berlin: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2010. ISBN 978-3-556-02609-0. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_00_00-Konzeption-Bildungsstandards.pdf.

¹⁴ <https://www.iqb.hu-berlin.de/de/>.

Spolkové krajiny sú zodpovedné za organizáciu a realizáciu testovania VERA, najmä za administratívne a logistické zabezpečenie testov, priebeh testovania na školách, spracovanie výsledkov na úrovni krajiny a za poskytovanie spätnej väzby školám. Samotná tvorba testových položiek, zostavovanie testových zošitov a odborná príprava testových materiálov však patrí do kompetencie IQB, ktorý tieto materiály vytvára centrálne pre všetky spolkové krajiny. Jednotlivé krajiny tak môžu upraviť najmä organizačné postupy a formy spätnej väzby, nie však obsah ani konštrukciu samotných testových úloh.

Tvorba a vývoj testových položiek sa realizuje na národnej úrovni pod vedením IQB. Vyškolení učitelia, ktorých nominuje príslušná nemecká spolková krajina, vytvárajú testové položky. Univerzitní odborníci na didaktiku predmetov ich preskúmajú a zhodnotia a následne, pred ich použitím na celoštátnej úrovni, ich akademickí odborníci na testovanie z IQB podrobia empirickým analýzám na vzorke žiakov, aby posúdili ich vhodnosť a náročnosť. Skutočné testy VERA obsahujú len tie položky, ktoré boli úspešne empiricky overené. Po didaktických a štatistických analýzach položiek ich IQB zostaví do testových zošitov.

Vývoj kritériálnych testov

Základom vývoja testových položiek sú vzdelávacie štandardy¹⁵, ktoré sú záväznými kritériami vo všetkých nemeckých školách. Vytvárajú ich spoločne odborníci v didaktike, skupiny učiteľov a štátna správa. Vzdelávacie štandardy sú orientované na výstupy a opisujú kompetencie špecifické pre daný predmet, ktorými by mali žiaci disponovať na konci určitej etapy ich vzdelávacej dráhy (tzv. normatívne štandardy). Pod kompetenciou sa pritom rozumie schopnosť aplikovať pri riešení problémov nadobudnuté vedomosti a zručnosti z príslušných predmetov.

V prvom štádiu tvorby testov je nevyhnutná operacionalizácia vzdelávacích štandardov. Jej vhodným základom je to, že vzdelávacie štandardy v Nemecku sú detailne rozpracované a špecifikované s ohľadom na didaktický rozmer daného predmetu a psychologické aspekty danej vekovej kategórie žiakov. Ďalším krokom je vypracovanie špecifikácií testov s detailným opisom toho, aké bude ich zameranie a čo budú testy overovať. Špecifikácie testov sú podkladom na vývoj testových úloh zo strany vyškolených učiteľov. Na tvorbu úloh zo strany učiteľov nadväzuje ich optimalizácia; tá sa uskutočňuje tzv. predpilótnymi štúdiami.

Predpilótné štúdium sa realizujú na malej vzorke (približne 50 – 100) žiackych odpovedí pripadajúcich na úlohu. Následne sa výsledky štatisticky spracovávajú a zisťujú sa frekvencie riešení úloh. Odpovede na dané úlohy sa kontrolujú a autorom úloh je poskytnutá spätná väzba. Cieľom predpilótných empirických štúdií je odpovedať na otázky, či úlohy fungujú primerane a či sú zrozumiteľné pre cieľovú populáciu. Nasleduje revidovanie úloh, ktoré neboli v predpilótnej štúdii úspešné, a doplnenie alebo prepracovanie kľúčov správnych odpovedí. Výsledky týchto predpilótných štúdií sú poskytnuté didaktikom a psychometrikom na analýzu a následnú optimalizáciu učiteľmi (vyškolení autori úloh).

Na výsledky predpilótnej štúdie nadväzuje **pilótna štúdia na kontrolu kvality úloh**. Realizuje sa na väčšom počte žiakov (veľkosť vzorky do cca 5 000 žiakov). Vzorka zodpovedá heterogénnej populácii žiakov, pričom nejde o prísne reprezentatívnu vzorku v úzkom význame slova. Počet žiakov sa stanovuje tak, aby počet odpovedí na položku bol postačujúci na zrealizovanie štatistickej analýzy.

¹⁵ INSTITUTE FOR EDUCATIONAL QUALITY IMPROVEMENT (IQB). Educational Standards. Institute for Educational Quality Improvement (IQB) [online]. Dostupné na: <https://www.iqb.hu-berlin.de/en/bista>.

Dizajn pilotného testu je nastavený tak, že úlohy sú nakombinované do viacerých súborov (tzv. setov), a tie do rôznych testových zošitov. Každý žiak odpovedá iba na istú časť úloh, ktoré sú zahrnuté v jeho testovom zošite. Žiaci neriešia tie isté úlohy, testové zošity sú odlišné a sú zostavené na základe rotujúcej kombinácie modulov. Z hľadiska vyhodnotenia pilotných testov je kľúčové, aby bolo zabezpečených aspoň 350 odpovedí na každú úlohu.

Po pilotnej štúdii nasleduje tzv. **štandardizačná štúdia** na stanovenie stupníc kompetencií. Tá sa už realizuje na veľkej reprezentatívnej vzorke žiakov (N = 10 000). Vyberajú sa do nej iba vysokokvalitné úlohy na základe svojich vyhovujúcich psychometrických vlastností.

Vývoj modelov úrovni spôsobilostí

Pre každé vyhodnotenie testu v štandardizačnej štúdii je dôležité vytvoriť príslušný model úrovne znalostí. Model popisuje stanovenie štandardov vrátane definovania hraničných skóre (Cut Score), ktoré oddeľujú určité úrovne znalostí. **Každá úroveň obsahuje opis kompetencií, ktorými disponujú testovaní žiaci s dosiahnutým skóre v danom rozmedzí.** Dáta zo štandardizačnej štúdie sú následne analyzované na základe veľkej skupiny položiek zohľadňujúc obťažnosť položiek na vzorke testovaných žiakov. Obťažnosť položky priamo súvisí s úrovňou schopnosti testovaných žiakov, keďže sú vyjadrené na rovnakej škále.

Štandardizačný proces je nastavený tak, že odborníci pracujú individuálne a v skupinách na dosiahnutí **konsenzu o hraničných skóre**. Odborníci pracujú s položkami, ktoré sú usporiadané v poradí podľa ich empirickej obťažnosti a na ich základe vytvárajú modely úrovni znalostí.

Tab. 2 Definovanie výkonových štandardov v Nemecku¹⁶

Typ štandardu	Čo vyjadruje	Ako sa určuje	Koho charakterizuje	Účel
Očakávaný štandard (Regelstandard)	Očakávaná úroveň kompetencií, ktorú má dosiahnuť „bežný“ žiak	Ad-hoc v pedagogických dokumentoch	Priemerných žiakov	Stanovuje normu – očakávaný výkon pre väčšinu
Minimálny štandard (Mindeststandard)	Najnižšia prijateľná úroveň kompetencií potrebná pre ďalšie vzdelávanie	Post-hoc na základe empirických dát	Žiakov s najnižším výkonom	Identifikácia žiakov, ktorí potrebujú podporu
Optimálny štandard (Maximal-/Exzellenzstandard)	Úroveň kompetencií dosahovaná najlepšími žiakmi pri priaznivých podmienkach učenia	Post-hoc podľa výkonov najúspešnejších žiakov	Žiakov s vysokým výkonom	Ukazuje, čo je dosiahnuteľné pri excelentnom učení

¹⁶ Institut zur Qualitätsentwicklung im Bildungswesen (IQB). *Kompetenzstufenmodelle zu den Bildungsstandards*. [online]. Berlin. Dostupné na: <https://www.iqb.hu-berlin.de/de/bista/entwicklung/kompetenzstufenmodelle/>.

Národné hodnotiace štúdie a sledovanie trendov

IQB pravidelne od roku 2009 skúma a zverejňuje správy o tom, či a do akej miery žiaci dosahujú vzdelávacie ciele obsiahnuté vo vzdelávacích štandardoch. Národné hodnotiace štúdie sú súčasťou komplexného systému hodnotenia vzdelávania v Nemecku, ktorého cieľom je spolu s ďalšími opatreniami podporovať výučbu orientovanú na rozvoj a nadobúdanie kompetencií a zabezpečovať kvalitu vzdelávania. Výsledky národnej hodnotiacej štúdie sa analyzujú na úrovni školského systému každej spolkovej krajiny. Zo zverejnených výsledkov nie je možné vyvodzovať závery o výsledkoch jednotlivých škôl, tried alebo žiakov. Výsledky sa uvádzajú na úrovni 16 spolkových krajín Nemecka a poskytujú krajinám spätnú väzbu o úrovni vedomostí, ktoré žiaci v ich školách dosiahli v testovaných predmetoch.

Od roku 2015 výsledky kontinuálneho monitoringu umožňujú aj vyhodnocovanie trendov na školách v rámci celého Nemecka. IQB periodicky realizuje štúdie IQB-Bildungstrend, ktoré sledujú dosahovanie vzdelávacích štandardov a umožňujú porovnávať vývoj výsledkov v čase (tzv. *IQB Trendy vo výsledkoch žiakov*) so zameraním na príslušnú cieľovú skupinu žiakov. Uvedené trendové štúdie sa realizujú na reprezentatívnych vzorkách škôl v jednotlivých spolkových krajinách z vybraných vyučovacích predmetov – nemecký jazyk a matematika (4. ročník ZŠ), nemecký, anglický, francúzsky jazyk, matematika a prírodné vedy (9. ročník ZŠ). Pre vybrané školy je účasť v testovaní povinná. Testovanie z príslušných predmetov v danom ročníku prebieha vo vopred stanovenom testovacom okne (spravidla napr. apríl – jún), žiaci riešia testy orientované na vzdelávacie štandardy a vyplňajú dotazníky. Okrem toho sú predmetom výskumu aj rodičovské a učiteľské dotazníky. Štúdia je metodologicky aj dizajnovo inšpirovaná výskumom OECD PISA. Napr. v trendovej štúdii v roku 2024 žiaci riešili v deň prieskumu najprv rôzne úlohy z matematiky a/alebo prírodných vied a niekoľko úloh na logické uvažovanie a jazykové zručnosti, potom vyplňali dotazník.

Od roku 2015 poskytujú výsledky IQB-Bildungstrend prehľad o dlhodobých vývojových trendoch vo výsledkoch žiakov¹⁷ v dosahovaní vzdelávacích štandardov. Testovanie prebieha na jar daného roka, avšak vzhľadom na rozsah zberu dát a komplexnú štatistickú analýzu sa výsledky zverejňujú až na jeseň nasledujúceho kalendárneho roka.

Výsledky každej spolkovej krajiny sú prezentované v porovnaní s ostatnými spolkovými krajinami a priemerom Nemecka ako celku. Keďže ide o kritériálne orientované testy, výsledky každej krajiny obsahujú porovnávacie dáta o dosiahnutom vzdelávacom štandarde, t. j. aké percento žiakov sa nachádza pod minimálnym štandardom, aké percento dosahuje minimálny štandard, a aké percento žiakov presahuje minimálny štandard, resp. dosahuje optimálny štandard. Prezentovanie výsledkov týchto trendových meraní sa orientuje na zdôrazňovanie silných a slabých stránok vzdelávania v spolkových krajinách. Pre slovenský vzdelávací systém je tiež veľmi inšpiratívne, že súčasťou testovania sú aj doplnkové štúdie o faktoroch dôležitých pre zlepšenie kvality vzdelávania (napr. potreba vzdelávania učiteľov).

¹⁷ BUSSE, J., BECKER, B., WEIRICH, S. a SCHIPOLOWSKI, S. IQB Trends in Student Achievement 2018. Journal of Open Psychology Data [online]. 2022, roč. 10, č. 1, čl. 18. Dostupné na: <https://doi.org/10.5334/jopd.71>.

TRENDY výsledkov žiakov na primárnom stupni vzdelávania

Testovanie sa uskutočňuje na vzorkách škôl každej spolkovej krajiny na konci 4. ročníka z predmetov nemecký jazyk a matematika. Testovanie na účel sledovania trendov sa realizuje a vyhodnocuje cyklicky každých 5 rokov.

TRENDY výsledkov žiakov na nižšom sekundárnom stupni vzdelávania

Testovanie sa uskutočňuje na vzorkách škôl každej spolkovej krajiny na konci 9. ročníka. Testovanie na účel sledovania trendov v nemeckom, anglickom a francúzskom jazyku sa vyhodnocuje každých 6 rokov. Podobne sa každých 6 rokov vyhodnocujú trendy v matematike a prírodných vedách (biológia, chémia, fyzika). Na vyhodnocovanie trendov sa, podobne ako v štúdiu PISA, využívajú kotviace položky, teda také, ktoré sú do testov zaraďované v rôznych rokoch opakovane a majú požadované kvalitatívne aj psychometrické parametre.

Dotazníkové prieskumy

Ako bolo uvedené vyššie, testovanie kompetencií v Nemecku je sprevádzané dotazníkovými prieskumami, ktoré mapujú kontext a dôležité faktory vzdelávania. Sú určené pre žiakov, rodičov, učiteľov a riaditeľov škôl. Prostredníctvom dotazníkov sa, obdobne ako v meraní PISA, zbierajú dáta o socio-demografických premenných a o podmienkach učenia sa žiakov v škole a doma, respektíve po škole.

Niekoľko príkladov tém žiackych dotazníkov:

- Sociálne zázemie a prisťahovalecké pozadie
- Aspekty kvality výučby (napr. kognitívna aktivácia)
- Využívanie školskej a mimoškolskej doplnkovej ponuky (napr. súkromné hodiny doučovania)
- Hodnotenie vlastnej motivácie (napr. záujmy súvisiace s predmetom)
- Aspekty vyučovania v škole a medzipredmetové zručnosti (napr. stratégie učenia)
- Očakávania žiaka týkajúce sa jeho kariéry
- Používanie digitálnych médií

Niekoľko príkladov tém dotazníkov pre učiteľov a riaditeľov:

- Dosiahnuté vzdelanie a ďalšie absolvované vzdelávania
- Prístup k inklúzii žiakov so ŠVVP
- Dodatočné školské aktivity (napr. doučovanie)

Testovanie s využitím počítačov – výhody a nevýhody

V posledných rokoch sa niekoľko nemeckých spolkových krajín rozhodlo zmeniť spôsob testovania VERA z testovania v papierovej forme na testovanie pomocou počítača. Elektronické testovanie znižuje náklady na tlač a administráciu testov, zvyšuje efektívnosť testovania a vyhodnocovania a v prípade položiek s výberom odpovede generuje okamžité výsledky (Ejim, 2017)¹⁸. Ak by praktická príprava a realizácia testov VERA zjednodušovala prácu učiteľov, mohli by byť testy viac učiteľmi akceptované. Učitelia by mali viac času a mohli by byť viac motivovaní venovať sa koncepčne relevantným činnostiam, ako je napríklad reflektovanie výsledkov testov svojich žiakov. Podľa sekvenčného modelu rozvoja vyučovacích hodín, ktorý navrhol

¹⁸ EJIM, S. An Overview of Computer Based Test [online]. 2017. Dostupné na: <https://doi.org/10.13140/RG.2.2.32040.88326>.

Helmke (2012)¹⁹, je predpokladom priaznivých zmien vo vyučovacích postupoch lepšia akceptácia výkonových testov a hlbšia reflexia výsledkov testov.

Zavedenie testovania s využitím počítača si však vyžaduje, aby výsledky boli porovnateľné s výsledkami získanými pri testovaní v papierovej forme, aby žiaci neboli týmto novým testovacím prostriedkom znevýhodnení. Táto otázka je obzvlášť dôležitá, keď sa majú výsledky výkonových testov porovnávať v čase. V prípade Nemecka, kde sa testovanie na počítači zaviedlo len čiastočne, musí byť zabezpečené aj porovnávanie skupín výsledkov žiakov, ktorí sú testovaní rôznymi spôsobmi (Bennett a kol., 2008)²⁰.

Jednou z praktických **výhod** počítačového testovania je, že už nie je potrebné tlačiť a uchovávať testové zošity (Thompson a Weiss, 2009)²¹. Využitie počítača môže za určitých podmienok umožniť meranie niektorých konštruktov s vyššou mierou šandardizácie. Ak inštrukcie a úlohy prečíta účastníkom počítačový systém, ako je to potrebné v prípade detí predškolského veku, vedie to k vyššiemu stupňu šandardizácie v porovnaní s tým, keď inštrukcie a úlohy čítajú učitelia, každý mierne odlišným – subjektívnym spôsobom. Automatizované skórovanie odpovedí na testové úlohy počítačovým systémom je tiež objektívnejšie ako hodnotenie rôznymi učiteľmi. Pomocou počítačového testovania možno merať viac konštruktov ako pri testovaní na papieri. To platí napríklad pre meranie multimediálnych vzdelávacích kompetencií, keďže do testov možno integrovať videá a animácie. Možné je aj dodatočné zhromažďovanie priebežných údajov, ako sú časy odpovedí a pohyby myšou, ako aj adaptívne testovanie podľa schopností žiakov (Csapó a kol., 2014)²².

Nevýhodou testovania s využitím počítača je, že počiatočné investície do testovacích prostredí sú veľmi vysoké. Vývoj testov je časovo náročný a vyžaduje starostlivé plánovanie. Niektoré formáty položiek, ktoré sa používajú pri papierovom testovaní, nie je možné z technických dôvodov implementovať do testovania s využitím počítača. Porovnateľnosť medzi týmito dvoma spôsobmi testovania tak môže byť v niektorých prípadoch znížená. Počas online testovania sa môžu vyskytnúť problémy s internetovým pripojením. Počítačová gramotnosť účastníkov môže mať tiež vplyv na ich výkon v teste (Kyllonen, 2009)²³.

Testovanie s využitím počítača a akademické výsledky

Základnou podmienkou zavedenia počítačového testovania VERA bolo, aby sa žiaci najprv prispôbili novému testovaciemu prostrediu a jeho požiadavkám. Interakcia s testovacím prostredím môže vyvolať dodatočnú kognitívnu záťaž a prinajmenšom v prvej fáze testovania môže znížiť kapacitu pracovnej pamäte

¹⁹ HELMKE, A. Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts [Quality of Classroom Practices and Teachers' Professionalism. Diagnosis, Evaluation and Improvement of Lessons]. 4. vyd. Seelze: Kallmeyer, 2012. ISBN 978-3-7800-1009-4.

²⁰ BENNETT, R. E., BRASWELL, J., ORANJE, A., SANDENE, B., KAPLAN, B. a YAN, F. Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP. The Journal of Technology, Learning and Assessment [online]. 2008, roč. 6, č. 9. Dostupné na: <https://ejournals.bc.edu/index.php/jtla/article/view/1639>.

²¹ THOMPSON, N. A. a WEISS, D. J. Computerized and adaptive testing in educational assessment. In: SCHEUERMANN, F. a BJÖRNSSON, J. (eds.). The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing [online]. 2009, roč. 2009, č. JRC49408, s. 127 – 133. Dostupné na: <https://publications.jrc.ec.europa.eu/repository/handle/JRC49408>.

²² CSAPÓ, B., MOLNÁR, G. a NAGY, J. Computer-based assessment of school readiness and early reasoning. Journal of Educational Psychology [online]. 2014, roč. 106, č. 3, s. 639 – 650. Dostupné na: <https://doi.org/10.1037/a0035756>.

²³ KYLLONEN, P. C. New constructs, methods, & directions for computer-based assessment. In: SCHEUERMANN, F. a BJÖRNSSON, J. (eds.). The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing [online]. 2009, roč. 2009, č. JRC49408, s. 151 – 156. Dostupné na: <https://publications.jrc.ec.europa.eu/repository/handle/JRC49408>.

žiacov na premýšľanie o úlohách a odpovediach (napr. Sweller, 2020)²⁴. Preto môžu žiaci s nižším študijným prospechom, ktorí majú väčšie problémy s konceptuálnym riešením úloh, ťažšie zvládať adaptáciu na nové požiadavky testovania. Tento spôsob uvažovania je však len čiastočne podporený predchádzajúcimi empirickými dôkazmi. Štúdia (Eno, 2011)²⁵ ukázala, že žiaci s lepšími výsledkami boli úspešnejší v testovaní na počítači než v papierovom testovaní. Jerrim a kol. (2018)²⁶ nezistili významnú interakciu medzi spôsobom testovania a úspešnosťou v teste. V štúdii Støle a kol. (2020)²⁷ boli dievčatá s najvyššou úrovňou čitateľského výkonu znevýhodnené testovaním na počítači.

Nasledujúce zistenia zachytávajú počiatočnú fázu prechodu VERA na počítačové testovanie a vychádzajú najmä zo zberu dát z roku 2019. Predstavujú preto historický empirický príklad toho, ako môže forma administrácie testu ovplyvniť výsledky žiakov, a nie opis súčasného stavu digitalizácie testovania VERA. Výsledky štúdie nemeckých výskumníkov Wagner a kol. (2022)²⁸ odhalili, že výkony žiakov v testoch VERA z nemeckého jazyka v roku 2019 boli významne nižšie v teste na počítači ako v teste v papierovej forme. Tento výsledok sa týkal testovacích oblastí pravopis a čítanie v oboch testových zošitoch, ktoré vyplňali rôzne vzorky žiakov s rôznou úrovňou študijných výsledkov. Veľkosti týchto vplyvov boli pomerne vysoké. V testovom zošite 1 bol rozdiel medzi výsledkami z testu na počítači a v papierovej forme približne 0,6 štandardnej odchýlky; pri testovom zošite 2 bol rozdiel približne 0,3 štandardnej odchýlky. Žiaci teda dosiahli v teste VERA administrovanom na počítači horšie výsledky ako v teste na papieri. Výsledky sú v súlade s výsledkami štúdií Eyre (2017)²⁹, Dahan Golan a kol. (2018)³⁰ a Støle a kol. (2020), v ktorých žiaci dosiahli v testoch zameraných na porozumenie textu nižšie skóre úspešnosti na počítači ako v papierových testoch.

Podľa autorov tejto štúdie mohol byť horší výkon žiakov v teste na počítači spôsobený tým, že žiaci ešte neboli na tento spôsob testovania zvyknutí. Väčšina testov a skúšok v triede sa stále píše v papierovej forme. Testovanie s využitím počítača ponúka niekoľko výhod, ale zahŕňa aj problémy, ako je napr. písanie dlhých viet pomocou klávesnice alebo posúvanie obrazovky pri hľadaní konkrétnej informácie na vyriešenie úlohy. Podľa autorov si žiaci najprv musia zvyknúť na plnenie týchto požiadaviek, a až potom sa môžu plne sústrediť na pochopenie a koncepčné riešenie úloh. Okrem toho si aj učitelia musia zvyknúť na realizáciu testov

²⁴ SWELLER, J. Cognitive load theory and educational technology. Educational Technology Research and Development [online]. 2020, roč. 68, č. 1, s. 1 – 16. Dostupné na: <https://doi.org/10.1007/s11423-019-09701-3>.

²⁵ ENO, L. P. Comparing the Reading Performance of High-Achieving Adolescents: Computer-Based Testing versus Paper/Pencil [online]. South Orange: Seton Hall University, 2011. Dizertačná práca. Dostupné na: <https://search.proquest.com/docview/1018431561>.

²⁶ JERRIM, J., MICKLEWRIGHT, J., HEINE, J.-H., SALZER, C. a McKEOWN, C. PISA 2015: How big is the 'mode effect' and what has been done about it? Oxford Review of Education [online]. 2018, roč. 44, č. 4, s. 476 – 493. Dostupné na: <https://doi.org/10.1080/03054985.2018.1430025>.

²⁷ STØLE, H., MANGEN, A. a SCHWIPPERT, K. Assessing children's reading comprehension on paper and screen: A mode-effect study. Computers & Education [online]. 2020, roč. 151, čl. 103861. Dostupné na: <https://doi.org/10.1016/j.compedu.2020.103861>.

²⁸ WAGNER, I., LOESCHE, P. a BIßANTZ, S. Low-stakes performance testing in Germany by the VERA assessment: analysis of the mode effects between computer-based testing and paper-pencil testing. European Journal of Psychology of Education [online]. 2022, roč. 37, s. 531 – 549. Dostupné na: <https://doi.org/10.1007/s10212-021-00532-6>.

²⁹ EYRE, J. On or off screen reading in a digital world. set: Research Information for Teachers [online]. 2017, č. 1, s. 53 – 58. Dostupné na: <https://doi.org/10.18296/set.0072>.

³⁰ DAHAN GOLAN, D., BARZILLAI, M. a KATZIR, T. The effect of presentation mode on children's reading preferences, performance, and self-evaluations. Computers & Education [online]. 2018, roč. 126, s. 346 – 358. Dostupné na: <https://doi.org/10.1016/j.compedu.2018.08.001>.

na počítači. Učitelia, ktorí dohliadali na testovanie na počítači, boli pravdepodobne neistejší a mohli mať väčšie technické problémy než učitelia, ktorí využívali papierový spôsob testovania, čo mohlo negatívne ovplyvniť atmosféru počas testovania.

Celkovo výsledky spomínanej nemeckej štúdie odhalili, že existuje viac dôkazov v neprospech počítačovej formy testov VERA v porovnaní s papierovou formou. Tento efekt bol konzistentný v rôznych oblastiach testovania a v rôznych testových zošitoch. Na základe týchto zistení výsledky testov VERA s využitím počítača v roku 2019 v predmete nemecký jazyk neboli priamo porovnávané s výsledkami testov v papierovej forme z predchádzajúcich rokov alebo z iných spolkových krajín.

Ako uvádzajú autori uvedenej štúdie, znevýhodnenie žiakov v testovaní na počítači by sa mohlo v nasledujúcich rokoch znížiť, ak by si žiaci čoraz viac zvykali na toto nové testovacie médium a jeho požiadavky. Autori odporúčajú, aby učitelia častejším využívaním digitálnych technológií získavali väčšiu istotu pri realizácii počítačového testovania. Súčasne by sa tento vývoj mohol urýchliť častejším začleňovaním testov a skúšok s využitím počítača do bežných postupov v triede.

Učitelia by sa mali pri tvorbe a vykonávaní elektronických formátov testov viac zdokonaľovať a riaditelia škôl by ich mohli motivovať k používaniu tohto nového spôsobu testovania. Medzi ďalšie odporúčania patrí, aby riaditelia škôl v súvislosti so zavádzaním testovania na počítači podporovali medzi učiteľmi spoluprácu. Taktiež by učiteľom mohol byť pridelený kolega s technickými znalosťami, ktorý by im pomáhal pri vykonávaní testovania s využitím počítača.

Zo záverov štúdie vyplýva, že tendenciu dosahovať nižšie skóre v testovaní VERA s využitím počítača mali žiaci s horšími študijnými výsledkami a s nemeckým jazykom ako nedominantným jazykom. Učitelia by mali preto venovať dostatočnú pozornosť príprave žiakov na spôsob, ktorým bude testovanie v ich triede realizované.

Syntetické zhrnutie – Nemecko (a prínosy pre Slovensko)

Nemecký systém externého hodnotenia žiakov predstavuje dlhodobu stabilnú a odborne ukotvenú model kritériálneho testovania, ktorý je pevne previazaný s jasne definovanými vzdelávacími štandardmi. Porovnávacie testy VERA sú navrhnuté ako diagnostický nástroj, ktorého cieľom je poskytovať školám a učiteľom spoľahlivú spätnú väzbu o dosahovaní očakávaných kompetencií žiakov v kľúčových fázach ich vzdelávacieho procesu.

Kľúčovou charakteristikou nemeckého prístupu je dôsledné oddelenie hodnotiacej a kontrolnej funkcie testovania. Výsledky testov VERA sa nepoužívajú na známkovanie žiakov, hodnotenie učiteľov ani na vytváranie porovnávacích rebríčkov škôl. Tento princíp významne prispieva k akceptácii testovania zo strany škôl a k jeho využívaniu ako nástroja podpory vyučovania, nie ako nástroja sankcie.

Za osobitne prínosný možno považovať model inštitucionálneho zabezpečenia testovania. Centrálna odborná príprava testových nástrojov prostredníctvom výskumno-akademickej inštitúcie (IQB) zaručuje vysokú odbornú a psychometrickú kvalitu testov, zatiaľ čo realizácia testovania a práca s výsledkami zostáva v kompetencii regionálnych aktérov. Tento model umožňuje kombinovať jednotné odborné štandardy s rešpektovaním regionálnej autonómie školských systémov.

Z hľadiska prínosu pre Slovensko je nemecká skúsenosť významná najmä v troch oblastiach: 1) ukazuje **dôležitosť jasne formulovaných vzdelávacích štandardov ako nevyhnutného východiska pre akékoľvek**

kriteriálne testovanie, 2) zdôrazňuje potrebu definovať **externé testovanie ako nástroj pedagogickej diagnostiky a podpory vyučovania, nie ako mechanizmus kontroly**, 3) poukazuje na **význam systematickej práce s dátami na úrovni systému, ktorá umožňuje sledovať dlhodobé trendy vo výsledkoch žiakov bez stigmatizácie jednotlivých škôl**.

Nemecký model tak predstavuje pre Slovensko relevantnú inšpiráciu pri ďalšom rozvoji kriteriálneho testovania, najmä **pri nastavovaní jeho účelu, interpretácii výsledkov a vzťahu medzi centrálnou odbornou prípravou testov a ich využívaním v školách**.

Limity prenositeľnosti a implikácie pre Slovensko

Pri prenose prvkov nemeckého systému externého hodnotenia do slovenského prostredia je potrebné zohľadniť viaceré systémové odlišnosti a z nich vyplývajúce implikácie pre tvorbu politík.

Nemecký vzdelávací systém je vysoko decentralizovaný, pričom spolkové krajiny disponujú rozsiahlymi kompetenciami v oblasti implementácie hodnotiacich nástrojov. Slovenský systém je, naopak, výrazne centralizovanejší. Z toho vyplýva potreba zachovať silnú centrálnu koordináciu testovania, zároveň však postupne posilňovať autonómiu škôl pri práci s výsledkami a ich využívaní na zlepšovanie výučby.

Kľúčovým predpokladom fungovania kriteriálneho testovania v Nemecku je **stabilný a dlhodobo rozvíjaný systém vzdelávacích štandardov**. V slovenskom kontexte, kde je kurikulárna reforma ešte vo vývoji, je preto nevyhnutné prioritne pokračovať v ich precizovaní, operacionalizácii a metodickom ukotvení tak, aby mohli slúžiť ako spoľahlivý základ pre tvorbu testových nástrojov.

Významnú úlohu v nemeckom modeli zohráva silné inštitucionálne zázemie reprezentované nezávislou odbornou inštitúciou (IQB), ktorá zabezpečuje vysokú kvalitu testovania. Pre Slovensko z toho vyplýva **potreba systematicky budovať a stabilizovať odborné kapacity v oblasti psychometrie a edukometrie**, didaktiky a hodnotenia vzdelávania, ako aj jasne definovať kompetencie jednotlivých inštitúcií zapojených do procesu testovania.

Z hľadiska pedagogickej praxe je dôležitým faktorom rozvinutá kultúra práce s dátami a spätnou väzbou. **V slovenskom prostredí je preto potrebné cielene podporovať profesijný rozvoj učiteľov v oblasti diagnostického hodnotenia a interpretácie výsledkov testovania, aby sa externé merania stali reálnym nástrojom zlepšovania výučby**, nie iba formálnym prvkom systému.

Osobitnú pozornosť si vyžaduje aj zavádzanie digitálneho testovania. Skúsenosti z Nemecka poukazujú na možné rozdiely vo výkone žiakov v závislosti od formy testovania, najmä v počiatočných fázach jeho implementácie. Pre Slovensko z toho vyplýva potreba postupného zavádzania elektronického testovania, sprevádzaného metodickou podporou škôl, prípravou žiakov na digitálne prostredie a dôsledným zabezpečením porovnateľnosti výsledkov v čase.

V dlhodobejšej perspektíve vytvára digitalizácia podmienky aj na zavedenie počítačovo adaptívneho testovania. Pri tomto type testovania sa výber ďalších úloh a ich náročnosť prispôsobujú predchádzajúcim odpovediam žiaka. Žiak tak rieši úlohy, ktoré lepšie zodpovedajú aktuálnej úrovni jeho vedomostí a zručností. Adaptívne testovanie môže priniesť presnejší odhad výkonu v rôznych častiach výkonového spektra, skrátiť čas potrebný na testovanie a znížiť počet úloh, ktoré sú pre konkrétného žiaka neprimerane ľahké alebo náročné. Zároveň umožňuje efektívnejšie využívať rozsiahle banky úloh a poskytovať podrobnejšiu

diagnostickú spätnú väzbu o zvládnutých a ešte nerozvinutých oblastiach učenia. Tieto vlastnosti sú osobitne významné pri kritériálne orientovanom testovaní, pretože môžu podporiť presnejšie určovanie dosiahnutej úrovne výkonu a formulovanie odporúčaní pre ďalšie učenie žiaka.

Pre Slovensko by zavedenie adaptívneho testovania malo predstavovať postupný a výskumne overovaný proces. Predpokladá vytvorenie dostatočne rozsiahlych a kvalitne kalibrovaných bánk úloh, využitie vhodných psychometrických modelov, zabezpečenie technologickej infraštruktúry, ochrany údajov a prístupnosti testovania pre žiakov so špecifickými potrebami. Nevyhnutné bude tiež pilotné overovanie algoritmov výberu úloh, posúdenie porovnateľnosti adaptívnych a neadaptívnych foriem testu a zachovanie kontinuity výsledkov v čase. Perspektívnym riešením môže byť najskôr viacstupňové adaptívne testovanie³¹, pri ktorom sa žiakom nepridelujú jednotlivito vybrané úlohy, ale ucelené bloky úloh rozdielnej náročnosti. Takýto model je z hľadiska riadenia obsahu, interpretácie výsledkov a technickej realizácie vhodným medzistupňom medzi lineárnym elektronickým testom a plne adaptívnym testovaním.

Uvedené skutočnosti naznačujú, že nemecký model predstavuje cenný zdroj inšpirácie, jeho uplatnenie v slovenských podmienkach si však vyžaduje postupný a primerane prispôbený prístup, ktorý zohľadní osobitosti národného vzdelávacieho systému a súčasnú fázu jeho transformácie.

1.3 ESTÓNSKO – model kombinácie externej evalvácie a formatívneho hodnotenia

Estónsky model hodnotenia možno interpretovať v kontexte teórií formatívneho hodnotenia (napr. Black a Wiliam, 1998) a konceptu hodnotenia pre učenie (*assessment for learning*), ktoré zdôrazňujú význam spätnej väzby pre zlepšovanie učenia žiakov³².

Podľa medzinárodných štúdií je estónsky vzdelávací systém považovaný za vysoko výkonný systém, v ktorom žiaci dosahujú vynikajúce výsledky³³. Žiaci z rôznych sociálno-ekonomických prostredí majú dobrý prístup k vzdelávaniu, je tu malý podiel žiakov so slabými výsledkami a mnohí žiaci zo znevýhodneného prostredia dosahujú výborné výsledky³⁴. Okrem toho sa u estónskych žiakov spájajú dobré študijné výsledky s vysokou úrovňou životnej spokojnosti³⁵. V Estónsku sa vzdelanie vysoko cení. Princípy aj obsah súčasného vzdelávacieho systému boli stanovené po obnovení nezávislosti tejto krajiny v roku 1990³⁶.

³¹ OECD. *PISA 2025 Integrated Design*. Paris: OECD, 2023, s. 32–36.

³² BLACK, P. a WILIAM, D. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*. 1998, roč. 5, č. 1, s. 7 – 74. ISSN 0969-594X.

³³ OECD. Estonia Country Note, PISA 2018 Results: Where All Students Can Succeed [online]. Paris: OECD Publishing, 2019. Dostupné na: https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/pisa/publications/national-reports/pisa-2018/featured-country-specific-overviews/PISA2018_CN_EST.pdf.

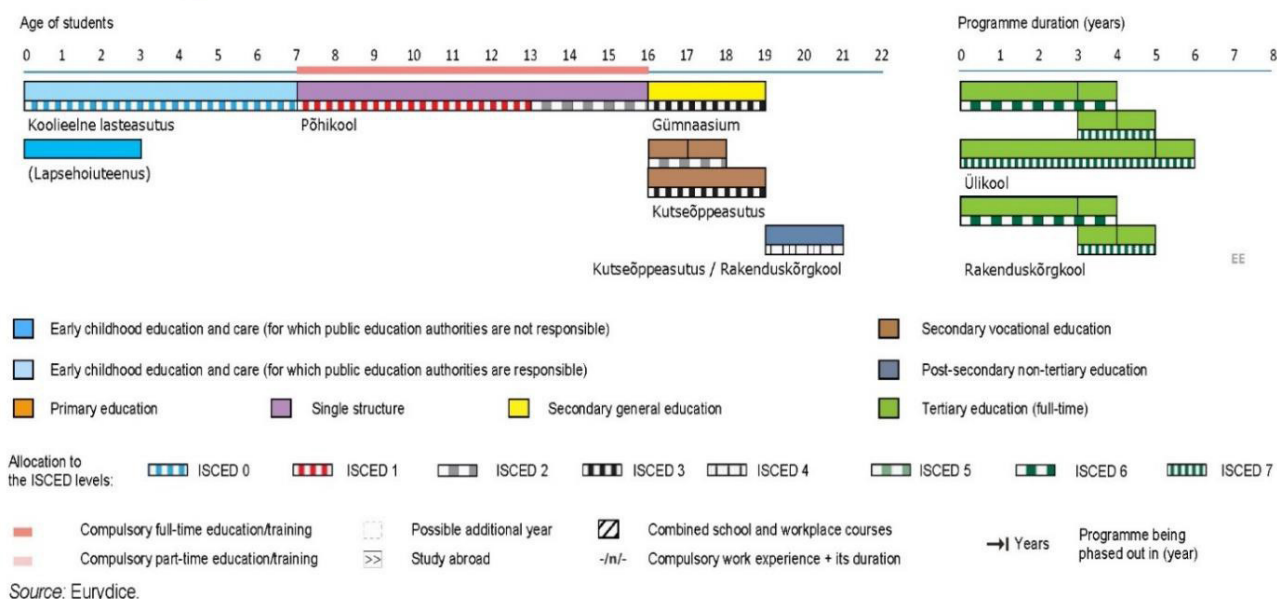
³⁴ OECD. Estonia Country Note, PISA 2018 Results: Where All Students Can Succeed [online]. Paris: OECD Publishing, 2019. Dostupné na: https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/pisa/publications/national-reports/pisa-2018/featured-country-specific-overviews/PISA2018_CN_EST.pdf.

³⁵ OECD. PISA 2022 Results: Factsheets – Estonia [online]. Paris: OECD Publishing, 2023. Dostupné na: https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/estonia_dafed886-en.html.

³⁶ LEES, M. Estonian Education System 1990-2016: Reforms and Their Impact [online]. 4liberty.eu, 2017. Dostupné na: https://4liberty.eu/wp-content/uploads/2016/08/Estonian-Education-System_1990-2016.pdf.

Estónsky vzdelávací systém

Estonia – 2023/2024



Obr. 5 Štruktúra vzdelávacieho systému v Estónsku³⁷

Základné vzdelávanie

Povinné základné vzdelávanie³⁸ sa realizuje od 1. do 9. ročníka. Rodičia môžu slobodne vybrať školu pre dieťa v predškolskom veku. Základná škola nadväzuje na predprimárne vzdelávanie, ktoré je deťom poskytované vo veku od 18 mesiacov do 7 rokov. Každé dieťa má právo získať základné vzdelanie v škole v blízkosti svojho bydliska. Uplatňujú sa princípy inkluzívneho vzdelávania, žiaci so špeciálnymi výchovno-vzdelávacími potrebami sa zvyčajne učia v bežnej triede svojej školy.

Základná škola je tvorená 3 stupňami: I. stupeň – 1. až 3. ročník; II. stupeň – 4. až 6. ročník; III. stupeň – 7. až 9. ročník.

Školy využívajú na hodnotenie vedomostí a zručností žiakov formatívne aj sumatívne hodnotenie. Pod pojmom formatívne hodnotenie sa rozumie hodnotenie prebiehajúce počas procesu vzdelávania. Učitelia poskytujú žiakom priebežnú ústnu alebo písomnú spätnú väzbu o ich výsledkoch, silných stránkach a nedostatkoch, pomáhajú podporovať nadšenie žiakov a usmerňujú ich vo vzdelávacej dráhe. Učitelia pri plánovaní cieľov a spôsobov učenia sa žiakov úzko spolupracujú s rodičmi, odbornými zamestnancami škôl a ďalším podporným personálom. Aspoň raz v priebehu školského roka škola uskutoční rozhovor za prítomnosti žiaka, triedneho učiteľa a rodiča, ktorý je zameraný na ďalšie študijné a rozvojové ciele žiaka.

Žiaci 9. ročníka, ktorí ukončujú vzdelávanie na základnej škole, prezentujú vlastnú tvorivú prácu, na ktorej pracujú počas celého tretieho stupňa štúdia. Okrem toho absolvujú tri záverečné centrálné zadávané a centrálné hodnotené sumatívne testy s jednotnými otázkami: z estónskeho jazyka a literatúry (alebo z estónčiny ako druhého jazyka), test z matematiky a test z ďalšieho predmetu podľa vlastného výberu žiaka.

³⁷ EURYDICE. Estonia: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurydice/estonia/overview>.

³⁸ EDUCATION ESTONIA. Estonian education system [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/about-education-system>.

Všeobecné stredoškolské vzdelávanie

Všeobecné stredoškolské vzdelanie sa získava na úrovni vyššieho stredného vzdelávania. Študijný program na strednej škole je rozdelený do povinných a dobrovoľných kurzov. Štúdium trvá 3 roky.

Na to, aby študenti úspešne absolvovali vyššiu strednú školu, musia zložiť štátne maturitné skúšky z troch povinných predmetov – materinského jazyka (alebo z estónčiny ako druhého jazyka), matematiky, cudzieho jazyka – formou centrálne zadávaných a centrálne hodnotených testov. Ďalej absolvujú školskú skúšku alebo vypracujú praktickú prácu. Spomedzi cudzích jazykov si študenti môžu vybrať skúšku v ruskom, anglickom, nemeckom alebo francúzskom jazyku. Štátne skúšky majú žiaci možnosť absolvovať len z anglického jazyka. V ruštine, nemčine a francúzštine sú štátne skúšky nahradené medzinárodnými skúškami odbornej spôsobilosti. Štátnu maturitu z cudzieho jazyka žiak nemusí vykonať, ak škole predloží platné osvedčenie o dosiahnutí úrovne B1 alebo vyššej v medzinárodne uznávanej skúške (napr. First Certificate in English – FCE, DALF C1, Goethe-Zertifikat C1). **Výsledky maturitných skúšok na strednej škole zohľadňujú vysoké školy pri prijímaní na vysokoškolské štúdium.** Dosiahnutie všeobecného stredného vzdelania oprávňuje študentov pokračovať v štúdiu na vysokej škole alebo získať odborné vzdelanie.

Odborné školstvo

Odborné vzdelávanie slúži na podporu vedomostí, zručností a postojov, profesijného know-how a sociálnej pripravenosti potrebnej na prácu, účasť na spoločenskom živote a účasť na procese celoživotného vzdelávania. Odborné vzdelávanie je bezplatné a možno ho získať po základnej škole ako stredné odborné vzdelanie (dĺžka: 3 – 4 roky) alebo len ako odborné vzdelanie bez všeobecného vzdelania (dĺžka: 3 mesiace až 2,5 roka). Školy úzko spolupracujú s podnikmi a zamestnávateľmi pri tvorbe učebných osnov a pri vytváraní príležitostí pre učňovské vzdelávanie.

Externé hodnotenie žiakov v Estónsku

V nasledujúcej časti je podrobne rozpracovaná téma externého hodnotenia žiakov v Estónsku na prvom a druhom stupni vzdelávania so zameraním na overovanie vzdelávacích štandardov³⁹. V polovici 90. rokov 20. storočia bol v Estónsku zavedený systém externého hodnotenia, ktorého cieľom je poskytovať spätnú väzbu o tom, ako žiaci dosahujú ciele vzdelávania uvedené v národných učebných osnovách. Estónsko sa zúčastňuje na medzinárodnom hodnotení vzdelávania od roku 2003. Prvou štúdiou bola štúdia IEA (International Association for the Evaluation of Educational Achievement) Trends in International Mathematics and Science Study (TIMSS) v roku 2003, po ktorej nasledoval Program OECD pre medzinárodné hodnotenie študentov (PISA) v cykloch od roku 2006 do súčasnosti, Medzinárodná štúdia OECD o vyučovaní a učení (TALIS) od roku 2008 do súčasnosti, Program OECD pre medzinárodné hodnotenie kompetencií dospelých (PIAAC) a ďalšie štúdie. Ako uvádza vo svojom článku Tire (2020), budúcnosť národného systému hodnotenia je viac zameraná na formatívne hodnotenie a využívanie digitálnych nástrojov, ktoré podporujú učenie a rozvoj každého žiaka.

³⁹ TIRE, G. Educational Assessment in Estonia. In: HARJU-LUUKKAINEN, H., McELVANY, N. a STANG, J. (eds.). Monitoring Student Achievement in the 21st Century: European Policy Perspectives and Assessment Strategies [online]. Cham: Springer, 2020, s. 119 – 129. Dostupné na: https://doi.org/10.1007/978-3-030-38969-7_10.

Podľa estónskej stratégie celoživotného vzdelávania do roku 2020⁴⁰ a jej následného strategického plánu Vzdelávanie 2035⁴¹ je hlavným cieľom hodnotenia žiakov podpora vzdelávania. Tieto strategické plány boli vytvorené na základe širokého zapojenia zainteresovaných strán, aby usmerňovali a riadili celý vzdelávací systém v Estónsku⁴².

V prípade hodnotenia žiakov vychádzajú plány zo systémového prístupu, keďže hodnotenie sa nevníma len ako nástroj na učenie sa žiakov, ale aj ako prostriedok na učenie sa systému a ľudí v ňom. Ministerstvo školstva a výskumu Estónska využíva nástroje formatívneho hodnotenia vrátane prieskumov, aby podnietilo strany zainteresované vo vzdelávaní k väčšej osobnej a kolektívnej zodpovednosti za dlhodobý rozvoj a zlepšovanie škôl, a zároveň investovalo menej času a pozornosti do externých kontrol.

Testy s nízkymi dôsledkami (low-stakes tests)

Jednou z inovácií, ktorú prijalo estónske ministerstvo školstva a výskumu, bolo vytvorenie série prevažne *low-stakes* a často dobrovoľných hodnotiacich nástrojov bez známkovania, ale so spätnou väzbou na celoštátnej úrovni⁴³. Údaje získané z týchto nástrojov hodnotenia majú podporiť prístup založený na dôkazoch so zameraním na zlepšenie vyučovania a učenia sa⁴⁴.

Externé hodnotenie výsledkov vzdelávania sa uskutočňuje prostredníctvom testov na overenie vzdelávacích štandardov. Týmito testami sa zisťuje úroveň všeobecných a predmetových kompetencií žiakov podľa štátneho vzdelávacieho programu, dosiahnutie cieľov vzdelávania a učenia sa v jednotlivých predmetoch a prierezových témach na príslušnom stupni vzdelávania.

Testy overujúce vzdelávacie štandardy štátneho kurikula na I. stupni (1. – 3. ročník) sú tieto⁴⁵:

- ✓ v 3. ročníku test z matematiky;
- ✓ v 3. ročníku test z estónskeho jazyka v škole s estónskym vyučovacím jazykom;
- ✓ v 3. ročníku test z ruského jazyka v škole s ruským vyučovacím jazykom;

Testy na overenie vzdelávacích štandardov na II. stupni (4. – 6. ročník) sú tieto:

- ✓ v 4. ročníku e-test z prírodovednej gramotnosti;
- ✓ v 6. ročníku e-test z estónskeho jazyka (alebo estónskeho jazyka ako druhého jazyka);
- ✓ v 6. ročníku e-test z matematiky;
- ✓ v 6. ročníku e-test z ruského jazyka v škole s ruským vyučovacím jazykom;

⁴⁰ MINISTRY OF EDUCATION AND RESEARCH. The Estonian Lifelong Learning Strategy 2020 [online]. Tartu: Republic of Estonia, Ministry of Education and Research, 2014. Dostupné na: https://www.educationestonia.org/wp-content/uploads/2022/12/estonian_lifelong_strategy.pdf.

⁴¹ MINISTRY OF EDUCATION AND RESEARCH. Education Strategy 2021-2035 [online]. Tartu: Republic of Estonia, Ministry of Education and Research, 2021. Dostupné na: https://www.educationestonia.org/wp-content/uploads/2022/12/haridusvaldkonna_arengukava_2035_kinnittaud_vv_eng_0-1.pdf.

⁴² EDUCATION ESTONIA. Voluntary, yet attractive and powerful low-stakes assessment [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/innovation/assessment>.

⁴³ JUKK, H. Assessment of mathematical competence using the Estonian national e-tests. Eesti Haridusteaduste Ajakiri. Estonian Journal of Education [online]. 2021, roč. 9, č. 2, s. 110 – 136. Tartu: University of Tartu. Dostupné na: <https://doi.org/10.12697/EHA.2021.9.2.05>.

⁴⁴ EDUCATION ESTONIA. Voluntary, yet attractive and powerful low-stakes assessment [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/innovation/assessment>.

⁴⁵ EURYDICE. Estonia – 4. Single-structure primary and lower secondary education: 4.3 Assessment in single-structure education [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/estonia/assessment-single-structure-education>.

Testy na overenie vzdelávacích štandardov na III. stupni (7. – 9. ročník):

✓ v 7. ročníku e-test z prírodovednej gramotnosti.

Moderné e-testy na určenie úrovne dosahovaného štandardu sú zadávané elektronicky a žiaci ich vyplňajú pomocou počítačov v informačnom testovacom systéme.

Výsledky testov sú základom pre sprievodné správy učiteľa, ako aj pre spätnú väzbu pre žiakov, keďže sa na jednej strane zameriavajú na to, čo žiaci vedia a dokážu, a na druhej strane odhaľujú, čo ešte potrebujú vedieť a naučiť sa lepšie. Konečný účel testov je formatívny tak pre žiakov, ako aj pre učiteľov.

To, že sú učitelia lepšie informovaní o silných stránkach a rozvojových potrebách žiakov v kombinácii s príslušnými odporúčaniami o stratégiách vyučovania, by malo viesť k zlepšeniu výučby jednotlivých žiakov aj skupín a v konečnom dôsledku k zlepšeniu učenia sa žiakov. Správy, ktoré žiaci dostávajú ako výstupy z testovania, obsahujú aj návrhy konkrétnych stratégií na zlepšenie procesu učenia sa, čím by mali žiakom umožniť, aby prevzali väčšiu kontrolu nad vlastným učením sa. Podľa estónskej legislatívy sa tieto testy nemôžu používať na známkovanie alebo ako základ pre rozhodovanie o budúcej vzdelávacej dráhe žiaka. Namiesto toho sú určené na podporu využívania hodnotenia ako účinného prostriedku na podporu vzdelávania žiakov a učiteľov.

Testy sú pre učiteľov voľne k dispozícii a žiaci ich môžu absolvovať online. Rovnako môžu učitelia vyplniť súvisiaci dotazníkový prieskum o ich vyučovacích postupoch, ktorý je pre učiteľov dostupný online. Použitie testov aj dotazníkových prieskumov je pre učiteľov dobrovoľné. Tieto nástroje majú inšpirovať autonómnych estónskych učiteľov k ďalšiemu uvažovaniu a diskusii o hodnote vyučovania všeobecných kompetencií vrátane zručností pri učení. Ponúkajú učiteľom a nepriamo aj žiakom príležitosť bez nátlaku sa naučiť, ako zlepšiť vyučovanie a/alebo učenie sa.

Spätná väzba

Testovací systém automaticky generuje spätnú väzbu vo forme individuálnych správ o každom žiakovi a súhrnných správ o triedach ako celku⁴⁶. Súhrnné správy identifikujú potreby rozvoja žiakov a pomáhajú učiteľom interpretovať výsledky. Keďže jednotlivé podpoložky testov pre žiakov a prieskumu učiteľov sú zosúladené, správy pomáhajú ozrejmiť, ktoré postupy uvádzané učiteľmi sa javia ako účinné a ktoré si vyžadujú ďalšiu pozornosť s cieľom lepšie podporiť jednotlivých žiakov alebo skupiny žiakov.

Správy o jednotlivých žiakoch alebo triedach nie sú priamo prístupované riaditeľom škôl. Testy majú slúžiť ako súčasť osobného súboru nástrojov učiteľa na zlepšenie vyučovania a učenia sa. Učiteľom sa však odporúča, aby sa poradili so školským psychológom pri interpretácii výsledkov, a aby požiadali riaditeľov škôl o podporu, ktorá by im uľahčila cielenú pomoc žiakom pri učení.

Keďže **triedni učitelia v Estónsku často učia tú istú triedu žiakov od 1. ročníka až do konca 6. ročníka, testy umožňujú učiteľom sledovať pokrok žiakov počas niekoľkých rokov**, a získať tak dôkazy o účinnosti vyučovacích stratégií a stratégií učenia sa žiakov. To znamená, že učitelia majú dobré podklady na to, aby počas niekoľkých rokov prispôbili svoje vyučovanie potrebám žiakov.

⁴⁶ KIKAS, E., MÄDAMÜRK, K., TREIAL, K., LUPTOVA, O., MALLEUS, E., SOODLA, P. a kol. Arvutipõhised hindamisvahendid õpi-, suhtlus- ja enesemääratluspädevuse hindamiseks I ja II kooliastmes [Computer-based Assessment Instruments for Measuring Learning, Communication and Self-regulation Skills in Key Stages I and II]. Tallinn: Tallinn University, 2018.

Testy na hodnotenie kompetencií

Estónski vytvárajú aj testy na overovanie matematických a jazykových kompetencií, ktorých cieľom je zistiť, akú úroveň žiak dosiahol, identifikovať možné dôvody nedosiahnutia očakávanej úrovne a vymedziť kognitívne a afektívne procesy (napr. myslenie, pamäť, motivácia, plánovanie), ktorým je potrebné venovať ďalšiu pozornosť. Tie sa považujú za kľúčové pri pomoci učiteľom a širokému okruhu žiakov pri budovaní všeobecnej schopnosti čítať a používať matematické vedomosti a zručnosti. Zavedené sú aj dobrovoľné testy digitálnych kompetencií zamerané na žiakov 8. ročníka.

Zameranie na budúcnosť

Estónske ministerstvo školstva a výskumu sa v nasledujúcich rokoch plánuje predovšetkým zamerať na hľadanie spôsobov, ako pomôcť väčšiemu počtu učiteľov a žiakov využívať hodnotenie ako nástroj podporujúci učenie a sebareflexiu. Zároveň naďalej zdokonaľuje testovacie nástroje. Väčšiu pozornosť plánuje venovať najmä zlepšeniu spätnej väzby, poskytovaniu informácií a skúmaniu spôsobov, ako zvýšiť využívanie spätnej väzby zo strany žiakov a učiteľov. Bude ponúkať ďalší profesijný rozvoj pedagógov s dôrazom na efektívnejšie využívanie testov a správ, pričom niektoré testy budú súčasťou systémového monitorovania a niektoré budú môcť byť školami využívané dobrovoľne.

Nástroje prieskumu

Ako protíváha k zdôrazňovaniu výsledkov bolo vyvinutých niekoľko prieskumných nástrojov na meranie spokojnosti žiakov, učiteľov a rodičov so vzdelávaním. Aj tieto nástroje boli iniciované Stratégiou celoživotného vzdelávania 2020⁴⁷, ktorá označila úroveň spokojnosti a pohody zainteresovaných strán v oblasti vzdelávania (t. j. well-being) za dôležitý ukazovateľ úspešnosti popri výsledkoch žiakov.

Prieskumy pomáhajú zviditeľniť tie faktory spojené so spokojnosťou a pohodou (ďalej aj „pohoda“, resp. „pohoda žiakov/učiteľov“), ktoré ovplyvňujú učenie sa žiakov. Od roku 2018 majú všetky všeobecnovzdelávacie školy možnosť každoročne zapojiť žiakov 4. aj 8. ročníka do prieskumu. Učitelia a rodičia sa prieskumu zúčastňujú každé tri roky.

Prieskumom spokojnosti žiakov sa zisťuje:

- celková spokojnosť v škole,
- spokojnosť s konkrétnymi aspektmi vzdelávacieho prostredia,
- spokojnosť s vonkajšími faktormi, ktoré môžu respondenti spájať so spokojnosťou v škole.

Žiacke prieskumy

Žiacke dotazníkové prieskumy sú koncipované tak, aby systematicky zachytávali široké spektrum faktorov ovplyvňujúcich vzdelávacie skúsenosti a pohodu žiakov. Metodologicky vychádzajú z rozlíšenia **zjavných (pozorovateľných)** a **latentných (nepozorovateľných)** premenných.

⁴⁷ MINISTRY OF EDUCATION AND RESEARCH. The Estonian Lifelong Learning Strategy 2020 [online]. Tartu: Republic of Estonia, Ministry of Education and Research, 2014. Dostupné na: https://www.educationestonia.org/wp-content/uploads/2022/12/estonian_lifelong_strategy.pdf.

Zjavné premenné sa zameriavajú na mapovanie konkrétnych prejavov správania a skúseností žiakov v školskom prostredí. Zahŕňajú najmä výskyt rôznych foriem rizikového a problémového správania, najmä rôzne formy šikanovania (verbálne, fyzické a online), ako aj správanie súvisiace s poškodzovaním majetku alebo porušovaním pravidiel. Tieto premenné umožňujú sledovať frekvenciu a rozšírenosť javov, ktoré majú priamy vplyv na bezpečnosť a sociálnu klímu školy.

Latentné premenné predstavujú komplexnejšie psychologické a sociálne konštrukty, ktoré nie je možné merať priamo. Ich zisťovanie je založené na súbore indikátorov reflektujúcich subjektívne prežívanie žiakov. Prieskumy sa v tejto oblasti zameriavajú najmä na aspekty osobnej autonómie a samostatnosti, kvalitu sociálnych vzťahov a pocit spolupatričnosti, vnímanú zmysluplnosť učenia, podporu zo strany spolužiakov a učiteľov, ako aj na charakter spätnej väzby v edukačnom procese. Súčasťou sledovaných oblastí sú aj faktory súvisiace s psychickou a fyzickou záťažou, ako sú únava, stres či frustrácia.

Takto koncipované žiacke prieskumy poskytujú komplexný obraz o tom, ako žiaci vnímajú vzdelávacie prostredie, vzťahy a procesy učenia. Získané údaje predstavujú dôležitý analytický podklad pre hodnotenie kvality školského prostredia a pre identifikáciu oblastí, ktoré si vyžadujú ciele intervencie v rámci školského rozvoja a systémových opatrení.

Prieskumy medzi učiteľmi

Dotazníkové prieskumy medzi učiteľmi sa zameriavajú na sociálne aj fyzické aspekty školského prostredia. Sociálne aspekty zahŕňajú napríklad výskyt šikanovania, aktivity učiteľov a vedenia školy či kvalitu vzťahov na pracovisku. Fyzické aspekty sa týkajú najmä stravovania, pracovných priestorov, dostupnosti podporných služieb a kvality učebných materiálov⁴⁸.

Každá škola dostáva podrobnú správu o tom, ako jej školská komunita hodnotila jednotlivé faktory zahrnuté v prieskume. Výsledky sú prezentované anonymizovane na úrovni jednotlivých otázok, pričom súhrnné údaje umožňujú vyvodiť závery aj o širších latentných premenných.

Výsledky jednotlivých škôl sa porovnávajú s národnými priermi. Školy tak každoročne získavajú prehľad o miere spokojnosti žiakov so vzdelávacím prostredím a vzdelávacími procesmi, ako aj o tom, ako sa tieto výsledky líšia od celonárodného priemeru. Každé tri roky školy zároveň dostávajú správy o spokojnosti učiteľov s ich pracovnými podmienkami a procesmi učenia žiakov, ako aj o spokojnosti rodičov s procesom vzdelávania ich detí.

Dôležitým aspektom je, že samotné prieskumy fungujú ako nástroj prepojenia a participácie, ktorý respondentom poskytuje možnosť aktívne sa vyjadriť k fungovaniu školy a podieľať sa na jej zlepšovaní. Získané správy pomáhajú žiakom, rodičom, učiteľom aj vedeniu škôl lepšie porozumieť, ktoré aspekty vzdelávacieho prostredia deti vnímajú pozitívne a ktoré negatívne, ako aj ich celkovej pohode v škole⁴⁹.

⁴⁸ INNOVE. Üldhariduskoolide 2018. aasta rahuloluküsitluste tulemused. Aruanne [General Education Schools 2018 Satisfaction Survey Results: Report]. Tallinn: Innove, 2019.

⁴⁹ VALK, A. Õpilaste ja õpetajate heaolu üle-Eestilises rahuloluküsitluses, lühiülevaade 2017. a tulemustest [Student and Teacher Well-Being in a Pan-Estonian Satisfaction Survey, Overview of Results from 2017]. Tartu: Ministry of Education and Research, 2018.

V estónskom vzdelávacom systéme predstavujú tieto údaje významný zdroj informácií, ktoré síce nie sú priamo normatívne, ale podnecujú rozsiahle odborné diskusie na úrovni škôl. Verejnosť má prístup k všeobecnému hodnoteniu spokojnosti v jednotlivých školách, ako aj k údajom o dochádzke a výskyte šikanovania. Zároveň je dostupný anonymizovaný národný prehľad všetkých výsledkov. Kľúčové zainteresované strany na základe podrobnejších národných analýz identifikujú problematické oblasti a v spolupráci so školami sa zapájajú do hľadania riešení. Platí, že ak sú problémy viditeľné a pomenované, zvyšuje sa pravdepodobnosť ich systémového riešenia.

Digitálne technológie vo vzdelávaní

Väčšina estónskych škôl využíva elektronické školské riešenia (napr. *eKool*, *Stuudium*)⁵⁰. Tieto inovatívne nástroje poskytujú rodičom, učiteľom a deťom jednoduchý spôsob spolupráce aj organizovania všetkých informácií potrebných na vyučovanie a učenie.

Už niekoľko rokov tu existuje rozvinutý ekosystém startupov pre vzdelávacie spoločnosti a školy majú veľkú podporu v podobe elektronických služieb zo súkromného sektora.

Od roku 2014 majú Estónci vypracovanú stratégiu celoživotného vzdelávania, ktorá zahŕňa aj program digitálnej transformácie. Cieľom tohto programu je pomôcť rozvíjať digitálne kompetencie učiteľov aj študentov. IT-kurzy a výučbové materiály pomohli začleniť digitálne technológie do vzdelávacieho procesu s cieľom rozvíjať digitálne kompetencie.

Ďalším míľnikom bolo, že Estónsko si stanovilo cieľ digitalizovať všetky vzdelávacie materiály už v roku 2015. Úspech digitálnej transformácie estónskeho vzdelávacieho systému závisí od dôkladného profesionálneho rozvoja aj odbornej prípravy učiteľov či odborníkov na vzdelávacie technológie.

Okrem výučby vedomostí a zručností v digitálnej oblasti estónske školstvo vo veľkej miere využíva množstvo inteligentných riešení: digitálne databázy, digitálne učebnice, elektronické učebné materiály, digitálne triedne denníky, digitálne hodnotenia, rôzne aplikácie a programy.

Syntetické zhrnutie – Estónsko (a prínosy pre Slovensko)

Estónsky systém hodnotenia žiakov je charakteristický dôsledným zameraním na podporu učenia sa žiakov, profesijnú autonómiu učiteľov a systematickú prácu so spätnou väzbou. **Kritériálne testovanie je v Estónsku vnímané predovšetkým ako nástroj formatívneho hodnotenia v zmysle konceptu *assessment for learning*, ktorého cieľom je pomáhať učiteľom a školám lepšie porozumieť vzdelávacím potrebám žiakov a cielene upravovať vyučovanie.**

Kľúčovým prvkom estónskeho prístupu je kombinácia povinných externých hodnotení na konci základného a stredného vzdelávania s rozsiahlym systémom dobrovoľných, prevažne elektronicky administrovaných *low-stakes* testov. Tieto testy sú navrhnuté tak, aby poskytovali zrozumiteľnú a prakticky využiteľnú spätnú väzbu žiakom aj učiteľom, bez priameho vplyvu na klasifikáciu alebo rozhodovanie o ďalšej vzdelávacej dráhe žiakov. Tento princíp významne prispieva k vysokej miere akceptácie externého hodnotenia v školách.

⁵⁰ EDUCATION ESTONIA. How did Estonia become a new role model in digital education? [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/how-did-estonia-become-a-new-role-model-in-digital-education>.

Významnou črtou estónskeho modelu je prepojenie testovania s hodnotiacimi správami a odporúčaniami, ktoré smerujú ku konkrétnym krokom v učení sa žiakov. Estónsky systém zároveň cielene dopĺňa testovanie kompetencií o dotazníkové prieskumy zamerané na podmienky učenia, motiváciu, školskú klímu a pohodu žiakov. Tým vytvára komplexnejší obraz o fungovaní vzdelávacieho procesu a podporuje učenie sa celého systému, nielen jednotlivcov.

Pre Slovensko je estónska skúsenosť mimoriadne inšpiratívna najmä v niekoľkých oblastiach. Ukazuje, že kritériálne testovanie môže byť **účinným nástrojom zlepšovania kvality vzdelávania**, ak je jasne oddelené od výberových a sankčných mechanizmov. Zdôrazňuje **význam kvalitnej spätnej väzby**, ktorá presahuje samotné skóre a ponúka **odporúčania využiteľné v každodennej pedagogickej praxi**. Zároveň poukazuje na to, že dôvera v **profesionalitu učiteľov**, dobrovoľnosť využívania nástrojov a systematická podpora ich práce môžu viesť k dlhodobo vysokým vzdelávacím výsledkom aj vysokej miere pohody žiakov.

Estónsky model tak predstavuje pre Slovensko **silnú inšpiráciu pri rozvoji formatívne orientovaného kritériálneho testovania**, pri budovaní **kultúry spätnej väzby** a pri **prepájaní hodnotenia s podporou autonómie škôl a učiteľov**.

Limity prenositeľnosti estónskeho modelu

Napriek tomu, že estónsky model hodnotenia žiakov predstavuje inšpiratívny príklad efektívneho prepojenia externého hodnotenia a formatívneho prístupu, jeho priamy prenos do slovenského vzdelávacieho prostredia naráža na viaceré systémové, kultúrne a kapacitné limity, čo možno interpretovať aj v kontexte **rizík nekritického preberania vzdelávacích politík (*policy borrowing*)**.

Jedným z kľúčových predpokladov fungovania estónskeho modelu je vysoká miera dôvery v profesionalitu učiteľov a relatívne vysoká úroveň ich pedagogickej autonómie. Učitelia v Estónsku majú dostatočný priestor na interpretáciu výsledkov hodnotenia a ich využívanie pri plánovaní výučby. V slovenskom kontexte je však profesijná autonómia učiteľov dlhodobo oslabovaná kombináciou normatívneho riadenia, administratívnej záťaže a nižšej systémovej podpory, čo môže obmedzovať efektívne využívanie formatívnych nástrojov.

Ďalším významným limitom je rozdielna úroveň pripravenosti systému na prácu s dátami. Estónsky model je založený na systematickom využívaní dát z testovania a prieskumov, ktoré sú prepojené s konkrétnymi odporúčaniami pre pedagogickú prax. Takýto prístup si vyžaduje nielen kvalitné analytické nástroje, ale aj rozvinuté dátové kompetencie učiteľov a vedenia škôl. **V slovenskom prostredí je práca s dátami často menej systematická a využívanie výsledkov hodnotenia na úrovni triedy a školy zostáva limitované.**

Významnou podmienkou estónskeho modelu je aj vysoká úroveň digitalizácie vzdelávacieho systému. Elektronické testovanie, automatizované generovanie spätnej väzby a prepojenie rôznych dátových zdrojov umožňujú efektívne fungovanie *low-stakes* testovaní. Hoci Slovensko v posledných rokoch zaznamenalo pokrok v oblasti digitalizácie, rozdiely v technickej infraštruktúre a digitálnych kompetenciách škôl a učiteľov môžu predstavovať bariéru pre plošnú implementáciu podobného modelu.

Limitujúcim faktorom je aj odlišná kultúra hodnotenia. Estónsky systém je postavený na princípe, že externé testovanie neslúži na sankcie ani selekciu, ale primárne na podporu učenia. V slovenskom prostredí má však testovanie historicky silnejšiu selekčnú a kontrolnú funkciu, čo ovplyvňuje aj jeho vnímanie zo strany škôl, učiteľov, žiakov a rodičov. **Prechod k formatívne orientovanému modelu preto predpokladá nielen zmenu nástrojov, ale aj zmenu myslenia a očakávaní všetkých aktérov.**

Dôležitým limitom je aj kapacita systému poskytovať kvalitnú metodickú podporu. Estónsky model je sprevádzaný systematickým profesijným rozvojom učiteľov a dostupnosťou prakticky využiteľných odporúčaní vychádzajúcich z výsledkov hodnotenia. V slovenskom kontexte sa postupne buduje systém podpory zavádzania inovácií do škôl.

Napokon, významnú úlohu zohráva aj stabilita a kontinuita vzdelávacej politiky. Estónsko dlhodobo uplatňuje konzistentnú stratégiu rozvoja vzdelávania, ktorá umožňuje postupné budovanie systému hodnotenia a dôvery medzi aktérmi. V podmienkach Slovenska, kde dochádza k častejším zmenám priorít a nástrojov, môže byť implementácia komplexnejších systémových zmien náročnejšia.

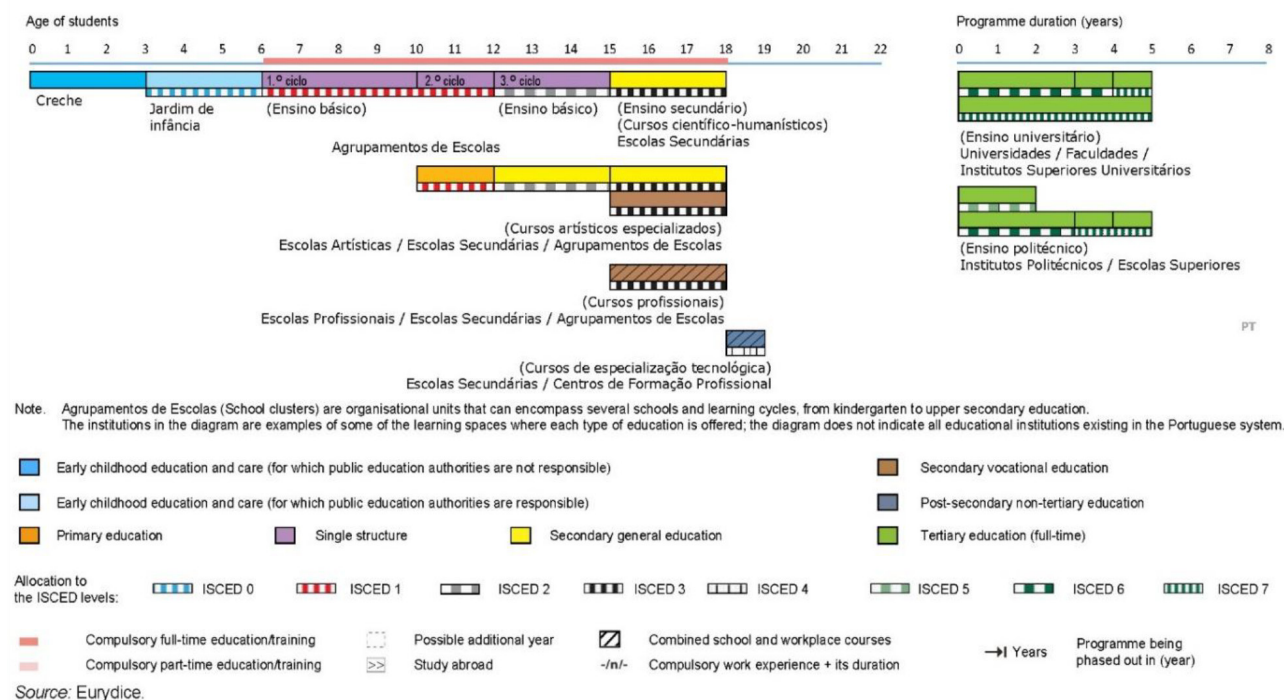
Z uvedených dôvodov možno konštatovať, že estónsky model nie je možné jednoducho preniesť do slovenského prostredia ako celok. Jeho prínos však spočíva v identifikácii princípov, ktoré sú adaptovateľné – najmä **posilnenie formatívnej funkcie hodnotenia, rozvoj práce s dátami, zlepšenie kvality spätnej väzby a postupné budovanie dôvery v profesionálnu autonómiu učiteľov. Implementácia týchto prvkov si však vyžaduje postupný, systémovo podporovaný a kontextovo citlivý prístup.**

1.4 PORTUGALSKO – reformy po roku 2000 a ich vývoj

Portugalsko je v súčasnosti považované za systém s výrazným zlepšením vzdelávacích výsledkov. Zo vzdelanostne zaostávajúcej krajiny v minulom storočí sa portugalský vzdelávací systém podľa výsledkov štúdie PISA dostal na úroveň priemeru krajín OECD. Najvýraznejší progres bol zaznamenaný po roku 2005, teda v období, keď Portugalsko implementovalo viaceré školské reformy inšpirované aj výsledkami a analýzami štúdie PISA. Vývoj naznačuje časovú súvislosť medzi reformami a zlepšením výsledkov, pričom ich presný podiel na tomto zlepšení nie je možné jednoznačne určiť.

Portugalský vzdelávací systém

Portugal – 2023/2024



Obr. 6 Štruktúra vzdelávacieho systému v Portugalsku⁵¹

⁵¹ EURYDICE. Portugal: Key features of the education system [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/portugal/overview>.

Základné vzdelávanie

Predškolská výchova a starostlivosť⁵² v Portugalsku je rozdelená do dvoch fáz, pričom sa líši ponukou pre mladšie (do troch rokov) a staršie deti (od troch rokov do začiatku povinnej školskej dochádzky vo veku 6 rokov). V Portugalsku základné vzdelávanie pokrýva primárne a nižšie sekundárne vzdelávanie v rámci jednej štruktúry. Základné vzdelávanie trvá deväť rokov a je rozdelené do troch cyklov:

1. cyklus: 1. – 4. ročník, vek 6 – 9 rokov
2. cyklus: 5. – 6. ročník, vek 10 – 11 rokov
3. cyklus: 7. – 9. ročník, vek 12 – 14 rokov

Prvý a druhý cyklus zodpovedajú primárnemu vzdelávaniu a tretí cyklus zodpovedá nižšiemu strednému vzdelávaniu. Cieľom všeobecného základného vzdelávania je poskytnúť žiakom všeobecné vzdelanie, aby mohli pokračovať v ďalšom štúdiu v súlade s princípmi, hodnotami a zručnosťami, ktoré sú uvedené v profile študentov ukončujúcich povinné vzdelávanie.

Neoddeliteľnou súčasťou práce učiteľov je hodnotenie žiakov prostredníctvom systematického zberu informácií a dát nevyhnutných na prijímanie vhodných rozhodnutí na zlepšenie kvality vzdelávania žiakov. Portugalské školy využívajú formatívne a sumatívne formy hodnotenia. Formatívne hodnotenie je kontinuálne a systematické, sumatívne pozostáva z celkového hodnotenia výsledkov žiakov na účely klasifikácie a certifikácie.

Za interné sumatívne hodnotenie na konci každého polroka školského roka sú zodpovední učitelia a vedenie školy. Externé sumatívne hodnotenie žiakov a škôl zabezpečuje ministerstvo školstva; pozostáva z dvoch typov testov, ktoré sú zadávané všetkým žiakom:

- ✓ štandardizované testy (*low-stakes tests*);
- ✓ celoštátne záverečné testy pri ukončení 3. cyklu (*high-stakes tests*).

Od školského roku 2024/2025 Portugalsko zaviedlo nový model externého hodnotenia žiakov. Jeho súčasťou sú povinné celoplošné skúšky monitorovania učenia *Provas de Monitorização da Aprendizagem* (ModA) v **4. a 6. ročníku** a celoštátne záverečné externé skúšky na konci **9. ročníka**. Model posilňuje porovnateľnosť výsledkov medzi školskými rokmi a ročníkmi, pričom na zabezpečenie porovnateľnosti umožňuje využívať kotviace položky. Súčasne zavádza včasnejšie reportovanie výsledkov a digitálnu alebo hybridnú administráciu externých skúšok v základnom vzdelávaní⁵³.

Výsledky žiakov a škôl v skúškach ModA sa uvádzajú v individuálnej správe o výsledkoch skúšok ModA a v školskej správe o výsledkoch skúšok ModA. Individuálna správa obsahuje kvantitatívne hodnotenie výkonu žiaka v sledovaných gramotnostiach, oblastiach a dimenziách spolu s opisom dosiahnutej výkonnej úrovne. Školská správa vychádza z agregovaných individuálnych výsledkov a poskytuje údaje za školu vrátane ich porovnania s výsledkami na úrovni obce a celého Portugalska.

⁵² EURYDICE. Portugal: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/portugal/overview>.

⁵³ PORTUGAL. MINISTÉRIO DA EDUCAÇÃO, CIÊNCIA E INOVAÇÃO. Novo modelo de avaliação externa dos alunos a partir de 2024/2025 [online]. 2024. Dostupné na: <https://portugal.gov.pt/gc24/comunicacao/comunicados/novo-modelo-de-avaliacao-externa-dos-alunos-a-partir-de-20242025>.

Oba typy správ poskytujú kvantitatívne výsledky aj ich opisnú interpretáciu, a predstavujú tak dôležité podporné nástroje pre žiakov, rodičov, učiteľov a školy. Umožňujú identifikovať silné stránky a oblasti vyžadujúce zlepšenie a podporujú učiteľov a vedenie školy pri plánovaní pedagogických a didaktických intervencií zameraných na skvalitňovanie učenia žiakov.

Všetci žiaci 9. ročníka na konci tretieho cyklu základného vzdelávania absolvujú externé testy z matematiky a portugalského jazyka (portugalského jazyka ako druhého jazyka alebo ako cudzieho jazyka). Testy vychádzajú z učebných osnov štátneho kurikula pre 3. cyklus a hodnotia študijné výsledky žiakov. Od školského roku 2022/23 sa štandardizované testy zadávajú žiakom elektronicky prostredníctvom počítačov a od roku 2024 sa aj záverečné externé testy pre 9. ročník vykonávajú na počítačoch.

Vyššie sekundárne vzdelávanie

Vyššie sekundárne vzdelávanie zahŕňa tri roky školskej dochádzky (10., 11. a 12. ročník) a v súčasnosti je povinné pre všetkých študentov vo veku do 18 rokov. Od roku 2009 Portugalsko zaviedlo pre deti a mládež vo veku 6 – 18 rokov vzdelávanie, ktoré je všeobecné, povinné a bezplatné.

Hodnotenie žiakov sa, podobne ako pri základnom vzdelávaní, realizuje formatívne aj sumatívne. Učitelia pri internom sumatívnom hodnotení používajú rôzne nástroje na hodnotenie vychádzajúce z metodických pokynov, ako aj individuálnych charakteristík žiakov (napr. písomné testy, experimentálne práce, projekty, portfóliá, výskum atď.).

Na to, aby žiaci ukončili vyššie sekundárne vzdelávanie, musia úspešne absolvovať všetky predmety zahrnuté v študijnom programe, ktorý navštevovali. Žiaci prírodovedno-humanitných kurzov musia zložiť aj záverečné národné skúšky (externé testy) zo štyroch predmetov: z portugalčiny a troch ďalších vybraných predmetov počas dvoch posledných ročníkov štúdia (11. a 12. ročník). Záverečné štátne skúšky, ktorými žiaci ukončujú stredoškolské štúdium, slúžia aj na prijímanie na vysoké školy.

Externé hodnotenie žiakov v Portugalsku

V nasledujúcej časti budú podrobnejšie analyzované faktory, ktoré viedli k úspechom portugalského vzdelávacieho systému v nadväznosti na externé hodnotenie žiakov. Politici v Portugalsku sa zaoberajú medzinárodnými hodnoteniami vzdelávacích systémov už od prvých výsledkov Stredomorského regionálneho projektu OECD na konci 60. rokov 20. storočia. Medzinárodné meranie TIMSS⁵⁴ 1995 bolo prvým rozsiahlym porovnávacím hodnotením, ktoré ukázalo, že portugalskí žiaci výrazne zaostávajú za rovesníkmi z 26 krajín. V tom čase tvorcovia vzdelávacej politiky nepovažovali TIMSS za platné meradlo špecifických vedomostí a zručností portugalských žiakov, pretože tie neboli v súlade s rámcom učebných osnov TIMSS (Carvalho a kol., 2017⁵⁵). Napriek tomu, že zistenia z TIMSS 1995 boli prijaté nepriaznivo, toto hodnotenie položilo základ zmien v politike hodnotenia a reforiem učebných osnov matematiky a prírodných vied. Bolotiež zlomovým bodom v uznaní potreby nielen ďalej hodnotiť portugalský vzdelávací systém podľa

⁵⁴ Poznámka: Medzinárodné meranie IEA TIMSS z matematickej gramotnosti a prírodovedy sa uskutočňuje na vzorkách žiakov 4. a 8. ročníka ZŠ.

⁵⁵ CARVALHO, L. M., COSTA, E. a GONÇALVES, C. Fifteen years looking at the mirror: On the presence of PISA in education policy processes (Portugal, 2000–2016). European Journal of Education [online]. 2017, roč. 52, č. 2, s. 154 – 166. Dostupné na: <https://doi.org/10.1111/ejed.12210>.

medzinárodných štandardov, ale aj venovať väčšiu pozornosť výsledkom žiakov v hlavných vyučovacích predmetoch (Crato, 2020⁵⁶; Marôco, 2020⁵⁷).

Vplyv štúdie PISA na zmeny vo vzdelávacom systéme Portugalska

Portugalsko sa ako člen OECD zúčastnilo prvého cyklu štúdie PISA 2000⁵⁸ a všetkých nasledujúcich cyklov. Podľa slov bývalého ministra školstva Nuno Crata, ktorý bol vo funkcii v rokoch 2012 až 2015, **sa PISA v Portugalsku vníma najmä „ako zrkadlo“, ktoré odráža, kde sa krajina nachádza v porovnaní s ostatnými krajinami** (Carvalho a kol., 2017⁵⁹), čo umožňuje tvorcom vzdelávacej politiky robiť zmeny vzdelávacích politík založené na dôkazoch.

João Marôco (2021)⁶⁰ vo svojom odbornom článku pojednávajúcom o vplyve štúdie PISA na zmeny vo vzdelávacom systéme Portugalska vysvetľuje, že neočakávané slabý výsledok Portugalska v štúdiu PISA 2000 pripravil pôdu pre schválenie série opatrení zameraných na podporu rôznych foriem mimoškolského štúdia a úpravu učebných osnov vyššieho sekundárneho vzdelávania vtedajším ministrom školstva Júliom Pedrosom. Ďalší minister, ktorý bol vo funkcii v rokoch 2003 až 2004, David Justino na základe slabých výsledkov v PISA a TIMSS podporil zavedenie národného testovania v Portugalsku, najprv v roku 2003 vo forme *low-stakes tests* (testy bez priameho dôsledku na žiaka) a v roku 2005 ako *high-stakes tests* (s významnými dôsledkami pre ďalšiu vzdelávaciu dráhu žiaka) z matematiky a portugalského jazyka na konci deviateho ročníka.

Nasledujúca ministerka školstva Maria de Lurdes Rodrigues (vo funkcii 2005 – 2009) označila štúdiu PISA za hlavný zdroj štatistických údajov o gramotnosti portugalských žiakov a jej význam pre podporu politík založených na dôkazoch. Zistenia zo štúdie PISA 2000 a 2003, podľa ktorých portugalskí žiaci dosahovali slabé výsledky v oblasti čitateľskej, matematickej a prírodovednej gramotnosti, podnietili a urýchlili zavedenie série programov a strategických plánov upevňujúcich politiky v oblasti vzdelávania.⁶¹ Počas svojho pôsobenia ministerka Rodriguesová zdôrazňovala súvislosť medzi slabými výsledkami v PISA a kultúrnym a sociálno-ekonomickým postavením rodín s cieľom **rozšíriť ekonomickú podporu pre žiakov z rodín s nízkymi príjmami, uľahčiť prístup k internetu a počítačom pre základné vzdelávanie a reorganizovať program prioritných intervenčných vzdelávacích území pre školy nachádzajúce sa v ekonomicky zaostalých oblastiach** (Afonso a Costa, 2009⁶²).

⁵⁶ CRATO, N. Curriculum and Educational Reforms in Portugal: An Analysis on Why and How Students' Knowledge and Skills Improved. In: REIMERS, F. M. (ed.). Audacious Education Purposes: How Governments Transform the Goals of Education Systems [online]. Cham: Springer, 2020, s. 209 – 231. Dostupné na: https://doi.org/10.1007/978-3-030-41882-3_8.

⁵⁷ MARÔCO, J. International large-scale assessments: Trends and effects on the Portuguese public education system. In: HARJU-LUUKKAINEN, H., McELVANY, N. a STANG, J. (eds.). Monitoring Student Achievement in the 21st Century: European Policy Perspectives and Assessment Strategies [online]. Cham: Springer, 2020, s. 207 – 222. Dostupné na: <https://doi.org/10.1007/978-3-030-38969-7>.

⁵⁸ Poznámka: Slovensko sa nezapojilo do 1. cyklu OECD PISA 2000, ale až o 3 roky neskôr do 2. cyklu PISA 2003.

⁵⁹ CARVALHO, L. M., COSTA, E. a GONÇALVES, C. Fifteen years looking at the mirror: On the presence of PISA in education policy processes (Portugal, 2000–2016). European Journal of Education [online]. 2017, roč. 52, č. 2, s. 154 – 166. Dostupné na: <https://doi.org/10.1111/ejed.12210>.

⁶⁰ MARÔCO, J. Portugal: The PISA Effects on Education. In: CRATO, N. (ed.). Improving a Country's Education [online]. Cham: Springer, 2021, s. 159 – 174. Dostupné na: https://doi.org/10.1007/978-3-030-59031-4_8.

⁶¹ Poznámka: Patrili k nim Program odbornej prípravy v oblasti experimentálneho vyučovania prírodných vied (2006), Národný program vyučovania portugalského jazyka (2007) zameraný na základné vzdelávanie, Akčný plán pre matematiku (2006) a Národný plán pre čítanie (2007).

⁶² AFONSO, N. a COSTA, E. The influence of the Programme for International Student Assessment (PISA) on policy decision in Portugal: the education policies of the 17th Portuguese Constitutional Government. Sísifo. Educational Sciences Journal [online]. 2009, č. 10, s. 53 – 64. Dostupné na: <https://sisifo.fpce.ul.pt>.

Ďalší významný vplyv výsledkov PISA na zmeny v portugalskej politike v oblasti vzdelávania nastal počas pôsobenia ministra školstva Nuno Crata (2012 až 2015). N. Crato nadviazal na prechádzajúce reformy a využil zjavnú stagnáciu výsledkov PISA v rokoch 2009 až 2012 na posilnenie kurikulárnych cieľov a výsledkov vzdelávania na základných a stredných školách, revíziu kurikulárnej štruktúry (zvýšenie počtu vyučovacích hodín portugalského jazyka, matematiky a prírodných vied) v rokoch 2013 – 2014, školskú autonómiu a na presadenie lepšej odbornej prípravy začínajúcich učiteľov. Osobitnú pozornosť venoval odbornému vzdelávaniu na stredných školách, čím diverzifikoval ponuku škôl, aby pokryli rôzne záujmy študentov (Crato, 2020⁶³). V rokoch 2014 – 2015 zaviedol aj skúšky s vysokými nárokmi (*high-stakes tests*) z matematiky a portugalského jazyka na konci cyklov – v 4. a 6. ročníku, ako aj hodnotenie kvalifikácie učiteľov a osvedčenie potrebné na vyučovanie. Testovanie v 4. a 6. ročníku, ako aj skúšky učiteľov napokon ukončila v roku 2016 novozvolená portugalská vláda. Kurikulárne reformy tohto ministra sledovali aj medzinárodné zosúladenie národných učebných osnov s tými, ktoré vyplývajú z rámcov PISA a najmä TIMSS (Marôco, 2020⁶⁴).

Pokrok Portugalska vo vzdelávaní

Viaceré štúdie preukázali, že zlepšenie výsledkov portugalských žiakov v štúdiu PISA 2015 bolo sprevádzané podobným zlepšením výsledkov v národných externých skúškach, čo poukazuje na ich vzájomnú súvislosť (napr. Marôco, 2018⁶⁵). Napriek tomu, že štúdia PISA sa nevyhla kritike, podobne ako na Slovensku, tvorcovia politik a verejnosť vo všeobecnosti využívajú túto štúdiu ako validný nástroj na porovnávanie efektívnosti vzdelávacích systémov a zavádzanie reforiem v oblasti vzdelávania (Breakspear, 2012⁶⁶; Phillips a Jiang, 2015⁶⁷). Pri prepojení politik s výsledkami PISA je však potrebné zohľadniť časový odstup medzi zmenou politiky a tým, kedy sa jej vplyv môže prejaviť vo výsledkoch PISA.

Zavedenie národných externých testovaní od roku 2005

Výsledky portugalských 15-ročných žiakov v testovaní PISA sa od roku 2005 výrazne zlepšili. Tento vývoj býva v odbornej literatúre interpretovaný ako indikátor pozitívneho smerovania zmien vo vzdelávacej politike, hoci priamu kauzálnu súvislosť medzi jednotlivými opatreniami a zlepšením výsledkov nie je možné jednoznačne preukázať.

Zavedenie externých národných testovaní z matematiky a portugalského jazyka v 9. ročníku ZŠ časovo korešponduje so zlepšením výsledkov PISA, pričom ide o korelačný, nie kauzálny vzťah. Ako uvádza vo svojom článku J. Marôco, pokiaľ ide o politiky v oblasti vzdelávania, viacerí autori sa zhodujú v tom, že príčinou zlepšenia je rozšírenie predškolského vzdelávania, diferenciacia pedagogických postupov, zlepšenie infraštruktúry škôl, podpora zosúladovania učebných osnov s medzinárodnými rámcami, zavedenie testov s vysokými nárokmi a rastúca ponuka odborných a profesijných kurzov pre žiakov mimo hlavného vzdelávacieho prúdu.

⁶³ CRATO, N. Curriculum and Educational Reforms in Portugal: An Analysis on Why and How Students' Knowledge and Skills Improved. In: REIMERS, F. M. (ed.). *Audacious Education Purposes: How Governments Transform the Goals of Education Systems* [online]. Cham: Springer, 2020, s. 209 – 231. Dostupné na: https://doi.org/10.1007/978-3-030-41882-3_8.

⁶⁴ MARÔCO, J. International large-scale assessments: Trends and effects on the Portuguese public education system. In: HARJU-LUUKKAINEN, H., McELVANY, N. a STANG, J. (eds.). *Monitoring Student Achievement in the 21st Century: European Policy Perspectives and Assessment Strategies* [online]. Cham: Springer, 2020, s. 207 – 222. Dostupné na: <https://doi.org/10.1007/978-3-030-38969-7>.

⁶⁵ MARÔCO, J. ...E pur si muove! Observador [online]. 11. január 2018. Dostupné na: <https://observador.pt/opiniao/e-pur-si-muove>.

⁶⁶ BREAKSPEAR, S. The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. *OECD Education Working Papers* [online]. 2012, č. 71, s. 1 – 31. Paris: OECD Publishing. Dostupné na: <https://doi.org/10.1787/5k9fdqffr28-en>.

⁶⁷ PHILLIPS, G. W. a JIANG, T. Using PISA as an International Benchmark in Standard Setting. *Journal of Applied Measurement* [online]. 2015. Dostupné na: <https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-26021-004&lang=pt-br&site=ehost-live&scope=site>.

Iní portugalskí výskumníci⁶⁸, ktorí sa zaoberali výsledkami štúdie PISA od roku 2000 do roku 2015, identifikovali nasledujúce hlavné faktory, ktoré vysvetľujú pozitívny vývoj Portugalska: (1) celkové výdavky na vzdelávanie v prepočte na obyvateľa sú porovnateľné s ostatnými členskými štátmi OECD, a to napriek tomu, že Portugalsko je z hľadiska OECD relatívne chudobnou krajinou – HDP na obyvateľa v roku 2018 bol v Portugalsku 49,8 tis. USD v porovnaní s cca 60,0 tis. USD v priemere OECD (OECD Data Explorer)⁶⁹; (2) pokrytie predškolskej výchovy takmer na 100 %; (3) primeraná špecifická a pedagogická príprava učiteľov, ich kompetencie a motivácia k vyučovaniu; (4) podpora žiakov zo strany rodičov a učiteľov, motivácia a vytrvalosť samotných žiakov; (5) školy v menej zvýhodnených ekonomických regiónoch dosahujúce výsledky nad očakávania a vzdelávacie projekty škôl zosúladené s regiónmi; a (6) úroveň vzdelania rodičov žiakov.

V prípade úspechu Portugalska je pozoruhodné, že napriek hospodárskej kríze v rokoch 2008 – 2013, keď sa národný HDP Portugalska znížil o 8 %, a celkovému negatívnemu vývoju výdavkov na vzdelávanie, Portugalsko dokázalo zlepšiť svoje celkové výsledky v štúdii PISA. Sekundárna analýza výsledkov PISA 2018 (Gomes a kol., 2019⁷⁰; Marôco, 2019⁷¹) ukázala, že hlavnými faktormi, ktoré podmieňujú výsledky portugalských žiakov, sú očakávania žiakov týkajúce sa ich budúceho povolania a sociálno-ekonomický a kultúrny status rodín.

Výsledky štúdie PISA 2000, ktoré sa vo vzťahu k Portugalsku nachádzali v spodnej časti rebríčka, neboli v Portugalsku prijaté až s takým „šokom“ ako v iných krajinách so slabými výsledkami (napr. v Nemecku). Portugalskí tvorcovia však tieto výsledky využili na zdôvodnenie a podporu reforiem na rôznych úrovniach systému. **Slabé výsledky štúdie PISA boli interpretované ako argument pre potrebu úpravy učebných osnov v kľúčových predmetoch**, konkrétne v portugalskom jazyku, matematike a prírodných vedách, a to na úrovni základného aj stredného vzdelávania. Národné programy boli zamerané na podporu čitateľských návykov a časové navýšenie vyučovania portugalského jazyka a matematiky. Na základe porovnania s ostatnými participujúcimi krajinami v štúdii PISA sa podporovali aj požiadavky na učiteľov a ich odbornú prípravu.

Najvýraznejší progres vo výsledkoch portugalskej štúdie PISA bol zaznamenaný medzi cyklami PISA 2006 a PISA 2009 (pozri obr. 1 – 3), teda po tom, ako boli v roku 2005 zavedené externé testy z matematiky a portugalského jazyka pre 9. ročník. Aj pokiaľ ide o vývoj štúdie PISA v rokoch 2015 až 2018, prírodné vedy boli jedinou oblasťou so štatisticky významným poklesom, pričom ide o oblasť, pre ktorú neexistuje ekvivalentné národné externé hodnotenie. Tento fakt býva niektorými autormi interpretovaný ako možný indikátor významu systematického hodnotenia, hoci takýto vzťah nie je možné jednoznačne potvrdiť. V niektorých krajinách bol pozorovaný súbeh zavedenia *high-stakes* testovania a zmien vo výsledkoch štúdie PISA, čo viedlo k úvahám o možnej súvislosti medzi týmito javmi (Bergbauer a kol., 2018⁷²). Tieto zistenia však neumožňujú jednoznačne vyvodiť kauzálny vzťah.

⁶⁸ FERREIRA, A. S., FLORES, I. a CASAS-NOVAS, T. Porque melhoraram os resultados PISA em Portugal? Estudo longitudinal e comparado (2000-2015) [online]. Lisboa: Fundação Francisco Manuel dos Santos, 2017. ISBN 978-989-8863-36-2. Dostupné na: <https://ffms.pt/sites/default/files/2022-08/introducao-ao-estudo-porque-melhoraram-os-resultados-pisa-em-portugal.pdf>.

⁶⁹ OECD Data Explorer, Annual GDP and consumption per capita, údaje za rok 2023. Dostupné na: https://data-explorer.oecd.org/vis?df%5Bdf%5D=OECD.SDD.NAD&df%5Bdf%5D=dsDisseminateFinalDMZ&df%5Bdf%5D=DSD_NAMAIN10%40DF_TABLE1_EXPENDITURE_HCPC.

⁷⁰ GOMES, J. F., HORMIGO, I. a MARÔCO, J. Primeiras lições do PISA: dois países a seguir e Portugal no mundo. Iniciativa Educação [online]. 3. november 2019. Dostupné na: <https://www.iniciativaeducacao.org/pt/ed-on/artigos/estatisticas/primeiras-licoes-do-pisa-dois-paises-a-seguir-e-portugal-no-mundo>.

⁷¹ MARÔCO, J. PISA 2018: Portugal é ainda a maior história de sucesso na Europa? Iniciativa Educação [online]. 2. november 2019. Dostupné na: <https://www.iniciativaeducacao.org/pt/ed-on/artigos/estatisticas/pisa-2018>.

⁷² BERGBAUER, A. B., HANUSHEK, E. A. a WOESSMANN, L. Testing. NBER Working Paper [online]. 2018, č. 24836. Cambridge, MA: National Bureau of Economic Research. Dostupné na: <https://doi.org/10.3386/w24836>.

Zavedenie testovania s vysokými dôsledkami patrí medzi faktory, ktoré mohli prispieť k zlepšeniu výsledkov v testovaniach PISA. Vzhľadom na súbežné pôsobenie viacerých reforiem (napr. kurikulárne zmeny, podpora učiteľov, rozšírenie predškolského vzdelávania) však nie je možné určiť jeho relatívnu váhu ani ho označiť za jednoznačne rozhodujúci faktor. Dôležitú úlohu mohla zohrávať aj reforma učebných osnov, vzdelávacích cieľov a organizácie vo vyučovaní a triedach. Zmeny v nastavení a štruktúrovaní národného kurikula boli v rámci **komplexného uplatnenia novej vzdelávacej politiky sprevádzané zavedením pravidelného systému národného vyhodnocovania výsledkov vzdelávania po každom vzdelávacom cykle (po štvrtom, šiestom a deviatom ročníku), úpravami vo vzdelávaní a hodnotení učiteľov, ako aj v dômyselných stratégiách tvorby a vydávania učebníc.**⁷³

Zmeny vzdelávacích politík po roku 2015

S nástupom nasledujúcej vlády (2015 – 2019) však došlo k zásadným zmenám, ktoré zrušili externé testovanie v štvrtom a šiestom ročníku aj overovacie skúšky učiteľov. Nový minister školstva⁷⁴ zrušil *high-stakes* testy pre štvrtý a šiesty ročník, tie boli nahradené diagnostickými *low-stakes* testami pre 2., 5. a 8. ročník z portugalského jazyka, matematiky a niekoľkých ďalších predmetov. Nový minister tiež zaviedol flexibilitu učebných osnov a zvýšil dôraz na rozvoj kľúčových kompetencií žiakov. Podľa niektorých odborníkov dôsledkom zrušenia externých *high-stakes* testov a zníženia náročnosti vzdelávacích cieľov môže byť samotné zníženie výkonov žiakov v testoch PISA. Až traja zo štyroch portugalských žiakov uviedli, že na test PISA 2018 vynaložili menej úsilia, ako keby sa test započítaval do ich známok. Treba však uviesť, že ide o údaj vychádzajúci zo subjektívneho hodnotenia žiakov. Po prvýkrát zároveň klesla miera účasti portugalských žiakov (76 %) pod štandardom štúdie PISA, ktorý bol 80 %.

Aktuálny vývoj po roku 2024 priniesol ďalšiu zmenu modelu externého hodnotenia. Portugalsko zaviedlo ModA v 4. a 6. ročníku, posilnilo hodnotenie na konci jednotlivých cyklov, porovnateľnosť výsledkov v čase prostredníctvom kotviacich položiek a digitálnu administráciu a klasifikáciu.

Napriek pozitívnemu vývoju portugalských žiakov štúdia PISA ukazuje, že výsledky žiakov sú v rámci krajiny výrazne asymetrické. Rozdiel vo výsledkoch medzi najúspešnejšími a najmenej úspešnými regiónmi v krajine zodpovedá približne dvom rokom školskej dochádzky. Očakávania žiakov vo vzťahu k ich kariére a sociálno-ekonomické postavenierodín sú hlavnými determinantmi výsledkov portugalských žiakov. Tieto vplyvy zostávajú prítomné vo všetkých cykloch štúdie PISA.

Celkovo sa však portugalský vzdelávací systém, ktorý v období 2003 – 2006 zaostával aj za slovenským systémom vzdelávania (PISA – matematická a prírodovedná gramotnosť), dostal podľa meradla štúdie PISA na úroveň priemeru krajín OECD. Napriek tomuto evidentnému progresu v období po roku 2011 sa Portugalsko vo vzdelávacej a kurikulárnej politike posunulo smerom k väčšej flexibilitě kurikula. Nasledujúce cykly štúdie PISA ukážu, či sa Portugalsko ubera správnym smerom.

⁷³ PUPALA, B. Portugalská cesta k úspechu: Konzervatívny oblúk v kurikulárnej politike. Pedagogické rozhľady [online]. 2022, č. 3, s. 27 – 32. Banská Bystrica: NIVAM. Dostupné na: <https://nivam.sk/wp-content/uploads/2025/02/pr-3-2022.pdf>.

⁷⁴ Poznámka: Tiago Brandão Rodrigues (funkčné obdobie 2015 – 2019).

Syntetické zhrnutie – Portugalsko (a prínosy pre Slovensko)

Portugalský vzdelávací systém predstavuje príklad krajiny, ktorá dokázala v relatívne krátkom časovom období výrazne zlepšiť vzdelávacie výsledky žiakov prostredníctvom systematickej práce s dátami a dôsledného prepojenia hodnotenia so vzdelávacou politikou. Významnú úlohu v tomto procese zohralo medzinárodné meranie PISA, ktoré portugalskí tvorcovia politik vyžívali ako referenčný rámec a východisko pre rozhodovanie založené na dôkazoch.

Kľúčovým prvkom portugalského prístupu bolo zavedenie národných externých testovaní, najmä *high-stakes* testov z matematiky a portugalského jazyka na konci nižšieho sekundárneho vzdelávania. Zavedenie týchto testov sa časovo prekrývalo s najvýraznejším zlepšením výsledkov portugalských žiakov v štúdiu PISA medzi cyklami v rokoch 2006 a 2009, čo naznačuje možnú súvislosť medzi týmito opatreniami a vývojom výsledkov. Skúsenosť Portugalska zároveň ukazuje, že pozitívny vývoj výsledkov nebol dôsledkom samotného testovania, ale kombinácie viacerých opatrení, medzi ktoré patrili úpravy učebných osnov, podpora predškolského vzdelávania, posilnenie kľúčových predmetov a zlepšenie prípravy učiteľov.

Neskorší posun od *high-stakes* testov k diagnostickým a formatívne orientovaným *low-stakes* testom poukazuje na **snahu hľadať rovnováhu medzi jasne definovanými nárokmi na výkon žiakov a podporou autonómie škôl**. Zároveň oslabenie motivačnej funkcie externého hodnotenia môže byť jedným z faktorov spájaných so stagnáciou alebo miernym poklesom výkonov, pričom tento vzťah je potrebné interpretovať opatrne vzhľadom na pôsobenie viacerých paralelných zmien vo vzdelávacom systéme.

Pre Slovensko je portugalská skúsenosť cenná najmä v tom, že dokumentuje význam politickej kontinuity, konzistentného smerovania reforiem a jasného účelu externého hodnotenia. Ukazuje, že **národné testovanie môže pôsobiť ako silný impulz zmeny, ak je súčasťou širšej reformnej stratégie**, a ak existuje zhoda na jeho význame medzi tvorcami politik, odbornou verejnosťou a školami. Zároveň upozorňuje na riziká spojené s častými zmenami koncepcie hodnotenia a na **potrebu citlivého vyvažovania medzi formatívnou podporou učenia a motiváciou k dosahovaniu vysokých výkonov**.

Limity prenositeľnosti portugalského modelu

Pri interpretácii portugalského modelu externého hodnotenia a jeho možnej aplikovateľnosti v podmienkach Slovenska je potrebné zohľadniť viacero kontextových a systémových faktorov, ktoré významne ovplyvňujú jeho fungovanie a výsledky.

Jedným z kľúčových limitov je **komplexnosť a súbežnosť realizovaných reforiem**. Zlepšenie výsledkov portugalských žiakov v medzinárodných meraniach nemožno pripísať jednému konkrétnemu opatreniu, ako je napríklad zavedenie externého testovania, ale skôr kombinácii viacerých paralelne realizovaných zmien. Išlo najmä o reformy kurikula, posilnenie výučby kľúčových predmetov, rozšírenie predškolského vzdelávania, zlepšenie prípravy učiteľov a cielenú podporu znevýhodnených žiakov. Prenos samotného modelu testovania bez týchto sprievodných opatrení by preto pravdepodobne neviedol k porovnateľným výsledkom.

Významným faktorom je aj politická kontinuita a stabilita reformného smerovania. Portugalsko realizovalo reformy v oblasti hodnotenia a kurikula v relatívne dlhšom časovom horizonte, pričom aj napriek čiastočným zmenám po roku 2015 si systém zachoval určitú mieru kontinuity. V podmienkach Slovenska,

kde dochádza k častejším zmenám politických priorít a koncepcií, môže byť udržanie dlhodobej reformnej línie náročnejšie, čo znižuje predpoklady pre dosiahnutie podobných efektov.

Ďalším limitom je odlišný vzťah medzi externým hodnotením a motiváciou žiakov a škôl. Portugalský model v určitom období využíval aj *high-stakes* testovania, čo mohlo posilňovať motiváciu žiakov k dosahovaniu výkonu v kľúčových predmetoch. Prenos takéhoto prístupu do slovenského prostredia by však mohol viesť aj k **nežiaducim efektom, ako je zvýšený tlak na výkon, zužovanie kurikula alebo orientácia na „učenie pre testy“, najmä ak by nebol vyvážený silnými prvkami formatívneho hodnotenia.**

Limitujúcim faktorom je aj úroveň rozvoja podporných mechanizmov v systéme. Portugalský model je založený na využívaní dát z externého hodnotenia na úrovni škôl a tried, pričom tieto dáta sú interpretované a využívané v kontexte pedagogickej podpory, profesijného rozvoja učiteľov a intervenčných opatrení. V podmienkach Slovenska zatiaľ nie je systematické využívanie dát na tejto úrovni dostatočne rozvinuté, čo môže obmedziť efektívnosť podobných nástrojov.

Dôležitým aspektom je aj sociálno-ekonomický a kultúrny kontext, najmä pokiaľ ide o očakávania žiakov, podporu zo strany rodín a postavenie vzdelávania v spoločnosti. Výskumy poukazujú na to, že práve tieto faktory zohrávajú významnú úlohu pri vysvetľovaní výsledkov žiakov v Portugalsku. Prenos systémových opatrení bez zohľadnenia týchto širších determinantov môže viesť k obmedzeným alebo odlišným efektom.

Z uvedených dôvodov možno konštatovať, že portugalský model externého hodnotenia poskytuje cenné inšpirácie najmä **v oblasti systematickej práce s dátami, prepojenia hodnotenia s tvorbou vzdelávacej politiky a kombinácie rôznych typov testovania.** Jeho prenos do slovenského prostredia si však vyžaduje dôkladnú adaptáciu, zohľadňujúcu špecifiká národného kontextu, úroveň pripravenosti systému a potrebu zabezpečenia dlhodobej stability a podpory realizovaných zmien.

2. Metodologická príprava a vývoj kritériálnych testov na Slovensku

Táto kapitola približuje teoretické a metodologické východiská kritériálneho testovania a opisuje ich postupné uplatňovanie v slovenskom vzdelávacom systéme. Jej cieľom je vysvetliť, na akých konceptuálnych základoch je kritériálne hodnotenie postavené, v čom sa odlišuje od normatívneho testovania a aký účel plní v kontexte kurikulárnej reformy. Ťažisko kapitoly je kladené na medzinárodné teoretické východiská, metodologickú prípravu kritériálnych testov na Slovensku a na ich využitie ako nástroja diagnostiky, formatívneho hodnotenia a podpory učenia.

2.1 Teoretické a metodologické východiská kritériálneho testovania

Kritériálne orientované hodnotenie sa formovalo v kontexte rozvoja pedagogického merania v druhej polovici 20. storočia a rastúcej potreby presnejšie posudzovať dosahovanie vzdelávacích cieľov. Jeho vznik sa spája najmä s hľadaním alternatívy k normatívnemu hodnoteniu, ktoré interpretuje výkon jednotlivca vo vzťahu k výkonu ostatných žiakov, a s úsilím posudzovať výsledky učenia vzhľadom na vopred definované očakávania a požiadavky na výkon.

Teoretické základy kritériálneho testovania vychádzajú najmä z prác Glasera (1963)⁷⁵, ktorý patril k prvým autorom systematicky rozpracúvajúcim koncept kritériálne orientovaného, merania a z koncepcie zvládnutého učenia (*mastery learning*) formulovanej Bloomom (1968)⁷⁶. Tento prístup predpokladá, že výkon žiaka má byť posudzovaný vo vzťahu k vopred definovaným kritériám alebo štandardom, ktoré reprezentujú očakávanú úroveň vedomostí, zručností a kompetencií, a nie vo vzťahu k výkonu ostatných žiakov (Glaser, 1963; Bloom, 1968).

V nasledujúcich desaťročiach sa kritériálne testovanie etablovalo ako významný prístup k hodnoteniu výsledkov učenia, najmä v prostredí vzdelávania orientovaného na dosahovanie vzdelávacích cieľov a štandardov. Na rozdiel od normatívneho hodnotenia, ktoré poskytuje informáciu o relatívnej pozícii žiaka v rámci populácie, kritériálne testovanie umožňuje interpretovať výsledky z hľadiska miery dosiahnutia konkrétnych vzdelávacích cieľov. Rozvoj tohto prístupu bol sprevádzaný aj vytváraním špecifických metodologických a psychometrických postupov, najmä v oblasti stanovovania výkonových štandardov a hraničných skóre (*cut scores*), ktoré určujú úrovne zvládnutia hodnotenej oblasti (Cizek a Bunch, 2007)⁷⁷.

Kritériálne testovanie (ďalej aj „CR-testovanie“) vychádza z požiadavky hodnotiť výkon žiaka vo vzťahu k vopred vymedzeným vzdelávacím cieľom a kritériám ich dosiahnutia. Jeho historické formovanie, ako aj jeho súčasné rozvíjanie možno zasadiť do širšieho rámca pedagogických a psychologických teórií učenia a kognitívneho vývinu:

- **Behaviorálne teórie učenia** a z nich vychádzajúce prístupy k formulovaniu vzdelávacích cieľov zdôraznili potrebu vyjadrovať ciele prostredníctvom pozorovateľných a hodnotiteľných výkonov. Na tento princíp nadväzuje kritériálne testovanie tým, že zisťuje, do akej miery žiak preukázal konkrétne vymedzené vedomosti, zručnosti alebo kompetencie vo vzťahu k očakávaným vzdelávacím výstupom (Mager, 1962)⁷⁸.

⁷⁵ GLASER, R. Instructional Technology and the Measurement of Learning Outcomes: Some Questions. American Psychologist [online]. 1963, roč. 18, č. 8, s. 519 – 521. Dostupné na: <https://doi.org/10.1037/h0049294>.

⁷⁶ BLOOM, B. S. Learning for mastery. Evaluation Comment. 1968, roč. 1, č. 2, s. 1 – 12.

⁷⁷ CIZEK, G. J. a BUNCH, M. B. Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests. Thousand Oaks, CA: Sage Publications, 2007. ISBN 978-1-4129-1683-7.

⁷⁸ MAGER, R. F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962. 62 s.

- Hoci **teórie kognitívneho vývinu** nepredstavujú priame teoretické základy kritériálneho testovania, poskytujú dôležité východiská pre tvorbu a interpretáciu kritériálnych testov. Upozorňujú, že spôsob porozumenia, uvažovania a riešenia problémov sa v priebehu vývinu dieťaťa mení. Pri tvorbe kritériálnych testov je preto potrebné zosúladiť obsah, formu a kognitívnu náročnosť úloh s vekom, skúsenosťami a vývinovými osobitosťami cieľovej skupiny. Výkon žiaka nemožno interpretovať bez zohľadnenia toho, aké myšlienkové operácie a úroveň porozumenia si riešenie úlohy vyžaduje (Piaget a Inhelderová, 2014; Helus, 2009)⁷⁹.

- **Konštruktivisticky orientované prístupy** síce nepredstavujú východisko historického vzniku kritériálneho testovania, významne však ovplyvnili jeho súčasné podoby, najmä v oblasti autentického hodnotenia, výkonových úloh a posudzovania komplexnejších vzdelávacích výstupov. Učenie pritom chápu ako aktívne vytváranie a pretváranie poznania na základe predchádzajúcich skúseností žiaka a jeho interakcie so sociálnym a kultúrnym prostredím.

V súčasných prístupoch ku kritériálnemu hodnoteniu sa preto pozornosť neobmedzuje len na reprodukciu osvojených poznatkov, ale zahŕňa aj úlohy vyžadujúce vysvetľovanie, prepájanie súvislostí, aplikáciu poznatkov a riešenie problémov v zmysluplných alebo autentických situáciách.

Takéto úlohy umožňujú posudzovať nielen správnosť výsledku, ale aj kvalitu porozumenia a úroveň dosiahnutia vymedzených vzdelávacích cieľov (Vygotskij, 2017; Průcha, 2020)⁸⁰.

Uvedené teoretické východiská nemajú vo vzťahu ku kritériálnemu testovaniu rovnaké postavenie. Behaviorálne koncepcie a operacionalizácia vzdelávacích cieľov ovplyvnili predovšetkým jeho historické formovanie, zatiaľ čo kognitívne a konštruktivistické prístupy rozšírili požiadavky na obsahovú a kognitívnu kvalitu úloh a na interpretáciu výkonu žiaka. Súčasný kritériálny testovanie preto predstavuje konceptuálne a metodologicky osobitý prístup k hodnoteniu výsledkov vzdelávania, ktorý prepája jasne vymedzené vzdelávacie ciele a kritériá ich dosiahnutia, obsahovo ukotvené hodnotiace nástroje a interpretáciu výkonu žiaka vo vzťahu k týmto cieľom a kritériám.

2.2 Metodologická príprava CR-testovaní na Slovensku

Medzinárodná spolupráca

Plán obnovy a odolnosti (2021 – 2026)⁸¹ umožňuje zavádzať na Slovensku významné reformy s celospoločenskými dopadmi. Jednou zo zmien v oblasti vzdelávania sú aj inovácie v externých testovaniach a hodnoteniach žiakov (Komponent 7, Reforma 1) s cieľom zvýšiť kvalitu vzdelávacieho procesu, ktorého súčasťou je poskytovanie adresnejšej spätnej väzby žiakom a školám. Popri už zavedenom normatívnom testovaní sa na slovenských školách v kontexte kurikulárnej reformy postupne overujú a zavádzajú kritériálne testy (ďalej aj „CR-testy“ alebo „CRT“).

Tento proces nadväzuje na systematickú metodologickú prípravu vo vtedajšom Národnom ústave certifikovaných meraní vzdelávania (NÚCEM), ktorá bola spojená s vyškolením slovenských odborníkov v spolupráci s významnými výskumnými pracoviskami v Nemecku a Estónsku. Dôvodom tohto úsilia bola absencia externého merania založeného na kritériálnej interpretácii výsledkov. Takéto meranie umožňuje zisťovať, či a do akej miery žiak zvládol konkrétny súbor vedomostí a zručností definovaných

⁷⁹ PIAGET, J. a INHELDEROVÁ, B. Psychologie dítěte. Praha: Portál, 2014; Helus, Zdeněk. *Dítě v osobnostním pojetí*. 2., přepracované a rozšířené vydání. Praha: Portál, 2009.

⁸⁰ VYGOTSKIJ, L. S. Psychologie myšlení a řeči. 2., upravené vydání. Praha: Portál, 2017; Průcha, Jan. *Psychologie učení: teoretické a výzkumné poznatky pro edukační praxi*. Praha: Grada, 2020.

⁸¹ ÚRAD VLÁDY SR. Plán obnovy: Domov [online]. 2021. Dostupné na: <https://www.planobnovy.sk>.

vo vzdelávacích štandardoch, teda či dosiahol vopred stanovené vzdelávacie ciele, nezávisle od výkonu ostatných žiakov. Práve na tento účel sa využívajú kritériálne testy.

Skúsenosti Nemecka a Estónska výrazne prispeli k tomu, že Slovensko začalo systematicky budovať vlastný model kritériálneho testovania. Medzinárodná expertíza zo Stuttgartskej univerzity a z estónskeho centra Innove umožnila NÚCEM-u od roku 2019 vyvíjať prvé ucelené prototypy týchto testov v rámci európskeho reformného programu. Táto spolupráca vytvorila základ, na ktorom dnes môže NIVaM stavať a pokračovať tak v zavádzaní a rozvíjaní metodiky kritériálneho hodnotenia vo vzdelávaní.

Porovnanie kritériálnych a normatívnych testov z hľadiska účelu

V slovenskom vzdelávacom systéme sú na základných školách zavedené od roku 2005 externé celoplošné merania vzdelávacích výsledkov žiakov 9. ročníka realizované s cieľom dosiahnuť objektivitu a diferencovanosť hodnotenia žiackych výkonov. Doposiaľ realizované externé testovania žiakov na základných školách (**Testovanie 9, Testovanie 5**) hodnotia žiaka z pohľadu jeho relatívneho výkonu, teda v porovnaní s priemerom a percentilom testovanej populácie žiakov. Na tento účel sa využívajú práve normatívne (rozlišovacie) testy⁸². Takéto testy môžu obsahovať menší počet položiek, pokrývajú väčší obsah učiva a sú zadávané pri výstupe zo vzdelávania na základnej škole. Normatívne testy preto neposkytujú priamu a detailnú interpretáciu miery zvládnutia konkrétnych štandardov, a preto doterajšie výstupy z celoplošných testovaní poskytujú do značnej miery limitovanú spätnú väzbu z testovaných predmetov pre žiakov aj učiteľov.

Testovanie 9 (predtým MONITOR 9)⁸³, ako aj externá časť maturitnej skúšky môžu predstavovať v slovenskom vzdelávacom systéme *high-stakes testy*, teda testy s vysokými dôsledkami na vzdelávaciu dráhu žiakov. V súčasnosti sa pripravujú zmeny smerujúce k využitiu výsledkov externého testovania pri prijímaní na stredné školy; ich konkrétna podoba a harmonogram budú závisieť od schválenej právnej úpravy.⁸⁴ Podobný vývoj je aj napr. v Českej republike.⁸⁵

Jedna z reforiem Plánu obnovy a odolnosti v oblasti vzdelávania bola zameraná na zavedenie metodologicky iného typu merania s odlišne nastavenou interpretáciou vzdelávacích výsledkov žiakov, na kritériálne testovanie, ktoré napĺňa odlišný účel v porovnaní s normatívnym testovaním.

Odlišnosti medzi normatívnym a kritériálnym testovaním môže z hľadiska prehľadnosti lepšie priblížiť ich porovnanie podľa účelu.

⁸² Poznámka: Viac o normatívnych testoch tu: KUBIŠ, T. a kol. 2015. Metodika tvorby testových úloh a testov [online]. Dostupné na: <https://www.etest.sk/data/att/fdd/609.2e2bed.pdf>.

⁸³ Poznámka: V rokoch 2003 a 2004 bol MONITOR 9 realizovaný na monitorovanie vedomostí a zručností žiakov končiacich 9. ročník ZŠ ako pilotný projekt na reprezentatívnej vzorke žiakov. Bližšie informácie: <https://www2.nucem.sk/sk/merania/narodne-merania/testovanie-9/o-testovani-9>.

⁸⁴ Poznámka: Podľa novely zákona 245/2008 Z. z. (školský zákon) sa upravuje jednotná prijímacia skúška ako externe zadávané a centrálné vyhodnocované testy od školského roku 2027/2028. Mali by sa tým odstrániť testy realizované na jednotlivých stredných školách, ktoré rovnako ako Testovanie 9 overovali u žiakov vedomosti z matematiky a vyučovacích jazykov.

⁸⁵ STATNIPRIJIMACKY.CZ. Přijímačky Cermat: testy, řešení, tipy [online]. Dostupné na: <https://www.statniprijimacky.cz/cermat>.

Tab. 3 Porovnanie kritériálnych testov (CRT) a normatívnych testov (NRT) podľa ich účelu a charakteristík⁸⁶

Charakteristika / Účel	Kritériálne testy (CRT)	Normatívne testy (NRT)
Primárny účel	Meranie zvládnutia učiva podľa vopred stanovených kritérií alebo štandardov	Porovnanie výkonu žiaka s výkonom ostatných žiakov (norma skupiny)
Otázka, ktorú test zodpovedá	„Vie žiak to, čo má vedieť?“	„Kde sa žiak v porovnaní s ostatnými nachádza?“
Meranie zvládnutia učiva	Áno – zisťuje úroveň zvládnutia špecifikovaného učebného obsahu, zručností, kompetencií	Čiastočne – úroveň zvládnutia učiva nie je primárnou funkciou NR testov, zisťuje sa relatívna pozícia žiaka oproti výsledkom ostatných testovaných žiakov
Diagnostika a intervencie	Identifikácia medzier vo vedomostiach, zručnostiach, podklad pre individualizovanú výučbu	Menej vhodné na diagnostiku, slúži skôr na selekciu alebo klasifikáciu
Plánovanie výučby	Výsledky testov pomáhajú upraviť učebné osnovy, individualizovať učenie	Neposkytuje detailné informácie pre úpravu výučby
Stanovenie úrovne zvládnutia	Porovnanie s kritériami a štandardmi, nie s inými žiakmi	Stanovenie poradia v skupine (percentilové poradie)
Výber, triedenie, porovnávanie	Nie je typickým účelom	Áno – používa sa pri prijímacích skúškach, identifikácii talentovaných žiakov
Stanovenie poradia	Nie – hodnotí sa absolútne zvládnutie testu	Áno – určuje „najlepší“, „priemerný“, „podpriemerný“ výkon v skupine
Štatistické porovnanie	Poskytuje informácie o zvládnutí štandardov; sekundárne môže slúžiť ako podklad na úpravu kurikúl	Primárne určené na porovnávanie žiakov, škôl, regiónov, ročníkov alebo populácií
Identifikácia krajných hodnôt (extrémov)	Skôr na zistenie nedostatkov v učive	Áno – odhaľuje žiakov s nadpriemernými schopnosťami alebo tých, ktorí potrebujú podporu
Typické použitie	Certifikácia, formatívne hodnotenie, diagnostika	Prijímacie skúšky, IQ testy, štandardizované testy

Ilustračný príklad rozdielov medzi kritériálnym a normatívnym testovaním

Rozdiel medzi kritériálnym a normatívnym testovaním možno názorne ilustrovať na posudzovaní odbornej spôsobilosti elektrikára. Pri tomto posudzovaní nie je rozhodujúce, aké miesto elektrikár zaujíma medzi ostatnými príslušníkmi profesie, ale či dokáže vykonať elektroinštalačné práce správne a bezpečne, v súlade so stanovenými technickými postupmi a bezpečnostnými požiadavkami. Jeho výkon sa teda posudzuje predovšetkým vo vzťahu k vopred vymedzeným odborným kritériám, nie podľa jeho relatívneho postavenia medzi ostatnými elektrikármi.

⁸⁶ POPHAM, W. J. Classroom Assessment: What Teachers Need to Know. 8. vyd. Boston: Pearson, 2017. ISBN 978-0-13-405386-8. (adaptované)

Analogicky vo vzdelávaní odpovedá kritériálne testovanie predovšetkým na otázku, do akej miery žiak dosiahol vymedzené vzdelávacie ciele alebo výkonové štandardy. Normatívne testovanie naproti tomu poskytuje informáciu najmä o relatívnom postavení žiaka v porovnaní s výsledkami referenčnej skupiny. Kým normatívna interpretácia môže ukázať, že žiak dosiahol lepší výsledok než určitý podiel jeho rovesníkov, kritériálna interpretácia sa usiluje opísať, ktoré vedomosti a zručnosti žiak pravdepodobne ovláda a na akej úrovni.

Oba prístupy predstavujú rozdielne spôsoby interpretácie výsledkov a ich vhodnosť závisí od účelu hodnotenia. Normatívna interpretácia je užitočná napríklad pri **zisťovaní relatívneho postavenia žiaka alebo pri výbere obmedzeného počtu uchádzačov**. Kritériálna interpretácia je vhodná vtedy, keď potrebujeme **posúdiť dosiahnutie konkrétnych vzdelávacích cieľov, identifikovať silné a slabšie stránky výkonu alebo určiť ďalšie kroky učenia**. Nejde preto o vzájomne sa vylučujúce prístupy; výsledok toho istého testovania možno za určitých podmienok interpretovať z oboch hľadísk.

Kritériálne zaradenie žiaka do výkonovej úrovne však nemožno považovať za absolútne potvrdenie alebo vyvrátenie jeho kompetentnosti. Výsledok je vždy spojený s určitou **mierou neistoty** vyplývajúcou z obmedzeného rozsahu testu, chyby merania a spôsobu stanovenia hraničného skóre. Najvyššie riziko nesprávnej klasifikácie vzniká pri výsledkoch nachádzajúcich sa v blízkosti hranice medzi dvoma výkonovými úrovňami. Pri ich interpretácii je preto potrebné zohľadniť neistotu odhadu a podľa významu rozhodnutia výsledok doplniť ďalšími dôkazmi o výkone žiaka.

Presnosť merania možno podporiť využitím **modelov teórie odpovede na položku** (IRT), ktoré umožňujú odhadovať úroveň spôsobilosti žiaka aj mieru neistoty tohto odhadu na základe vlastností jednotlivých položiek. IRT a kritériálne hodnotenie pritom nepredstavujú konkurenčné prístupy. IRT poskytuje model merania a odhadu spôsobilosti, zatiaľ čo kritériálne ukotvenie umožňuje interpretovať tento odhad vo vzťahu k obsahovo vymedzeným vzdelávacím cieľom a výkonovým úrovňam.

Kritériálne orientované výsledky môžu vhodne podporovať aj formatívne hodnotenie, pokiaľ nie sú redukované na celkové skóre alebo jednoduché zaradenie žiaka. Učiteľovi môžu napríklad ukázať, ktoré čiastkové ciele žiak už dosiahol, pri ktorých typoch úloh má ťažkosti a aká následná podpora je primeraná. Žiak môže dostať konkrétnu informáciu o tom, čo už zvládol, v čom sa potrebuje zlepšiť a aký má byť jeho ďalší krok. **Formatívny účinok** však nevyplýva automaticky z kritériálnej povahy testu, ale **zo spôsobu, akým sa získané informácie využijú na úpravu vyučovania a ďalšieho učenia**.

Box 3: Kritériálne testovanie a teória odpovede na položku (IRT)

Kritériálne testovanie (*Criterion-Referenced Testing*, CRT) a teória odpovede na položku (*Item Response Theory*, IRT) označujú dva odlišné, ale navzájom sa dopĺňajúce koncepty.

Kritériálne testovanie predstavuje predovšetkým spôsob interpretácie výsledkov testovania, ktorý ovplyvňuje aj obsahový rámec a konštrukciu testu. Výkon žiaka sa pri ňom posudzuje vo vzťahu k vopred stanoveným vzdelávacím cieľom, štandardom alebo výkonovým úrovňam. Základnou otázkou kritériálneho testovania je: *Do akej miery žiak dosiahol požadované vzdelávacie ciele alebo úroveň kompetencie?*

IRT predstavuje psychometrický rámec zahŕňajúci modely, ktoré opisujú pravdepodobnosť určitej odpovede žiaka na testovú položku v závislosti od úrovne meranej latentnej schopnosti a vlastností položky. Podľa zvoleného modelu môže ísť najmä o obťažnosť položky, jej diskriminačnú schopnosť, prípadne parameter hľadania. Za podmienky primeranosti zvoleného modelu a splnenia jeho predpokladov umožňuje IRT umiestniť odhady schopností žiakov a parametre položiek na spoločnú škálu.

Využívanie modelov IRT môže:

- podporiť prepájanie rôznych testových foriem a testových zošitov;
- pri vhodnom dizajne prepojenia podporiť vytváranie a udržiavanie porovnateľných výkonových škál v čase;
- poskytovať informáciu o presnosti merania pri rôznych úrovniach schopnosti;
- vytvárať metodologické predpoklady na počítačovo adaptívne testovanie;
- podporiť efektívne využívanie kalibrovaných bánk testových úloh.

Kritériálna interpretácia výsledkov nevyžaduje použitie IRT. Kritériálny test možno vyhodnocovať aj pomocou postupov klasickej teórie testov (CTT). V rozsiahlych systémoch hodnotenia však môže spojenie kritériálnej interpretácie s vhodne zvolenými modelmi IRT podporiť porovnateľnosť testových foriem, vyjadrovanie neistoty merania a stabilitu výkonovej škály. Uvedené prínosy závisia od kvality testu a dát, primeranosti zvoleného modelu, splnenia jeho predpokladov a kvality použitého dizajnu prepojenia.

Elektronické testovanie možno realizovať aj bez použitia IRT. V dlhodobej perspektíve však môže budovanie bánk úloh, ktorých položky sú kalibrované pomocou IRT, vytvoriť podmienky na zavádzanie počítačovo adaptívneho testovania. Pri adaptívnom testovaní sa výber nasledujúcich úloh prispôsobuje priebežne aktualizovanému odhadu úrovne meranej schopnosti žiaka. Jeho zavedenie si vyžaduje dostatočne rozsiahlu, obsahovo vyváženú a kvalitne kalibrovanú banku položiek, vhodnú technologickú infraštruktúru a stabilné odborné kapacity.

Zjednodušene povedané: kým CRT odpovedá predovšetkým na otázku *Ako interpretujeme výsledok?*, IRT odpovedá na otázku *Ako modelujeme odpovede a odhadujeme meranú schopnosť?*

3. Metodika tvorby kritériálnych testov

Nasledujúca časť predstavuje základné princípy a kroky pri tvorbe kritériálnych testov. Metodika vychádza z tréningových materiálov odborníkov z Univerzity v Stuttgarte, ktoré boli adaptované na slovenské podmienky a postupne sa zavádzajú do systému externých meraní na Slovensku.

Medzi základné princípy kritériálneho testovania patrí predovšetkým:

- **Obsahové a kurikulárne ukotvenie:** Test musí byť navrhnutý tak, aby overoval konkrétne vedomosti, zručnosti alebo kompetencie vymedzené vo vzdelávacích cieľoch a štandardoch. Súbor testových úloh má zároveň primerane reprezentovať jednotlivé oblasti hodnoteného výkonu tak, aby poskytoval dostatočné dôkazy o miere dosiahnutia stanovených cieľov.
- **Transparentnosť kritérií:** Vzdelávacie ciele, očakávané výkony a kritériá ich dosiahnutia majú byť jasne vymedzené a zrozumiteľné učiteľom aj žiakom. Spôsob interpretácie výsledkov vrátane vymedzenia výkonových úrovní musí byť stanovený vopred.
- **Interpretácia nezávislá od výkonu ostatných žiakov:** Výsledok žiaka sa interpretuje vo vzťahu k vopred stanoveným kritériám dosiahnutia vzdelávacích cieľov, nie podľa jeho relatívneho postavenia v porovnaní s ostatnými žiakmi. Tým sa kritériálne orientované testovanie odlišuje od normatívne orientovaného testovania.

3.1 Základné kroky pri vývoji kritériálnych testov

Pri vývoji kritériálnych testov by sa mali uplatňovať nasledujúce kroky:

1. Definovanie vzdelávacích cieľov

Východiskom kritériálneho testovania sú jasne vymedzené vzdelávacie ciele odvodené od obsahového a výkonového štandardu. Na ich základe sa tvoria testové úlohy a stanovujú kritériá dosiahnutia jednotlivých úrovní výkonu.

2. Tvorba teoretického rámca

Musí byť jasne špecifikované, čo sa má testom zisťovať, pre koho je test určený a z akých úloh má byť zostavený. Teoretický rámec opisuje obsahové domény, testované kompetencie, kognitívne úrovne a formáty testových položiek. Zároveň zabezpečuje, aby všetky položky boli v súlade so vzdelávacími cieľmi definovanými v kurikulárnych dokumentoch. Teoretický rámec zároveň explicitne definuje očakávanú úroveň zvládnutia (*mastery*), ktorá bude neskôr použitá pri interpretácii výsledkov.

3. Vývoj testových položiek a testovacieho nástroja

Testové položky sú vytvárané tímom expertov (autorský tím) tak, aby priamo merali zvládnutie konkrétnych cieľov vzdelávania. Môžu mať rôzne formáty: výber z viacerých možností, priradovanie, zoraďovanie viet podľa kontextu, tvorba krátkych odpovedí, tvorivé úlohy na pochopenie celkového významu textu atď. Každá položka musí mať jasne definovaný cieľ, overovanú kompetenciu, kognitívnu úroveň a štruktúru správnej odpovede, pričom už v tejto fáze sa posudzuje jej vhodnosť pre psychometrické modelovanie. Neoddeliteľnou súčasťou procesu je systematicky budovaná databáza úloh, overovaných v rozsiahlych pilotných meraniach.

4. Realizácia pilotných testovaní a hlavného merania

Pilotné testovanie kritériálnych testov môže prebiehať vo viacerých postupných kolách (často 1 – 3), pričom ich počet a konkrétny dizajn závisí od cieľov projektu a rozsahu testovania. Každé z nich má odlišnú úlohu v procese vývoja testu a spoločne vytvárajú základ pre finálnu podobu testových zošitov používaných v hlavnom meraní. V rozsiahlych testovacích programoch môže pilotné overovanie prebiehať vo viacerých etapách.

V prvom pilotnom testovaní sa overuje kvalita samostatných testových položiek. Úlohy sú zaradené do viacerých testovacích zošitov (napr. päť zošitov po 24 položiek). Cieľom je získať základné informácie o ich psychometrických vlastnostiach – obťažnosti, diskriminačnej schopnosti alebo funkčnosti distraktorov. Výsledky prvého pilotu poskytujú autorom prvý orientačný obraz o tom, ako jednotlivé položky fungujú.

V druhom pilotnom testovaní sa okrem opätovného overovania parametrov položiek z prvého pilotu overujú aj klastre položiek – menšie tematicky alebo obsahovo previazané skupiny úloh. Testovanie môže prebiehať v menšom počte zošitov (napr. tri zošity po piatich položkách). Cieľom je overiť stabilitu parametrov položiek v čase v kontexte iných úloh, porovnať obťažnosť jednotlivých klastrov a identifikovať, či a ako niektoré položky v rôznych kombináciách menia svoje vlastnosti.

V treťom pilotnom testovaní sa overujú **kompletné testové zošity** (napr. tri zošity po 20 položiek), ktoré už pozostávajú výlučne z položiek a klastrov overených v predchádzajúcich pilotných meraniach, v ktorých sa preukázalo, že spĺňajú minimálne psychometrické kritériá. V tejto fáze sa sleduje najmä celková funkčnosť testových zošitov, rovnomernosť rozloženia obťažnosti položiek, psychometrické parametre ako validita a reliabilita, časová primeranosť. Na základe výsledkov tretieho pilotného testovania prebehne podrobné štatistické vyhodnotenie a externé recenzné posúdenie. Až následne sa definitívne rozhodne, ktoré položky, klastre a testové zošity budú zaradené do hlavného merania.

5. Dizajn plošných testovaní

Pri plošnom testovaní, resp. v hlavnom meraní rozsiahlych hodnotiacich štúdií, sa môže uplatňovať matrixový rotačný dizajn, ktorý sa využíva napríklad v štúdii PISA. V takomto dizajne žiaci v rámci jednej triedy nemusia riešiť rovnakú testovú formu. Jednotlivé testové formy, označované v papierovom testovaní aj ako testové zošity alebo booklet, sa zostavujú z viacerých blokov úloh. Tieto bloky sa podľa vopred stanoveného plánu kombinujú a opakujú v rôznych testových formách.

Každý žiak tak rieši iba určitú kombináciu blokov, pričom súbor všetkých použitých testových foriem umožňuje pokryť podstatne širší rozsah hodnoteného obsahu, než by mohol v stanovenom čase vyriešiť jeden žiak. Prekrytie úloh alebo blokov medzi testovými formami umožňuje ich psychometrické prepojenie. Pri použití vhodného modelu možno následne umiestniť parametre úloh a odhady výkonu žiakov na spoločnú škálu. Matrixový rotačný dizajn je preto vhodný predovšetkým na získavanie spoľahlivých výsledkov na úrovni hodnotenej populácie alebo jej skupín. Presnosť výsledku jednotlivého žiaka však môže byť vzhľadom na menší počet ním riešených úloh obmedzená.

Voľba testovacieho dizajnu musí vychádzať z účelu testovania a plánovanej úrovne vykazovania výsledkov. Matrixový rotačný dizajn je vhodný najmä na získavanie spoľahlivých výsledkov na úrovni populácie, školy alebo väčších skupín žiakov a môže poskytovať podklady na formatívne využitie,

napríklad na identifikovanie oblastí, v ktorých škola alebo trieda potrebuje posilniť výučbu. Ak má test poskytovať podrobnú spätnú väzbu jednotlivému žiakovi, musí obsahovať dostatočný počet vhodne zvolených úloh pre každú hodnotenú oblasť a zabezpečiť primeranú presnosť individuálnych výsledkov. Formatívny prínos teda nie je podmienený jedným konkrétnym dizajnom testu, ale primeranosťou poskytovaných informácií, ich zrozumiteľnou interpretáciou a následným využitím pri plánovaní ďalšieho učenia a vyučovania.

Matrixový rotačný dizajn nie je jedinou možnosťou pri plošnom kritériálnom testovaní. Niektoré národné testovania môžu pracovať s menším počtom testových foriem a s obmedzeným prekrytím úloh alebo blokov. Konkrétny dizajn sa volí podľa účelu testovania, šírky hodnoteného obsahu, požadovanej úrovne vykazovania výsledkov, požadovanej presnosti merania a organizačných možností realizácie. Aj v rámci jedného testovacieho programu sa môže líšiť podľa predmetu, ročníka alebo cyklu testovania.

Konečný dizajn hlavného merania vychádza z výsledkov pilotného testovania, v ktorom sa overujú vlastnosti úloh, fungovanie jednotlivých testových foriem, ich vzájomné prepojenie, primeranosť časového limitu a zvolený spôsob vyhodnocovania.

Napríklad súbor 60 úloh možno usporiadať do viacerých blokov, z ktorých sa podľa vopred stanoveného rotačného plánu zostavia jednotlivé testové formy. Niektoré úlohy alebo bloky sa medzi formami prekrývajú a plnia prepájaciu funkciu. Konkrétny počet blokov, testových foriem a prepájacích úloh musí vychádzať z účelu testovania, požadovanej presnosti výsledkov a empirických zistení pilotného merania.

6. Stanovovanie štandardov – hraničné skóre

Ak sa majú výsledky kritériálneho testovania interpretovať prostredníctvom výkonových úrovní, dôležitým krokom je stanovenie hraničných skóre (*cut scores*), ktoré vymedzujú hranice medzi susediacimi úrovňami výkonu. Môže ísť napríklad o základnú, strednú, pokročilú a vyššiu pokročilú úroveň. Každá hranica vyjadruje minimálnu úroveň výkonu potrebnú na zaradenie žiaka do príslušnej vyššej kategórie.

Stanovovanie hraničných skóre je súčasťou širšieho procesu stanovovania výkonových štandardov (*standard setting*). Jeho úlohou je prepojiť obsahové opisy jednotlivých úrovní výkonu s konkrétnymi hodnotami na skórovej škále. Hraničné skóre pritom nepredstavuje prirodzene danú ani čisto štatisticky odvodenú hodnotu. Je výsledkom formalizovaného odborného úsudku založeného na účele testovania, obsahu testu, opisoch výkonových úrovní a dostupných empirických údajoch⁸⁷.

Na stanovovanie hraničných skóre sa používajú rôzne metódy, ktoré sa líšia predmetom expertného posudzovania aj mierou využitia empirických údajov. Medzi známe postupy patria Angoffova metóda, metóda záložky (*Bookmark*), metóda posudzovania súborov žiackych výkonov (*Body of Work*) a metóda kontrastných skupín. Výber metódy závisí od typu testu, charakteru hodnoteného výkonu, dostupných údajov a spôsobu, akým sa majú výsledky interpretovať.

Pri Angoffovej metóde odborní posudzovatelia najskôr vytvoria spoločnú predstavu o žiakovi, ktorého výkon sa nachádza práve na hranici príslušnej úrovne. Následne pri každej položke odhadujú

⁸⁷ CIZEK, G. J. a BUNCH, M. B. *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Thousand Oaks, CA: Sage Publications, 2007; Štuka, Č. – Vejražka, M. 2021. *Testování a hodnocení studentů na VŠ*. Praha: Univerzita Karlova, Nakladatelství Karolinum, 2021, s. 64–70.

pravdepodobnosť, s akou by takto vymedzený hraničný žiak odpovedal správne. Agregovaním posudkov naprieč položkami a posudzovateľmi sa získa odporúčaná hodnota hraničného skóre. Ak sa stanovuje viac výkonových úrovní, analogický postup sa uskutočňuje pre každú hranicu osobitne.

Metóda záložky vychádza z empiricky kalibrovaných položiek zoradených podľa náročnosti. Posudzovatelia v takto usporiadanom súbore určujú miesto, od ktorého už položky spravidla presahujú očakávaný výkon žiaka na hranici danej úrovne. Metóda *Body of Work* sa opiera o posudzovanie ucelených ukážok reálneho výkonu žiakov, zatiaľ čo metóda kontrastných skupín porovnáva rozdelenia výsledkov skupín, ktoré boli na základe nezávislého externého posúdenia označené napríklad za kompetentné a ešte nekompetentné.

Empirické údaje nenahrádzajú odborný úsudok, ale pomáhajú posudzovateľom porozumieť vlastnostiam položiek a možným dôsledkom navrhovaných hraníc. V závislosti od zvolenej metódy možno využiť údaje o náročnosti a rozlišovacej schopnosti položiek, rozdelení výsledkov, predpokladanom podiele žiakov v jednotlivých úrovniach a neistote klasifikácie v okolí navrhovaných hraníc. Niektoré metódy, napríklad Bookmark alebo metóda kontrastných skupín, empirické údaje priamo vyžadujú; Angoffovu metódu možno uskutočniť aj pred administráciou testu, hoci empirické údaje môžu podporiť diskusiu posudzovateľov a posúdenie dôsledkov ich rozhodnutia.

Kvalita výsledného štandardu sa posudzuje najmä podľa obsahovej primeranosti opisov výkonových úrovní, výberu a prípravy posudzovateľov, konzistentnosti ich úsudkov, reprodukovateľnosti postupu a dôsledkov klasifikácie. Dôležité je preskúmať, ako navrhnuté hranice fungujú v rôznych relevantných skupinách a aké je riziko nesprávnej klasifikácie žiakov, najmä pri výsledkoch v bezprostrednej blízkosti hraničného skóre. Validita sa pritom nevzťahuje iba na test alebo stanovenú hranicu ako také, ale na konkrétne interpretácie výsledkov a rozhodnutia, ktoré sa na ich základe prijímajú. Čím závažnejšie dôsledky má takéto rozhodnutie pre žiaka, tým presvedčivejšie musia byť dôkazy podporujúce zamýšľanú interpretáciu a použitie výsledku⁸⁸.

Hraničné skóre teda predstavuje syntézu obsahového vymedzenia očakávaného výkonu, systematicky získaného expertného úsudku a relevantných empirických dôkazov. Samo osebe však neodstraňuje neistotu merania ani nevytvára ostrý rozdiel medzi dvoma inak veľmi podobnými výkonmi. Pri interpretácii výsledkov žiakov v blízkosti hranice je preto potrebné uvádzať neistotu klasifikácie a pri rozhodnutiach s významnými dôsledkami využívať aj ďalšie dôkazy o výkone žiaka.

7. Tvorba štatistických a psychometrických analýz

Psychometrické analýzy poskytujú empirické dôkazy o kvalite testových položiek, fungovaní testu a presnosti výsledkov. Podporujú posudzovanie toho, či výsledky testu možno spoľahlivo interpretovať zamýšľaným spôsobom a používať na stanovený účel. Samy osebe však nepreukazujú validitu testu. Validita sa vzťahuje na konkrétne interpretácie a spôsoby použitia výsledkov a opiera sa o širší súbor obsahových, teoretických a empirických dôkazov.

⁸⁸ KANE, M. T. Validating the Interpretations and Uses of Test Scores. Journal of Educational Measurement [online]. 2013, roč. 50, č. 1, s. 1 – 73. ISSN 0022-0655. Dostupné na: <https://doi.org/10.1111/jedm.12000>.

Pri kriteriálnom testovaní sú psychometrické analýzy dôležité aj pri stanovovaní výkonových štandardov a hraničných skóre. Poskytujú informácie o fungovaní jednotlivých položiek a testu ako celku, o presnosti odhadu výkonu a o neistote klasifikácie žiakov do výkonových úrovní. Hraničné skóre však nie je čisto štatisticky odvodenou hodnotou. Ako bolo spomenuté, je výsledkom formalizovaného procesu, ktorý prepája obsahové vymedzenie výkonových úrovní, odborný úsudok a relevantné empirické údaje.

Psychometrické analýzy zahŕňajú najmä:

- **Položkovú analýzu:**

- obťažnosť položky, ktorá sa v klasickej teórii testov (CTT) pri dichotomicky skórovaných položkách vyjadruje spravidla podielom žiakov, ktorí odpovedali správne;
- rozlišovaciu schopnosť položky, teda mieru, do akej položka rozlišuje medzi žiakmi s rozdielnou úrovňou sledovanej spôsobilosti;
- súlad položky s obsahovým vymedzením hodnoteného konštruktú a jej vzťah k ostatným položkám alebo k celkovému skóre;
- vplyv položky na presnosť a vnútornú konzistentnosť výsledkov testu;
- fungovanie distraktorov pri položkách s výberom odpovede.

- **Analýzu testu ako celku:**

- rozdelenie dosiahnutých skóre a výskyt efektu podlahy alebo stropu;
- presnosť a spoľahlivosť výsledkov;
- štandardnú chybu merania, prípadne podmienenú štandardnú chybu merania v rôznych častiach škály skóre;
- primeranosť časového limitu;
- obsahové pokrytie hodnotenej oblasti;
- pri klasifikačnom použití aj konzistentnosť a presnosť zaradenia žiakov do výkonových úrovní.

- **Modely teórie odpovede na položku (Item Response Theory, IRT):**

- odhad parametrov položiek a úrovne latentnej spôsobilosti žiakov;
- modelovanie pravdepodobnosti správnej odpovede v závislosti od spôsobilosti žiaka a vlastností položky;
- posudzovanie zhody položiek a dát so zvoleným modelom;
- vyjadrenie presnosti merania v rôznych častiach škály;
- podporu tvorby spoločnej škály, prepájania testových foriem a adaptívneho testovania.

- **Analýzu rozdielneho fungovania položiek (DIF):**

- zisťovanie, či majú žiaci z rôznych relevantných skupín pri rovnakej úrovni sledovanej spôsobilosti rozdielnu pravdepodobnosť správnej odpovede;
- následné obsahové posúdenie, či zistený rozdiel predstavuje relevantnú vlastnosť hodnoteného konštruktú alebo možné skreslenie položky.

- **Analýzu vnútornej štruktúry testu:**

- posudzovanie toho, či vzťahy medzi položkami zodpovedajú predpokladanej jednorozmernej alebo viacrozmernej štruktúre hodnoteného konštruktú;
- overovanie predpokladov potrebných na interpretáciu celkového skóre alebo čiastkových skóre.

- **Prepájanie a porovnávanie testových foriem:**

- tzv. *linking*, ktorého cieľom je umiestniť výsledky rôznych testov alebo foriem do vzájomného vzťahu

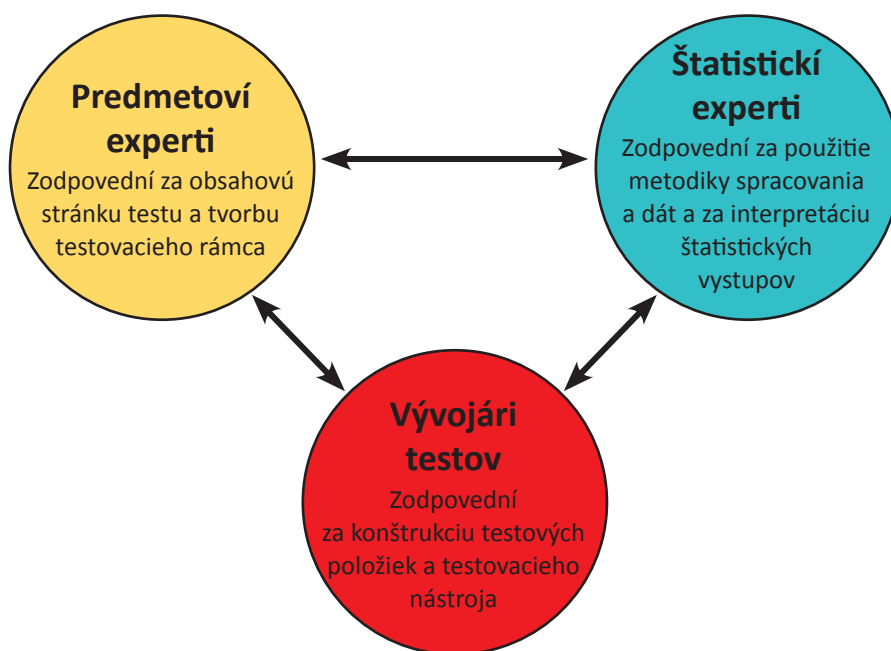
- alebo na spoločnú škálu;
- o tzv. *equating*, ktorý za prísnejších podmienok umožňuje vzájomne zameniteľnú interpretáciu skóre z rovnocenných foriem toho istého testu;
- o sledovanie stability škály a výkonových štandardov medzi cyklami testovania.

Pri CRT je potrebné venovať osobitnú pozornosť presnosti merania v okolí hraničných skóre. Dôležitá nie je iba celková reliabilita testu, ale aj pravdepodobnosť konzistentného a správneho zaradenia žiaka do príslušnej výkonovej úrovne. Posudzovať treba najmä neistotu klasifikácie a riziko falošne pozitívnych a falošne negatívnych rozhodnutí.

Výsledky psychometrických analýz tak tvoria významnú časť empirických dôkazov využívaných pri stanovovaní a hodnotení výkonových štandardov. Umožňujú posúdiť kvalitu položiek, presnosť výsledkov, fungovanie škály a dôsledky navrhovaných hraničných skóre. Konečné hranice medzi výkonovými úrovňami však musia byť zároveň obsahovo zmysluplné, metodologicky obhájiteľné a transparentne zdokumentované.

3.2 Tvorba CR-testov z hľadiska rolí expertov

Na tvorbe CRT sa podieľajú rôzni odborníci, ktorí sa špecializujú na určitú fázu vývoja testu. Pri tvorbe je od začiatku potrebné prihliadať na to, že výstupy z testov majú byť primárne určené pre učiteľov a ich využitie vo vzdelávacom procese na účely zlepšovania výsledkov vzdelávania žiakov.



Obr. 7 Diagram rolí expertov pri vývoji testov

Na vývoji testov sa podieľajú:

Predmetoví experti – zodpovední za obsahovú stránku testu a tvorbu testovacieho rámca. Nevyhnutné je, aby túto rolu zabezpečovali učitelia s aprobáciou na daný predmet, respektíve vzdelávaciu oblasť, ktorá bude testovaná.

Vývojári testov – zodpovední za konštrukciu testových položiek a testovacieho nástroja. Túto rolu by mali zastávať odborníci na testovanie, ktorí sa kontinuálne vzdelávajú v oblasti meraní výsledkov vzdelávania, a aj učitelia aprobovaní na testovaný predmet, respektíve vzdelávaciu oblasť.

Štatistickí experti – zodpovední za správne využitie metodiky spracovania dát a za interpretáciu štatistických výstupov. Túto rolu zabezpečujú odborníci na pedagogické meranie, najmä psychometrici, edukometrici, štatistici a dátoví analytici. Psychometrici a edukometrici musia byť zapojení už pri tvorbe položiek, aby sa zabezpečila kompatibilita úloh s plánovaným IRT modelom.

3.3 Vývoj kritériálnych testov na Slovensku

V tejto časti sa vysvetľujú kľúčové procesy vývoja kritériálnych testov a ich plánovaná implementácia na Slovensku. Text vychádza z odborných školení, ktoré pre slovenských testovacích špecialistov realizovali experti z Univerzity v Stuttgarte. Cieľom tejto časti je ukázať, ako budú princípy a postupy prezentované nemeckými odborníkmi adaptované na podmienky slovenského systému testovania a akým spôsobom budú tvoriť základ metodiky kritériálneho testovania v NIVaM. Uvádzané postupy budú zároveň doplnené konkrétnymi príkladmi a ukážkami výstupov vytvorených odborným tímom NIVaM v rámci projektu *Tvorba a vývoj kritériálnych testov na Slovensku* (Plán obnovy a odolnosti, Komponent 7).

Teoretický rámec

Základným východiskom pre vyvíjanie kritériálneho testu v štádiu jeho prípravy je vymedzenie teoretického rámca. Na jeho základe je možné komplexne a detailne opísať: obsah, ktorý bude testovaný, pre koho je test určený, čo sa ide testom zisťovať, z akých úloh bude zostavený a pod. Na základe teoretického rámca dokážu tvorcovia testu charakterizovať každú úlohu z určitých definovaných hľadísk, a to z obsahového zamerania, rozsahu požiadaviek na výkon žiaka a z hľadiska úrovni testovaných kompetencií. Vymedzenie teoretického rámca musí plne rešpektovať platné vzdelávacie štandardy štátneho vzdelávacieho programu (ŠVP).

V rámci vyvíjania kritériálnych testov na Slovensku odborníci NIVaM postupne vytvárajú teoretické rámce z jednotlivých testovaných predmetov, jedným z nich je **teoretický rámec z matematiky**⁸⁹. Pri matematike sa používa jeden spoločný rámec, ktorý nečlení matematiku na samostatné dokumenty pre aritmetiku, algebru či geometriu, ale opisuje kľúčové matematické kompetencie a kognitívne procesy naprieč celým predmetom. Všeobecný charakter rámca umožňuje posúdiť pokrytie a vyváženosť banky úloh z hľadiska matematických kompetencií v testoch. Vzhľadom na jeho obsiahlosť je potrebné ho vždy prispôbiť na konkrétne testovanie (počet úloh, čas testovania, dostupnosť vhodných úloh a pod.).

Teoretický rámec pre matematiku zároveň predstavuje operacionalizáciu meraného konštruktu, teda špecifickej podoby matematických vedomostí a kompetencií definovaných v kontexte štátneho kurikula (ŠVP). Pozostáva z viacerých častí, ktoré spoločne vymedzujú, čo sa má testovaním overovať a akým spôsobom majú byť úlohy konštruované. Prvou časťou je **rozsah požiadaviek**, ktorý opisuje charakteristiky úloh a ich kognitívnu náročnosť. Ide najmä o typy kognitívnych procesov, ktoré má žiak pri riešení úlohy aktivovať, a o úrovne ich požadovaného zvládnutia. Rozsah požiadaviek teda neurčuje len formu úloh, ale aj to, akú kognitívnu hĺbku majú dosahovať. Druhou časťou rámca sú **matematické kompetencie**. Patria sem oblasti ako argumentácia, komunikácia, práca so symbolickým jazykom a reprezentáciami, modelovanie a matematizácia či používanie matematických nástrojov a pomôcok. V rámci teoretického rámca nie sú tieto kompetencie uvedené len ako zoznam, ale sú aj stručne opísané v tom, čo presne predstavujú a ako sa v žiackych riešeniach prejavujú. Rámec zároveň definuje, aké úrovne osvojenia týchto kompetencií sa predpokladajú pri rôznych typoch úloh, aby bolo možné vytvárať položky primerané rôznym stupňom náročnosti.

⁸⁹ FICOVÁ, L. a PICHANIČOVÁ, I. Kritériálny test ako nástroj na mapovanie absolútneho výkonu žiaka. Pedagogické rozhľady [online]. 2022, č. 3, s. 20 – 26. Banská Bystrica: NIVAM. Dostupné na: <https://nivam.sk/wp-content/uploads/2025/02/pr-3-2022.pdf>.

Ďalšou súčasťou teoretického rámca je **obsahová stránka testovania**. Základným východiskom je štátny vzdelávací program (ŠVP) pre predmet matematika, v ktorom je učivo rozdelené do piatich tematických okruhov: Čísla, premenné a početné výkony; Vzťahy, funkcie, tabuľky a diagramy; Geometria a meranie; Kombinatorika, pravdepodobnosť a štatistika; Logika, zdôvodňovanie a dôkazy. Keďže ŠVP nevymedzuje všetky požiadavky v podrobnosti potrebnej na bezprostrednú tvorbu testových úloh, teoretický rámec tieto požiadavky na účely testovania ďalej odborne rozpracúva. Spresňuje najmä obsah, očakávané výkony a kognitívne činnosti, ktoré majú byť testovaním zachytené, pričom toto rozpracovanie musí zostať v súlade so ŠVP.

Na základe takto rozpracovaného rámca sa vytvára špecifikácia testu a následne konkrétne testové úlohy. Každá úloha sa priradzuje k vymedzenému vzdelávaciemu cieľu, obsahovej oblasti a požadovanej kognitívnej činnosti. Samostatná úloha pritom poskytuje iba čiastkový dôkaz o výkone žiaka; záver o miere dosiahnutia vzdelávacieho cieľa alebo zvládnutia určitej oblasti musí vychádzať zo súboru obsahovo reprezentatívnych úloh.

Takto prepojený rámec vytvára predpoklady na konzistentnú tvorbu úloh, primerané zastúpenie hodnotených oblastí a obsahovo zmysuplnú interpretáciu výsledkov. Primeranosť náročnosti, spoľahlivosť výsledkov a validitu ich interpretácie je však potrebné ďalej overovať expertným posúdením, pilotným testovaním a psychometrickými analýzami.

V rámci vývoja kriteriálneho testu z matematiky pre 1. cyklus ZŠ boli v NIVaM-e domény teoretického rámca rozpracované nasledovne:

Vybrané boli obsahové komponenty⁹⁰, ktoré sa obsahovo zameriavali na:

- Čísla a operácie s číslami,
- Závislosti, vzťahy a prácu s údajmi,
- Geometriu.

Obsahové komponenty musia byť vybrané tak, aby vytvorené položky mali dostatočný rozsah obťažnosti na pokrytie celého výkonového spektra populácie, čo je podmienkou pre stabilné IRT škálovanie.

Tabuľka 4 sumarizuje **obsah jednotlivých komponentov**, z ktorých boli vytvorené testové položky.

Tab. 4 Obsahové komponenty z matematiky

Čísla a operácie s číslami	Závislosti, vzťahy a práca s údajmi	Geometria
Obsahový komponent Čísla a operácie s číslami je zameraný na vytváranie a rozvíjanie predstavy žiakov o prirodzených číslach a ich význame a osvojovanie si skúsenosti s riešením úloh s prirodzenými číslami; chápanie jednoduchých aritmetických operácií s prirodzenými číslami, porozumenie vlastnostiam operácií a nadobúdanie skúsenosti s použitím operácií pri riešení matematických úloh.	Obsahový komponent Závislosti, vzťahy a práca s údajmi je zameraný na rozvíjanie schopnosti pozorovať, objaviť a opísať pravidlo jednoduchšej postupnosti a závislosti; vyhľadať a triediť údaje podľa rôznych kritérií; rozumieť rôznym záznamom údajov, vedieť doplniť jednoduché tabuľky, schémy a postupnosti; rozumieť jednoduchým kombinatorickým situáciám a vedieť ich riešiť, používať organizačný princíp pri hľadaní počtu možností, poznať jednoduché reprezentácie kombinatorických situácií.	Obsahový komponent Geometria je zameraný na rozvoj geometrického uvažovania prostredníctvom manipulačných činností a nadobúdanie skúseností s geometrickými pojmami v prirodzenom prostredí.

⁹⁰ MINISTERSTVO ŠKOLSTVA, VEDY, VÝSKUMU A ŠPORTU SR. Vzdelávacie štandardy: Vzdelávacia oblasť Matematika a informatika [online]. Dostupné na: https://www.minedu.sk/data/files/11820_matematika-a-informatika.pdf.

• Prirodzené čísla v obore do 10 000 • Sčítanie a odčítanie prirodzených čísel v obore do 1 000 • Násobenie a delenie prirodzených čísel v obore do 100 • Číselné výrazy v riešení úloh	• Základy práce so vzormi a postupnosťami • Základy práce s jednoduchými závislosťami a vzťahmi • Základy riešenia jednoduchých kombinatorických situácií • Pozorovanie jednoduchých pravdepodobnostných situácií • Základy práce s údajmi	• Orientácia v rovine a priestore • Základy skúmania geometrických útvarov v rovine • Základy skúmania geometrických útvarov v priestore • Jednoduché postupy merania a určovania miery • Základy skúmania a tvorby súmerných útvarov
--	--	---

V tabuľke 5 je uvedený **rozsah požiadaviek na výkon žiaka**, ktorý deklaruje charakteristickú vlastnosť každej úlohy s akcentom na jej kognitívnu náročnosť. Pre potreby gradácie úrovne jednotlivých úloh zaradených do testovacích nástrojov možno rozsah požiadaviek sledovať naprieč týmito úrovňami.

Tab. 5 Testovaný rozsah požiadaviek na výkon žiaka z matematiky

• základné porozumenie matematickým pojmom, • reprodukcia a priame použitie základných pojmov, tvrdení a postupov v jednej oblasti matematiky a kontexte, z ktorého je zrejmé, ktorý pojem, metódu, algoritmus má žiak použiť, • vykonávať rutinné postupy podľa priamych pokynov v explicitných situáciách – úloha obsahuje presné pomenovanie algoritmu (pojmu) nutného na jej riešenie, • schopnosť určiť, ktorý pojem, metódu, algoritmus má žiak použiť pri riešení úlohy – úloha neobsahuje presné pomenovanie matematického aparátu nutného na jej riešenie, • schopnosť prepojenia rôznych oblastí matematiky v rámci rôznych obsahových komponentov – vykonávanie výpočtov, procedúr, riešenie úloh, ktoré si vyžadujú prepojenie viacerých známych tematických celkov, • spojenie viacerých metód, postupov, výpočtových algoritmov, • úloha si vyžaduje vytvoriť plán činnosti, • úlohy sú charakterizované potrebou posúdenia, hodnotenia pravdivosti, • riešenie problémových situácií, • nadobudnuté vedomosti je žiak schopný použiť na riešenie úloh, kde z kontextu nie je automaticky zrejmé, ktoré matematické vedomosti (ktorý tematický celok) je nutné použiť, • úlohy s prebytočnou, resp. chýbajúcou informáciou, • porozumieť matematickým pojmom a používať ich v kontextoch, ktoré sa líšia od tých, v ktorých boli osvojené a precvičené, napr. nepriamo formulované úlohy.

Treťou časťou teoretického rámca sú testované matematické **kompetencie** obsahujúce kritériá, ktoré budú testovými položkami hodnotené (pozri tabuľku 6).

Tab. 6 Testované matematické kompetencie

Argumentácia	• posúdiť správnosť jednotlivých výpočtových procedúr a algoritmov, • na základe osvojených vedomostí, resp. jednoduchých kritérií rozhodnúť o príslušnosti prvku, resp. objektu k určitej množine, • uvádzať konkrétne príklady na základe platných tvrdení, • posúdiť správnosť postupov riešenia úloh, • sformulovať svoje myšlienky s použitím správnych matematických pojmov v logickej nadväznosti,
--------------	---

	• posúdiť relevantnosť výsledkov vzhľadom na kontext úlohy, • zdôvodniť pravdivosť alebo nepravdivosť tvrdenia, • sledovať, posudzovať a samostatne vytvárať reťazce matematických argumentov, • objektívne posúdiť možné riešenia problému na základe argumentov (vlastných aj cudzích), • na základe objaveného kritéria rozhodnúť o príslušnosti prvku, resp. objektu k určitej množine.
Komunikácia	• zapamätať si, vybaviť si a rozpoznať vedomosti a jednotlivé druhy poznatkov z dlhodobej pamäti, • porozumieť, vyjadriť ústne a písomne jednoduchý matematický obsah, nie je nutné použitie matematickej terminológie, • porozumieť, vyjadriť ústne a písomne zložitejší matematický obsah, pričom je nutná matematická terminológia, • aplikovať určitý postup, preukázať schopnosť zmysluplne použiť zovšeobecnenia v konkrétnych situáciách, • zrozumiteľne sprostredkovať a interpretovať úvahy, riešenia alebo výsledky, • porozumieť, ústne a písomne prezentovať zložitý matematický obsah, • urobiť rozbor matematického obsahu, ktorý sa týka vzájomných vzťahov a porozumieť písomným vyjadreniam o ňom, • rozlíšiť podstatné a nepodstatné údaje v matematickom obsahu, • posúdiť pravdivosť vyhlásení týkajúcich sa matematického obsahu, • vyjadriť správne zložitý matematický obsah, pričom je nutná matematická terminológia.
Symbolický jazyk a reprezentácia	• pracovať so základnými vzorcami v podobe, ako sú dané, • čítať s porozumením text obsahujúci matematické symboly a znaky, • rozpoznať a rozumieť štandardným reprezentáciám matematických objektov, • nájsť informácie z jednoduchých grafov, tabuliek, textov, obrázkov, • zapísať svoje myšlienky v logickej nadväznosti pomocou matematických symbolov a znakov, • pracovať s výrokmi a výrazmi obsahujúcimi symboly, • zvoliť si a použiť reprezentáciu matematických objektov, • prepojiť, nájsť súvislosti v údajoch z rôznych reprezentácií a modelov, • zaobchádzať so zložitými výrokmi a výrazmi alebo s neznámym symbolickým či formálnym jazykom, porozumieť tomuto jazyku a prekladať z tohto jazyka do prirodzeného jazyka, • prepojiť, nájsť súvislosti v údajoch zo zložitých reprezentácií a modelov, • porozumieť, tvorivo kombinovať a vytvárať neštandardné reprezentácie, • zvoliť si niektorú z rôznych foriem reprezentácie, prevádzať a rozlišovať medzi rôznymi formami reprezentácie, • vytvoriť jednoduché interpretácie údajov, systematické usporiadania údajov, jednoduché štatistické súbory, grafy, tabuľky, piktogramy.
Modelovanie a matematizácia	• rozpoznať, vybaviť si a využívať známe modely, • previesť známe modely (a ich výsledky) do reality a naopak a interpretovať výsledky známeho modelu, • rozpoznať, vybaviť si a využívať modely, ktoré nie sú pre žiakov bežné, • pracovať s modelmi komplexných konkrétnych situácií, • vybrať/vytvoriť vhodný matematický model, • overovať výsledný model, posudzovať model na základe jeho analýzy a kritického zhodnotenia, • spojiť použitie viacerých známych modelov, • previesť zložitejšie modely (a ich výsledky) do reality a naopak a interpretovať výsledky tohto modelu.

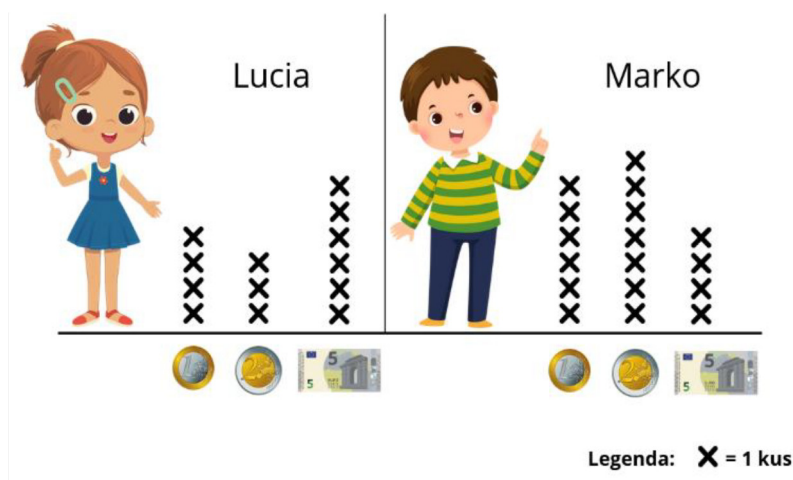
Nástroje a pomôcky	• poznať a používať bežné pomôcky a nástroje v takých kontextoch a situáciách a takými spôsobmi, ktoré sú blízke tým, v ktorých bolo používanie týchto pomôcok a nástrojov zavedené a precvičené, • poznať a používať bežné pomôcky a nástroje v takých kontextoch a situáciách a takými spôsobmi, ktoré sa líšia od tých, v ktorých bolo používanie týchto pomôcok a nástrojov zavedené a precvičené, • spojiť použitie viacerých pomôcok a nástrojov na vyriešenie daného problému, • používať rôzne menej známe pomôcky a nástroje na základe návodov, resp. inštrukcií, • uvedomiť si hranice možností pomôcok a nástrojov, • posúdiť vhodnosť použitia pomôcok a nástrojov na riešenie daného problému.
---------------------------	---

Každú položku z matematiky v kriteriálnom teste možno na základe týchto troch aspektov – obsahové zameranie, rozsah požiadaviek a úroveň kompetencií – vopred komplexne charakterizovať, čo je dôležité pre následný proces empirického vyhodnocovania výsledkov.

V nasledujúcej ukážke je uvedená vybraná testová položka CRT z matematiky pre 1. cyklus a jej detailná charakteristika (Obrázok 8 a Tabuľka 7).

Úloha

Lucia a Marko si do pokladničky sporili na bicykel. V pokladničke mali len 1-eurové, 2-eurové mince a 5-eurové bankovky. Počet mincí a bankoviek si zaznačili nasledovne.



Koľkokrát viac eur má v minciach Marko ako Lucia?

Správna odpoveď: 2

Obr. 8 Ukážka položky z matematiky pre 1. cyklus

Tab. 7 Charakteristika položky z matematiky v CRT

Obsahová štruktúra úlohy		Matematická gramotnosť
Cyklus		1.
Cieľ matematického vzdelávania		Používať prirodzené čísla, operácie s prirodzenými číslami a ich vlastnosti na riešenie jednoduchých aplikačných a kontextových úloh. Orientovať sa v grafe a používať ich pri riešení aplikačných úloh zameraných na vyhľadávanie, zaznamenávanie, triedenie, usporiadanie a interpretáciu údajov.
Komponent		Čísla a operácie s číslami Závislosti, vzťahy a práca s údajmi
Výkonový štandard		Vykonávať násobenie a delenie v obore do 100 aj mimo oboru násobilky využitím algoritmov, vlastností operácií a vzťahov medzi nimi. Žiak vie používať grafy na triedenie, usporiadanie a správnu interpretáciu údajov v reálnom živote a v aplikačných úlohách.
Obsahový štandard		Násobenie a delenie prirodzených čísel v obore do 100 Základy práce s údajmi
Matematické praktiky		znalosť a využívanie rôznych reprezentácií násobenia a delenia, používanie násobenia a delenia pri riešení matematických úloh; používanie vzťahov medzi operáciami pri zjednodušovaní a optimalizácii výpočtov, čítanie a zapisovanie násobenia a delenia; používanie symbolických znakov násobenia a delenia; identifikovanie a zdôvodnenie použitia násobenia a delenia v obore do 100 pri riešení jednoduchých úloh a interpretovanie výsledkov, používanie rôznych reprezentácií údajov (graf) a orientácia v nich, používanie grafu ako nástroja na riešenie aplikačných úloh s kvantitatívnymi aj kvalitatívnymi údajmi v rôznych oblastiach života, interpretovanie údajov v konkrétnych kontextoch a situáciách; diskutovanie o riešení jednoduchých aplikačných úloh s údajmi.
Typ úlohy		Úloha s tvorbou krátkej odpovede
Bloomova taxonómia		Analyzovať konceptuálne poznatky
Zaradenie podľa teoretického rámca testovania	Matematický obsah	Čísla a operácie s číslami Násobenie a delenie prirodzených čísel v obore do 100 Závislosti, vzťahy a práca s údajmi Základy práce s údajmi
	Rozsah požiadaviek	schopnosť prepojenia rôznych oblastí matematiky v rámci rôznych obsahových komponentov – vykonávanie výpočtov, procedúr, riešenie úloh, ktoré si vyžadujú prepojenie viacerých známych tematických celkov
	Kompetencie	sformulovať svoje myšlienky s použitím správnych matematických pojmov v logickej nadväznosti , porozumieť , vyjadriť písomne zložitejší matematický obsah, pričom je nutná matematická terminológia , rozpoznať a rozumieť štandardným reprezentáciám matematických objektov, nájsť informácie z grafov, tabuliek, textov, obrázkov, zapísať svoje myšlienky v logickej nadväznosti pomocou matematických symbolov a znakov, rozpoznať, vybaviť si a využívať modely , ktoré nie sú pre žiakov bežné .
Predpokladaný čas riešenia		4 minúty
Popis k úlohe		Aplikačná úloha je zameraná na násobenie prirodzených čísel v obore násobilky do 100 s nutnosťou identifikovania potrebných informácií z grafu. Žiak musí rozumieť základným matematickým pojmom, aby dokázal identifikovať, ktorú aritmetickú operáciu na výpočet danej úlohy potrebuje. Riešenie úlohy je viackrokové. Žiak musí z grafu identifikovať početnosť 1-eurových a 2-eurových mincí, čo využije na výpočet nasporenej sumy každého dieťaťa v minciach. Následne žiak určí, koľkokrát viac eur má v minciach jedno z detí. Žiak pracuje s textom súvislým aj s textom nesúvislým, pričom využíva základné matematické pojmy. V prípade nesúvislého textu pracuje žiak s grafom, ktorý nemá štandardné stvárnenie. Zo strany žiaka je potrebné danú situáciu analyzovať z pohľadu potrebných (počet mincí) a nepotrebných informácií (počet bankoviek) na vyriešenie danej aplikačnej úlohy.

Dizajn testovacieho nástroja

Dizajn testovacieho nástroja používaný na Slovensku vychádza, podobne ako VERA v Nemecku, z tvorby klastrov, teda menších súborov úloh s rovnakým počtom položiek, vyváženým obsahom a porovnateľnou úrovňou obťažnosti. Z týchto klastrov sa zostavujú testové zošity (booklety), ktoré predstavujú výslednú podobu testu pre jednotlivých žiakov. Pre každý teoretický rámec je vytvorený súbor klastrov a z nich odvodených zošitov, pričom všetky zošity sú vzájomne porovnateľné z hľadiska náročnosti aj obsahového pokrytia.

Pri plošnom testovaní NIVaM neuplatňuje plnohodnotný rotačný (matrixový) dizajn, ale model viacerých **testových zošitov (bookletov)**, ktorý vychádza z prístupov používaných napr. v nemeckom testovaní VERA. V tomto modeli existuje obmedzený počet testových zošitov, ktoré sú konštruované tak, aby mali porovnateľnú obťažnosť a obsahové pokrytie. Jednotlivé zošity obsahujú **čiasťovo spoločné (prekrývajúce sa) klastre úloh**, ktoré umožňujú ich vzájomné psychometrické prepojenie (*linking*). Každý žiak rieši jeden konkrétny **testový zošit**, pričom rôzne skupiny žiakov môžu dostať rôzne varianty zošitov. Na rozdiel od matrixového dizajnu však nejde o systematickú rotáciu veľkého množstva blokov medzi zošitmi, ale o menší počet vopred zostavených testových zošitov.

Tento prístup umožňuje zachovať primeranú dĺžku testu pre jednotlivého žiaka (aby nebol nadmerne zaťažovaný), zabezpečiť porovnateľnosť výsledkov medzi zošitmi a zároveň pracovať s mierne širšou bankou úloh. Psychometrické prepojenie (*linking*) medzi jednotlivými zošitmi je zabezpečené prostredníctvom **spoločných položiek alebo klastrov**, pričom na tento účel postačuje čiastočné prekrytie úloh medzi testovými formami.

Použitie viacerých zošitov vychádza zo skutočnosti, že jeden žiak dokáže zmysluplne vyriešiť iba obmedzený počet úloh, a preto nie je účelné, aby každý riešil celý rozsah úloh pokrývajúcich celý teoretický rámec. Vytvorením viacerých variantov testu je možné efektívne rozdeliť obsah bez nadmerného zaťaženia jednotlivca.

Napríklad pri súbore 48 úloh možno vytvoriť štyri testové zošity po 24 úloh. Každý zošit obsahuje štyri klastre po šesť úloh. Vybrané klastre sa medzi zošitmi čiastočne prekrývajú, takže približne polovica úloh (v tomto prípade 12 úloh – dva súbory po šesť úloh) je v jednotlivých zošitoch odlišná. Toto čiastočné prekrytie zabezpečuje, že všetkých 48 úloh možno následne analyzovať v rámci jednej spoločnej škály a porovnať ich náročnosť.

Psychometrické vyhodnocovanie testov

Psychometrické postupy využívané pri kritériálnom testovaní sú často rovnaké ako pri normatívnych testoch. Líšia sa však spôsobom interpretácie výsledkov, ktorý je založený na vzťahu k vopred stanoveným výkonovým štandardom.

Často pritom využívajú pokročilé štatistické metódy, najmä postupy založené na probablistických modeloch testovej teórie, známej ako **teória odpovede na položku (Item Response Theory, IRT)**. IRT predstavuje súbor modelov, ktoré opisujú pravdepodobnosť správnej odpovede na testovú položku ako funkciu dvoch faktorov: úrovne latentnej schopnosti testovaného a parametrov položky (typicky obťažnosť, diskriminačná schopnosť a v niektorých modeloch aj pravdepodobnosť uhádnutia)⁹¹.

⁹¹ HAMBLETON, R. K., SWAMINATHAN, H. a ROGERS, H. J. Fundamentals of Item Response Theory. Newbury Park, California: SAGE Publications, 1991. 174 s. Measurement Methods for the Social Sciences, zv. 2. ISBN 978-0-8039-3647-8.

V najjednoduchšej podobe (Raschov model) IRT opisuje pravdepodobnosť správnej odpovede ako funkciu dvoch parametrov: úrovne latentnej spôsobilosti žiaka a obťažnosti položky. Rozšírené modely môžu zohľadňovať aj diskriminačnú schopnosť položky a pravdepodobnosť uhádnutia.

Modely IRT patria v súčasnosti medzi najpoužívanéjšie prístupy vo veľkoplošných štandardizovaných meraniach. Od 80. rokov 20. storočia sa stali štandardom v mnohých medzinárodných aj národných testovacích programoch, a to vďaka schopnosti spoľahlivo škálovať položky aj žiakov, umožniť prepojenie testových foriem a poskytovať stabilnejšie odhady schopnosti⁹².

Teória odpovede na položku (IRT) vychádza z koncepcie latentných premenných. Predpokladá, že pozorované odpovede na testové položky sú ovplyvnené nepozorovateľnou vlastnosťou jednotlivca, označovanou ako latentná spôsobilosť. V pedagogickom testovaní môže ísť napríklad o matematickú gramotnosť alebo inú vymedzenú vedomostnú či kompetenčnú oblasť. Správne a nesprávne odpovede žiaka predstavujú pozorované indikátory, na základe ktorých sa úroveň tejto spôsobilosti odhaduje.

IRT modely vyjadrujú pravdepodobnosť konkrétnej odpovede ako funkciu úrovne latentnej spôsobilosti žiaka a vlastností položky. Pri pevne danej položke spravidla platí, že **s rastúcou úrovňou spôsobilosti žiaka rastie pravdepodobnosť správnej odpovede.** Naopak, **pri rovnakej úrovni spôsobilosti je pravdepodobnosť správneho vyriešenia náročnejšej položky nižšia než pri menej náročnej položke.**

V procese odhadu sa táto logika využíva opačným smerom: vzorec odpovedí žiaka, najmä úspešné riešenie náročnejších položiek, podporuje odhad vyššej úrovne latentnej spôsobilosti. Tento záver sa však nevyvodzuje z jednej položky, ale zo súboru odpovedí a z ich zhody s použitým IRT modelom.

Zatiaľ čo klasická testová teória (CTT) hodnotí žiaka na základe počtu správne vyriešených položiek, kde vo väčšine prípadov za každú položku v teste žiak dostane jeden bod, IRT využíva informácie o vlastnostiach položiek, najmä o ich obťažnosti a v niektorých modeloch aj o diskriminačnej schopnosti, na **odhad latentnej úrovne žiaka.**

V jednom modeli sa kalibrujú obťažnosti položiek a odhadujú úroveň latentnej premennej žiakov na tej istej škále. Keď následne usporiadame položky na tejto škále od najľahšej po najobťažnejšiu, IRT model dokáže presnejšie (v porovnaní s CTT) modelovať úroveň žiaka podľa obťažnosti položiek, na ktoré žiak ešte odpovedá správne, a na ktoré už odpovedá nesprávne⁹³.

Okrem samotnej úrovne schopnosti žiaka vplyvajú na správnosť alebo nesprávnosť riešenia úlohy aj ďalšie faktory. Dôležitú úlohu zohrávajú **vlastnosti položky**, ako už spomínaná obťažnosť či rozlišovacia schopnosť, ktoré interagujú s úrovňou žiakovej schopnosti. Do výkonu vstupujú aj **náhodné situačné vplyvy** (napr. pozornosť, únava), ktoré sa v dátach prejavujú ako zdroj náhodnej variability a prispievajú k tomu, že ani pri rovnakej úrovni schopnosti nie je výsledok odpovede úplne predvídateľný.

⁹² CARLSON, J. E. a VON DAVIER, M. Item Response Theory. In: BENNETT, R. a VON DAVIER, M. (eds.). Advancing Human Assessment. Methodology of Educational Measurement and Assessment [online]. Cham: Springer, 2017, s. 127 – 174. Dostupné na: https://doi.org/10.1007/978-3-319-58689-2_5.

⁹³ ALLEN, M. J. a YEN, W. M. Introduction to Measurement Theory. Long Grove, IL: Waveland Press, 2002, s. 253–269. ISBN 978-1-57766-230-3.

Teória odpovede na položku tieto skutočnosti zohľadňuje tým, že modeluje pravdepodobnosť úspechu v úlohe ako funkciu parametrov položky a úrovne latentnej schopnosti. Ako uvádza de Ayala, IRT umožňuje na základe určitých východiskových predpokladov modelovať pravdepodobnosť výberu jednotlivých možností odpovede na danú položku ako funkciu vlastností tejto položky a miery latentnej premennej, ktorú sa snažíme merať (de Ayala, 2009)⁹⁴. Tento prístup poskytuje presný a robustný spôsob, ako popísať vzťah medzi vlastnosťami úloh, schopnosťami žiakov a náhodnými vplyvmi, ktoré môžu ovplyvniť výsledný výkon.

Je však potrebné zdôrazniť, že využitie IRT modelov samo o sebe neurčuje kritériálny charakter testovania. **IRT predstavuje nástroj na škálovanie a odhad schopností, zatiaľ čo kritériálna interpretácia výsledkov je zabezpečená až prostredníctvom stanovenia výkonových úrovní, ich prepojenia so vzdelávacími štandardmi a interpretáciou údajov.**

Pri kritériálnom testovaní, ktorého výsledky sa interpretujú prostredníctvom výkonových úrovní, zohráva kľúčovú úlohu **stanovenie hraničných skóre**, ktoré oddeľujú jednotlivé úrovne znalostí či spôsobilostí žiakov. Pri interpretácii výsledkov kritériálneho testovania predstavuje určenie týchto hraníc **jeden z kľúčových krokov, pretože umožňuje prepojiť výsledky merania s očakávanými vzdelávacími štandardmi.**

Každá definovaná úroveň musí obsahovať jasný opis kompetencií a vedomostí, ktoré žiaci na danej úrovni typicky preukazujú. Na interpretácii škály latentnej spôsobilosti a na vymedzení výkonových úrovní spolupracujú predmetoví experti a psychometricki. Psychometrické modely poskytujú empirický základ škály, zatiaľ čo predmetoví experti zabezpečujú jej obsahovú interpretáciu vo vzťahu k vzdelávacím cieľom.

Výsledný model s charakteristikou jednotlivých úrovní vzniká spoločnou, efektívne koordinovanou prácou predmetových expertov a psychometrikov. V mnohých postupoch stanovovania štandardov, najmä pri metóde Bookmark, odborníci pracujú s položkami zoradenými podľa empirickej obťažnosti odhadnutej pomocou IRT modelov.

Tak ako uvádzajú Cizek a Bunch, stanovenie štandardov je systematický a metodologicky riadený proces, v ktorom odborníci kombinujú profesijný úsudok s empirickými údajmi, aby dosiahli **konsenzus pri určovaní hraničných skóre**⁹⁵.

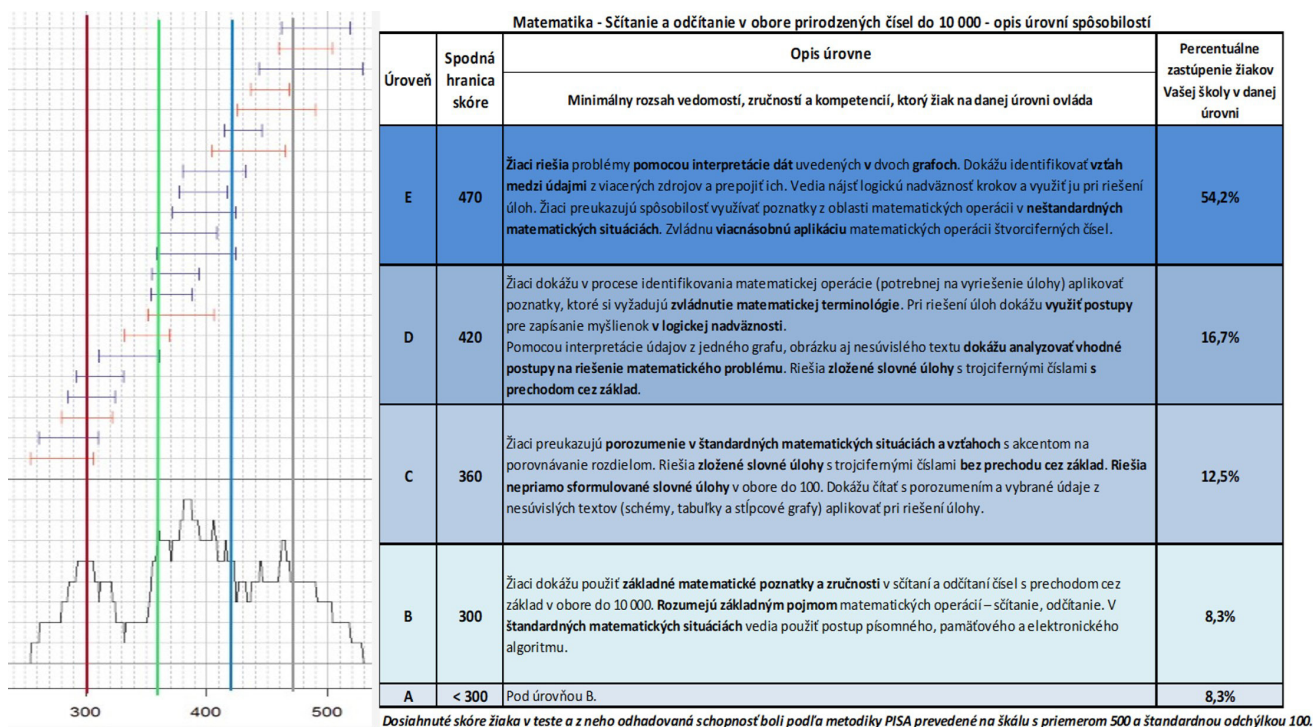
Každá výkonová úroveň je následne podrobne opísaná na základe množiny úloh charakteristických pre danú časť škály a na základe odborného posúdenia vedomostí, zručností a kompetencií, ktoré tieto úlohy reprezentujú.

Opisy úrovní by mali byť formulované tak, aby sa dali používať pri formatívnom hodnotení a správach pre školy obsahujúcich spätnú väzbu. Takto spracované výsledky umožnia školám získať kvalitnú spätnú väzbu o výkone žiakov v súvislosti s testovaným obsahom a poskytujú interpretácie, ktoré sú v súlade s koncepciou kritériálneho testovania.

⁹⁴ DE AYALA, R. J. The Theory and Practice of Item Response Theory. New York: The Guilford Press, 2009. ISBN 978-1-59385-869-8.

⁹⁵ CIZEK, G. J. a BUNCH, M. B. Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests. Thousand Oaks, CA: Sage Publications, 2007. ISBN 978-1-4129-1683-7. [\[archive.org\]](https://archive.org)

V nasledujúcej ukážke sú prezentované príklady výstupov kritériálne orientovaného testovania z matematiky založeného na využití modelov teórie odpovede na položku (IRT) a na procese stanovovania výkonových štandardov. Obrázok 9 ilustruje prepojenie IRT škály, hraničných skóre a obsahových opisov výkonových úrovní, ktoré umožňujú interpretovať výsledky žiakov z hľadiska dosiahnutej úrovne spôsobilosti. Pri každej výkonovej úrovni je zároveň uvedené percentuálne zastúpenie žiakov zaradených do príslušnej úrovne.



Obr. 9 Ukážka výkonových úrovní, hraničných skóre a opisov dosiahnutej spôsobilosti žiakov

4. Spätná väzba a využitie výsledkov kritériálnych testov

Táto kapitola sa zameriava na možnosti interpretácie a využitia výsledkov kritériálneho testovania ako nástroja spätnej väzby a formatívneho hodnotenia. Opisuje, akým spôsobom môžu byť výsledky CRT spracované do zrozumiteľných výstupov pre školy a učiteľov a ako podporujú identifikáciu úrovní zvládnutia učiva, silných stránok aj oblastí vyžadujúcich cielenejšiu pedagogickú podporu. Ťažisko kapitoly je kladené na využitie CRT v každodennej pedagogickej praxi – pri plánovaní výučby, poskytovaní spätnej väzby, podpore učenia a motivácie žiakov – a zároveň upozorňuje na metodologické limity a riziká nesprávnej interpretácie výsledkov. Kritériálne testovanie je prezentované ako významný, no nie jediný zdroj dôkazov o učení, ktorý má slúžiť predovšetkým na podporu rozvoja žiakov a zvyšovanie kvality vyučovania.

Box 4: Kedy má kritériálne testovanie formatívny účinok?

Kritériálny test nie je automaticky formatívnym hodnotením. Kritériálna interpretácia výsledku poskytuje informáciu o tom, do akej miery výkon žiaka zodpovedá vymedzeným vzdelávacím cieľom alebo výkonovým štandardom. Formatívny účinok vzniká až vtedy, keď sa získaná informácia využije na podporu ďalšieho učenia alebo na úpravu vyučovania.

Aby mohli výsledky testovania plniť formatívnu funkciu, mali by:

- poskytovať viac než iba celkové skóre alebo označenie výkonovej úrovne;
- byť obsahovo prepojené s konkrétnymi vzdelávacími cieľmi;
- poskytovať primerane spoľahlivé informácie o tom, ktoré oblasti žiak pravdepodobne zvláda, a v ktorých sa prejavujú ťažkosti;
- byť sprostredkované v zrozumiteľnej a pedagogicky využiteľnej podobe;
- viesť k určeniu primeraného ďalšieho kroku vo vyučovaní alebo učení;
- umožniť následne overiť, či poskytnutá podpora viedla k pokroku.

Učiteľ môže výsledky využiť napríklad na opätovné vysvetlenie nedostatočne zvládnutého učiva, výber vhodnejších úloh, diferenciaciu podpory alebo úpravu tempa vyučovania. Žiak môže získať informáciu o tom, čo už pravdepodobne zvládol, na čom má ďalej pracovať, a aký postup mu môže pomôcť.

Výsledok testu sám osebe nevysvetľuje príčinu výkonu žiaka. Nízke skóre môže súvisieť s nedostatočným porozumením učivu, ale aj s nepozornosťou, motiváciou, nedostatočnou znalosťou formátu úloh, jazykovou náročnosťou testu alebo ďalšími okolnosťami. Výsledky externého testovania je preto potrebné prepájať s pozorovaním učiteľa, žiackymi prácami, rozhovorom so žiakom a ďalšími dôkazmi o jeho učení sa.

Formatívne využitie výsledkov možno vyjadriť ako cyklus: **vzdelávací cieľ → získanie dôkazu o aktuálnom výkone → interpretácia výsledku → spätná väzba a podpora → ďalšie učenie → opätovné overenie pokroku.**

Testovanie má formatívny účinok iba vtedy, ak získané informácie vedú k ďalšiemu pedagogickému alebo učebnému kroku.

Vyhodnotenie kritériálnych testov sa realizuje s cieľom poskytnúť školám čo najadresnejšiu spätnú väzbu o miere zvládnutia testovaného učiva. Miera podrobnosti a presnosti tejto spätnej väzby závisí od spôsobu konštrukcie testu, použitého modelu merania a spôsobu interpretácie výsledkov. V testoch využívajúcich modely teórie odpovede na položku (IRT) a výkonové úrovne možno výsledky interpretovať vo vzťahu k opisom vedomostí, zručností a kompetencií charakteristických pre jednotlivé úrovne výkonu. Na základe takto definovaných úrovní je možné identifikovať žiakov, ktorí dosiahli stanovené výkonové kritériá, ako aj oblasti, v ktorých tieto kritériá zatiaľ nedosahujú. Pre školy môže byť zároveň užitočné sledovať agregované výsledky žiakov podľa jednotlivých výkonových úrovní.

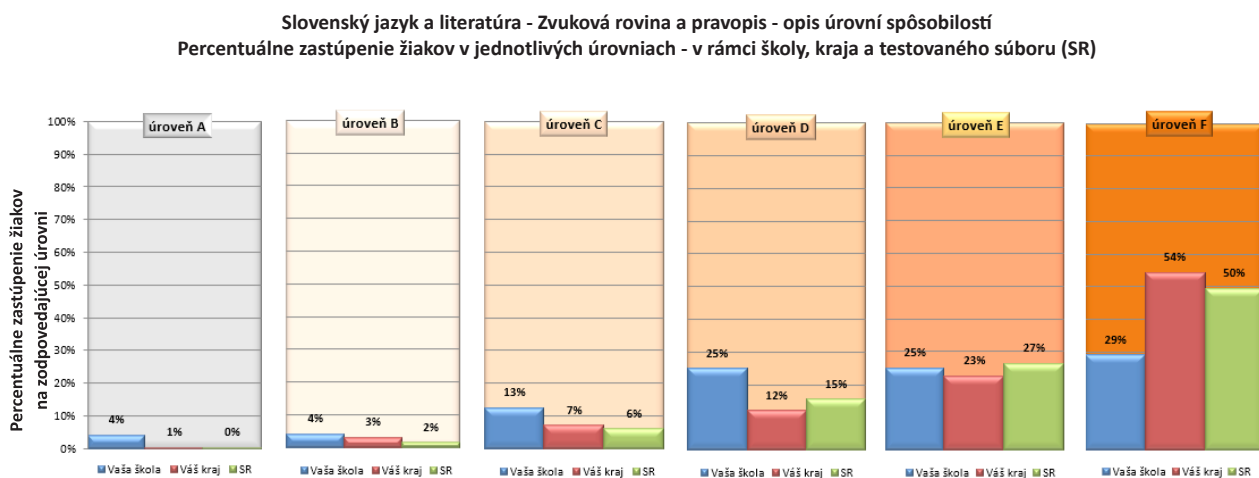
V ďalšej ukážke zo slovenského jazyka a literatúry, na obrázku 10 je znázornených šesť výkonových úrovní (úrovní spôsobilosti) vymedzených hraničnými skóre od úrovne A (menej ako 280 bodov), cez úroveň B (280 – 329 bodov) až po úroveň F (490 a viac bodov). Body v tomto prípade nepredstavujú hrubé skóre získané v teste, ale transformované hodnoty odhadov latentnej spôsobilosti žiakov získaných pomocou modelov teórie odpovede na položku (IRT). Tieto odhady boli následne prevedené na škálu s priemerom 500 a štandardnou odchýlkou 100, podobne ako v medzinárodných meraniach PISA. Najvyššiu výkonovú úroveň F dosahuje 29,2 % testovaných žiakov danej školy. Úrovne E a D sú zastúpené po 25,0 % žiakov. Úroveň C dosahuje 12,5 % žiakov, zatiaľ čo najnižšiu opísanú výkonovú úroveň B dosahuje 4,2 % žiakov. Rovnaký podiel žiakov (4,2 %) nedosiahol ani úroveň B a bol zaradený do úrovne A. Túto skupinu možno považovať za skupinu žiakov so zvýšeným rizikom nedostatočného zvládnutia sledovaných vzdelávacích cieľov, ktorí môžu vyžadovať zvýšenú pedagogickú pozornosť a cielenú podporu.

Na základe spracovania výsledkov kritériálneho testovania pomocou modelov IRT a výkonových úrovní stanovených procesom určovania štandardov je možné každej škole poskytnúť spätnú väzbu o výsledkoch jej žiakov z hľadiska na testovaný obsah vo forme výsledkového listu. Pre lepšiu predstavu je priložená ukážka spätnej väzby pre školu, ktorá vznikla v rámci experimentálneho overovania metodiky kritériálneho testovania.

Výsledky kritériálneho testu v uvedenom prehľadnom formáte predstavujú výstup pilotného testovania realizovaného NIVaM. Učiteľom poskytujú informáciu o tom, do akej miery žiaci dosiahli očakávané vzdelávacie ciele a výkonové štandardy v hodnotenej oblasti. Zároveň ponúkajú prehľad o zastúpení žiakov v jednotlivých výkonových úrovniach v rámci daného ročníka. Učiteľ dostáva grafický prehľad rozdelenia žiakov podľa úrovne dosiahnutých vedomostí, zručností a kompetencií. Takýto diagnostický výstup môže slúžiť ako podklad pre pedagogické rozhodovanie, plánovanie vyučovania a cielenú podporu žiakov. Umožňuje identifikovať skupiny žiakov, ktoré potrebujú dodatočnú podporu, a prispôbiť vyučovacie stratégie, formy a metódy tak, aby sa zvýšila pravdepodobnosť dosiahnutia požadovaných vzdelávacích cieľov a rozvoja požadovaných vedomostí, zručností a kompetencií.

Slovenský jazyk a literatúra - Zvuková rovina a pravopis - opis úrovni spôsobilostí				
Úroveň	Spodná hranica skóre	Opis úrovne		Percentuálne zastúpenie žiakov Vašej školy v danej úrovni
		Minimálny rozsah vedomostí a zručností, ktorú žiak na danej úrovni ovláda	Slovná zásoba	
F	490	Žiak pri riešení úlohy využíva analýzu obsahu textu. Žiak rieši úlohy zamerané na špecifické jazykové prvky (rôzne výnimky, napr. slabikotvorné hlásky). Žiak pri riešení úloh uplatňuje kognitívne procesy na úrovni konceptuálnych poznatkov. Žiak dokáže samostatne vyvodiť záver prostredníctvom analýzy.	slová z periferie, s ktorými sa žiak bežne nestretáva	29,2 %
E	430	Žiak rieši úlohy, v ktorých sa porozumenie kontextu stáva dôležitým pre riešenie úlohy. Žiak sa dokáže orientovať v textových jednotkách. Žiak využíva analýzu s ohľadom na kombinované prvky (napr. lexikálne, fonologické a ortografické). Identifikácia správneho riešenia si od žiaka vyžaduje zapojenie komplexného arzenálu vedomostí. Žiak uvažuje v kontexte a prepája vedomosti z rôznych rovin.	náročnejšie vyjadrovacie prostriedky a ich významové odtiene z rôznych vrstiev slovnej zásoby	25,0 %
D	380	Žiak dokáže riešiť úlohy, v ktorých sa kontext stáva sekundárnym sprievodným prvkom dôležitým pre správne riešenie úlohy. Žiak dokáže správne identifikovať význam slova. Žiak uplatňuje analýzu na úrovni slova. Žiak prepája vedomosti z rôznych rovin (napr. ortografia, fonetika, lexikológia) a očakáva sa od neho interpretácia na základnej úrovni.	slovná zásoba rozšírená o štylisticky príznakové slová	25,0 %
C	330	Žiak zvláda riešiť úlohy bez viazanosti na kontext na úrovni porozumenia alebo jednoduchších aplikačných úkonov. Žiak dokáže porozumieť dlhšej inštrukcii v zadaní úlohy. Žiak dokáže riešiť úlohy s distraktormi pozostávajúcimi z viacerých prvkov. Žiak triedi javy do definovaných kategórií podľa zadaných kritérií.	slová aj z okrajov slovnej zásoby, bežne používané v školskom prostredí	12,5 %
B	280	Žiak ovláda jednoúrovňové vedomosti – vyžadujúce od žiaka zapamätanie alebo porozumenie. Žiak má základné porozumenie kľúčovej informácie. Žiak vie riešiť izolované úlohy – bez súvislosti s kontextom. Žiak dokáže riešiť úlohy s distraktormi pozostávajúcimi z viacerých prvkov. Žiak sa riadi základnou inštrukciou uvedenou v otázke bez potreby vyhľadávania v texte.	slová z jadra slovnej zásoby (základná úroveň)	4,2 %
A	< 280	Úroveň A nie je možné prostredníctvom nastavenej metodiky opísať.		4,2 %
Dosiadnuté skóre žiaka v teste a z neho odhadovaná schopnosť boli podľa metodiky PISA prevedené na škálu s priemerom 500 a štandardnou odchýlkou 100				

Obr. 10 Výsledkový list – ukážka opisu úrovni spôsobilostí žiakov zo slovenského jazyka a literatúry



Na základe vybraného súboru úloh v teste je možné opísať len úrovne B až F.

Obr. 11 Ukážka výsledkového listu pre školu

4.1 Poskytovanie spätnej väzby na základe výsledkov kritériálnych testov

Kritériálne testovanie môže poskytovať učiteľom a žiakom zmysluplnú, diagnosticky orientovanú spätnú väzbu, ktorá podporuje učenie a rozvoj vedomostí, zručností a kompetencií žiakov. Ak sú výsledky interpretované prostredníctvom jasne definovaných výkonových úrovní a vzdelávacích cieľov, môžu sa stať významným zdrojom informácií pre formatívne hodnotenie. Spätná väzba by mala vychádzať z vopred stanovených vzdelávacích cieľov a kritérií úspešnosti a mala by žiakom aj učiteľom poskytovať informácie

o ďalších krokoch v učení (Black a Wiliam, 1998; Brookhart, 2008)⁹⁶.

Vzdelávacie ciele a kritériá:

- spätná väzba by mala byť konkrétna a mala by vychádzať z jasne definovaných, merateľných a operacionalizovaných kritérií, ktoré sú žiakom vopred komunikované;
- umožňujú žiakom pochopiť mieru dosiahnutia očakávaných výsledkov učenia a ich pozíciu v rámci postupnosti rozvoja učenia (*learning progression*);
- informácie o spätnej väzbe sú viazané na konkrétne výkonové úrovne, nie na sumárne skóre alebo počet bodov;
- výkonové úrovne by mali reflektovať postupné štádiá osvojovania si kompetencií.

Rozvoj kompetencií, nie porovnávanie:

- na rozdiel od normatívneho testovania CRT neporovnáva žiakov medzi sebou, ale posudzuje mieru dosiahnutia stanovených kritérií;
- výsledky CRT sa primárne interpretujú vo vzťahu k vopred vymedzeným vzdelávacím cieľom, kritériám alebo výkonovým štandardom, nie vo vzťahu k výsledkom ostatných žiakov. Percentilové poradie je naproti tomu normatívnym ukazovateľom, pretože vyjadruje relatívne postavenie žiaka v porovnávanej skupine (Nitko a Brookhart, 2014, s. 25 – 30).⁹⁷

Podpora učenia sa a sebareflexie

- spätná väzba má diagnostický charakter – identifikuje silné stránky a oblasti na zlepšenie;
- podporuje rozvoj sebahodnotenia a sebaregulácie učenia sa;
- obsahuje konkrétne odporúčania pre ďalšie kroky učenia sa, ktoré sú pre žiaka zrozumiteľné a realizovateľné;
- mala by byť poskytovaná v primeranom časovom horizonte, aby podporila efektívne učenie sa.

Učiteľ ako sprievodca učením

- učiteľ analyzuje výsledky kritériálneho testu a prispôsobuje výučbu potrebám žiakov (individualizácia a diferenciacia);
- využíva výsledky ako podklad pre pedagogické intervencie a plánovanie ďalších krokov;
- pracuje s profilom výkonových úrovní jednotlivých žiakov, nielen s agregovanými výsledkami triedy.

Transparentnosť a motivácia

- kritériá hodnotenia sú zrozumiteľné a dostupné, čo zvyšuje transparentnosť procesu učenia;
- spätná väzba je formulovaná konštruktívne a podporuje vnútornú motiváciu žiakov;
- formulácie spätnej väzby sú jasné, povzbudzujúce a zamerané na výkon, nie na osobu žiaka;
- pozitívne hodnotenie aj čiastkového pokroku podporuje rozvoj sebaúčinnosti (*self-efficacy*; Bandura, 1997)⁹⁸.

⁹⁶ BLACK, P. a WILIAM, D. Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice. 1998, roč. 5, č. 1, s. 7 – 74. ISSN 0969-594X.

BLACK, P. a WILIAM, D. Developing the Theory of Formative Assessment. Educational Assessment, Evaluation and Accountability [online]. 2009, roč. 21, č. 1, s. 5 – 31. ISSN 1874-8597. Dostupné na: <https://doi.org/10.1007/s11092-008-9068-5>.

BROOKHART, S. M. How to Give Effective Feedback to Your Students. Alexandria, VA: ASCD, 2008, s. 1–9. ISBN 978-1-4166-0736-6.

⁹⁷ NITKO, A. J. a BROOKHART, S. M. Educational Assessment of Students. 7. vyd. Boston: Pearson, 2014, s. 25–30. ISBN 978-0-13-343649-5.

⁹⁸ BANDURA, A. Self-efficacy: The Exercise of Control. New York: W. H. Freeman, 1997. ISBN 978-0716728504.

4.2 Využitie výsledkov kritériálnych testov na formatívne hodnotenie

Výsledky kritériálnych testov predstavujú významný zdroj informácií pre formatívne hodnotenie, ktorého cieľom je zlepšovať proces učenia a vyučovania. Formatívny potenciál kritériálneho testovania však nevyplýva automaticky z charakteru testu, ale zo spôsobu interpretácie výsledkov a ich následného využitia pri podpore učenia sa.

CRT je vhodné chápať ako jeden zo zdrojov dôkazov o učení (*evidence of learning*), nie ako jediný nástroj hodnotenia (OECD, 2013)⁹⁹, (Pellegrino, 2020)¹⁰⁰. Využitie výsledkov kritériálneho testu možno opísať prostredníctvom cyklu formatívneho hodnotenia: zisťovanie úrovne učenia, interpretácia výsledkov a pedagogická reakcia (Black a Wiliam, 2009)¹⁰¹.

Učitelia by mali výsledky kritériálnych testov využívať predovšetkým ako nástroj na zlepšenie kvality výučby, podporu učenia sa žiakov či kvalitnejšiu pedagogickú diagnostiku¹⁰². Učitelia môžu využiť kritériálne testovanie na formatívne hodnotenie žiakov prostredníctvom nasledovných procesov:

Diagnostika a plánovanie výučby

Výsledky CRT umožňujú učiteľom identifikovať úroveň zvládnutia učiva vo vzťahu k vzdelávacím cieľom a štandardom. Na základe týchto informácií môžu učitelia cielene prispôbiť obsah, metódy a tempo výučby potrebám jednotlivcov aj celej triedy.

Poskytovanie cielenej spätnej väzby

Výsledky CRT umožňujú poskytovať konkrétnu spätnú väzbu viazanú na výkonové úrovne. Učiteľ môže žiakom pomôcť identifikovať, čo zvládli, čo je potrebné zlepšiť, a aké kroky majú nasledovať. Súčasťou tohto procesu je podpora sebahodnotenia a aktívne zapojenie žiaka do riadenia vlastného učenia sa. Spätná väzba musí byť explicitne prepojená s výkonovými úrovňami, ktoré žiaci dosahujú.

Reflexia efektivity výučby

Učitelia môžu výsledky CRT využívať na reflexiu vlastnej pedagogickej praxe, identifikáciu účinných stratégií a oblastí vyžadujúcich zmenu. Výsledky slúžia ako podklad pre úpravu výučby a podporu učenia žiakov, nie na hodnotenie výkonu učiteľa.

Komunikácia s rodičmi a školou

Výsledky CRT môžu slúžiť ako podklad pre komunikáciu s rodičmi o tom, ktoré vzdelávacie ciele už žiak dosiahol, v ktorých oblastiach ešte potrebuje podporu, a aké ďalšie kroky môžu škola a rodina spoločne podporiť.

⁹⁹ OECD. Synergies for Better Learning: An International Perspective on Evaluation and Assessment. Paris: OECD Publishing, 2013. ISBN 978-9264190658.

¹⁰⁰ PELLEGRINO, J. W. Important Considerations for Assessment to Function in the Service of Education. Educational Measurement: Issues and Practice [online]. 2020, roč. 39, č. 3, s. 81 – 85. ISSN 1745-3992. Dostupné na: <https://doi.org/10.1111/emip.12372>.

¹⁰¹ BLACK, P. a WILIAM, D. Developing the Theory of Formative Assessment. Educational Assessment, Evaluation and Accountability [online]. 2009, roč. 21, č. 1, s. 5 – 31. ISSN 1874-8600.

¹⁰² Poznámka: Odborníci v NIVaM vypracovali pre učiteľov veľmi zaujímavý a praktický didaktický materiál o novom prístupe k hodnoteniu žiakov. Dostupné online: https://vzdelavanie21.sk/wp-content/uploads/2025/03/NIVAM_Hodnotenie_ziakov.pdf.

Pre vedenie školy môžu slúžiť ako jeden z nástrojov monitorovania kvality vzdelávania. Diagnostické využitie výsledkov však nezávisí iba od kritériálneho charakteru testu, ale aj od jeho konštrukcie, spôsobu škálovania a úrovne podrobnosti poskytovaných informácií. Výsledky by preto nemali byť redukované na jednoduché porovnávanie škôl, tried alebo žiakov.

Podpora motivácie a rozvoja žiaka

Ak sú výsledky CRT sprevádzané zrozumiteľnou spätnou väzbou a využité na stanovenie ďalších krokov učenia, môžu podporovať motiváciu žiakov tým, že im pomáhajú lepšie porozumieť vlastnému pokroku vo vzťahu k cieľom učenia.

Kritériálne testovanie umožňuje zachytiť aj čiastkové pokroky¹⁰³, čo je dôležité najmä pre žiakov so vzdelávacími ťažkosťami alebo zo znevýhodneného prostredia.

Box 5: Ukážka formatívneho využitia výsledkov školy – čítanie s porozumením

Nasledujúca tabuľka predstavuje zjednodušenú ilustračnú ukážku výsledkov, ktoré môže škola dostať po kritériálnom testovaní čítania s porozumením. Žiaci sú zaradení do výkonových úrovní podľa dosiahnutého skóre. Ku každej úrovni je priradený opis typického výkonu a uvedené percentuálne zastúpenie žiakov školy.

Uvedené opisy a percentuálne údaje sú modelové a neslúžia ako návrh konkrétnych hraničných skóre.

Výkonová úroveň	Ilustračný opis typického výkonu	Zastúpenie žiakov školy
Pokročilá	Žiak interpretuje aj nepriamo vyjadrené významy, prepája informácie z viacerých častí textu, posudzuje argumentáciu a svoje závery dokáže zdôvodniť dôkazmi z textu.	15 %
Vyššia stredná	Žiak vyvodzuje závery z nepriamo uvedených informácií, rozpoznáva vzťahy medzi časťami textu a spravidla dokáže svoje odpovede primerane zdôvodniť.	25 %
Stredná	Žiak vyhľadáva explicitné informácie, identifikuje hlavnú myšlienku a vyvodzuje jednoduché závery, ak sú v texte dostatočne podporené.	35 %
Základná	Žiak vyhľadáva jednoduché, priamo uvedené informácie, ale má ťažkosti pri prepájaní informácií, interpretácii implicitných významov a zdôvodňovaní odpovedí.	20 %
Pod základnou úrovňou	Žiak zatiaľ nepreukázal stabilné zvládnutie základných požiadaviek sledovaných testom. Ťažkosti sa môžu prejavovať už pri vyhľadávaní a porozumení priamo uvedených informácií.	5 %

Z výsledkov v modelovom príklade vyplýva, že 40 % žiakov školy dosahuje vyššiu strednú alebo pokročilú úroveň a pri práci s textom typicky zvláda aj vyvodzovanie záverov a interpretáciu nepriamo vyjadrených významov. Ďalších 35 % žiakov dosahuje strednú úroveň, na ktorej spravidla zvládajú vyhľadávanie explicitných informácií, identifikovanie hlavnej myšlienky a jednoduchšie vyvodzovanie záverov. Približne štvrtina žiakov sa nachádza na základnej úrovni alebo pod základnou úrovňou a môže potrebovať cielenejšiu podporu.

¹⁰³ ORBÁNOVÁ, D. Hodnotenie ako súčasť pedagogického procesu. Zborník vedeckých príspevkov [online]. Komárno: Univerzita J. Selyeho, 2019, č. 2, s. 253 – 259. Dostupné na: <https://uk.ujss.sk/dl/3334/22ORBANOVA.pdf>.

Učiteľ môže tieto výsledky využiť ako východisko na diferenciáciu vyučovania. Žiakom na základnej úrovni môže zadávať úlohy zamerané na prepájanie informácií z rôznych častí textu a postupné vyvodzovanie jednoduchých záverov. Žiaci na strednej úrovni môžu pracovať s úlohami vyžadujúcimi interpretáciu implicitných významov a zdôvodnenie odpovede konkrétnou časťou textu. Žiaci na vyšších úrovniach môžu posudzovať argumentáciu autora, porovnávať viaceré texty alebo formulovať a obhajovať vlastnú interpretáciu.

Výsledok školy však neodhaľuje príčiny výkonu jednotlivých žiakov a zaradenie do výkonovej úrovne nemožno chápať ako úplný opis všetkých ich čitateľských zručností. Učiteľ preto prepája informáciu z externého testovania s analýzou žiackych odpovedí, pozorovaním, rozhovorom so žiakom a ďalšími úlohami.

Formatívny prínos výsledkov vzniká až vtedy, keď učiteľ na ich základe určí ďalší pedagogický krok, poskytne žiakom primeranú spätnú väzbu a následne overí ich pokrok.

4.3 Limity interpretácie výsledkov kritériálneho testovania

Pri interpretácii výsledkov kritériálneho testovania (CRT) je potrebné zohľadniť nielen **špecifiká tohto typu merania**, ale aj **všeobecné obmedzenia pedagogického merania**, ktoré sa uplatňujú aj pri normatívnych (NR) testoch. Ich nereflektovanie môže viesť k nesprávnym záverom o výkone žiakov alebo kvalite výučby.

Všeobecné limity pedagogického merania (*platné pre CRT aj normatívne testovanie*)

Kontextuálna závislosť výsledkov: Výsledky testovania odrážajú výkon žiaka v konkrétnom čase a kontexte (napr. trieda, učiteľ, učebné podmienky) a nemožno ich bez ďalšieho overovania generalizovať na iné situácie alebo dlhodobejšiu úroveň schopností.

Obmedzená šírka hodnotených kompetencií: Testové nástroje zachytávajú len vybrané aspekty výkonu definované obsahom testu. Niektoré komplexné kompetencie, ako sú tvorivosť, spolupráca, sociálne zručnosti alebo vybrané aspekty kritického myslenia a komunikácie, možno pomocou štandardizovaného testovania zachytiť len čiastočne alebo nepriamo. Ich komplexné posúdenie si často vyžaduje využitie viacerých hodnotiacich postupov.

Vplyv vonkajších faktorov: Výkon žiaka môže byť ovplyvnený aktuálnym psychickým stavom, motiváciou, zdravotným stavom alebo prostredím, čo môže viesť k dočasným odchýlkam od jeho reálnej úrovne vedomostí a zručností.

Reliabilita merania: Výsledky testovania sú vždy do určitej miery ovplyvnené náhodnými zdrojmi variability. Reliabilita vyjadruje mieru presnosti a stability merania, teda do akej miery by test pri opakovanom použití za porovnateľných podmienok poskytoval podobné výsledky.

Validita interpretácie výsledkov: Validita vyjadruje, do akej miery sú interpretácie a závery vyvodzované z výsledkov testu primerané zamýšľanému účelu merania. Aj test s vysokou reliabilitou môže viesť k nepresným záverom, ak nemeria zamýšľaný konštrukt, alebo ak sú výsledky interpretované nevhodným spôsobom.

Chyba merania: Každý výsledok testovania je zaťažený určitou mierou neistoty. Rozlišovať možno náhodnú chybu merania, ktorá vzniká pôsobením premenlivých faktorov (napr. momentálna pozornosť alebo únava žiaka) a systematickú chybu, ktorá môže byť spôsobená napríklad charakteristikami testu alebo podmienkami

administrácie. Chyba merania má osobitný význam pri interpretácii výsledkov nachádzajúcich sa v blízkosti hraničných skóre (Nitko a Brookhart, 2014)¹⁰⁴.

Limitácie spätnej väzby: Výsledky testovania poskytujú informatívnu hodnotu len vtedy, ak sú odborne interpretované a pedagogicky spracované. Bez primeraného vysvetlenia zo strany učiteľa môžu byť výsledky nesprávne pochopené žiakmi, rodičmi alebo vedením školy.

Špecifické limity kritériálneho testovania

Obmedzená vhodnosť na normatívne testovanie: CRT sú konštruované na posudzovanie miery zvládnutia vopred stanovených cieľov a štandardov, nie na porovnávanie žiakov medzi sebou. Používanie percentilov, poradí alebo iných normatívnych ukazovateľov nie je primárnym účelom CRT. Hoci ich možno za určitých podmienok odvodiť, interpretácia výsledkov by mala vychádzať predovšetkým z miery dosiahnutia stanovených cieľov a výkonových štandardov.

Stanovenie výkonových úrovní: Určovanie hraníc medzi výkonovými úrovňami (*cut scores*) je založené na expertnom posúdení a obsahuje určitú mieru neistoty. Táto neistota musí byť zohľadnená pri rozhodovaní o dosiahnutí alebo nedosiahnutí požadovanej úrovne zvládnutia.

Riziko nesprávneho použitia: CRT neboli primárne konštruované na hodnotenie učiteľov alebo škôl. Ich využitie na účely porovnávania alebo vytvárania zodpovednostných mechanizmov môže viesť k nesprávnym interpretáciám a nežiaducim motivačným efektom.

Pri interpretácii výsledkov CRT je preto nevyhnutné:

- dôsledne rozlišovať medzi **všeobecnými limitmi pedagogického merania** a **špecifikami kritériálneho prístupu**,
- zdôrazňovať **diagnostickú a podpornú funkciu výsledkov**,
- a posudzovať výkon a pokrok žiakov **vo vzťahu k stanoveným cieľom a štandardom**, nie v porovnaní s ostatnými.

¹⁰⁴ NITKO, A. J. a BROOKHART, S. M. Educational Assessment of Students. 7. vyd. Boston: Pearson, 2014. ISBN 978-0-13-257201-7.

5. Implementácia systému kritériálnych testovaní na Slovensku

Kapitola opisuje plánovaný proces zavádzania systému kritériálneho testovania v slovenskom vzdelávacom systéme ako postupný, udržateľný a metodologicky podložený proces s dôrazom na využitie výsledkov CRT. Predstavuje hlavné prvky implementácie – budovanie bánk testových úloh, pilotné a hlavné merania naprieč vzdelávacími cyklami, využitie elektronického testovania a prepojenie výsledkov s dotazníkovými prieskumami. Ťažisko kapitoly je kladené na formatívny potenciál kritériálneho testovania, jeho využitie na podporu učenia sa, zvyšovanie kvality vyučovania a poskytovanie adresnejšej spätnej väzby školám, ako aj na potrebu dlhodobej udržateľnosti, ďalšieho výskumu a úzkeho prepojenia hodnotenia s kurikulárnymi zmenami.

Implementácia kritériálnych testovaní na Slovensku bude prebiehať postupne naprieč jednotlivými vzdelávacími cyklami. Výsledky testovaní nebudú využívané na individuálne rozhodnutia o vzdelávacej dráhe žiakov ani na iné rozhodnutia s vysokými dôsledkami (*high-stakes tests*). Testovania budú preto plniť predovšetkým diagnostickú, monitorovaciu a rozvojovú funkciu. Ich primárnym cieľom bude poskytovať školám, učiteľom a tvorcom vzdelávacej politiky informácie o miere dosahovania vzdelávacích cieľov a výkonových štandardov.

Výsledky budú slúžiť na identifikáciu oblastí, v ktorých žiaci dosahujú požadovanú úroveň výkonu, ako aj oblastí vyžadujúcich dodatočnú pedagogickú podporu. Formatívny potenciál testovania sa pritom nebude odvíjať od samotných výsledkov, ale od spôsobu ich interpretácie a využitia pri plánovaní ďalšieho vyučovania, poskytovaní spätnej väzby a nastavovaní pedagogických intervencií. Pri analýze výsledkov sa predpokladá využívanie moderných psychometrických postupov vrátane modelov teórie odpovede na položku (IRT), ktoré umožňujú presnejšie škálovanie výkonu a odhad neistoty merania.

Podobne ako v niektorých zahraničných systémoch budú testovania slúžiť predovšetkým na podporu učenia a zlepšovanie kvality vyučovania.

5.1 Banka úloh a pilotné testovania

Východiskom pre testovania sú reformy Plánu obnovy a odolnosti, v rámci ktorého odborníci NIVaM odboru hodnotenia a monitorovania vzdelávania (OHMV) vytvárajú od roku 2022 banky testových položiek z:

- čitateľskej a jazykovej gramotnosti pre 1., 2. a 3. cyklus;
- matematickej gramotnosti pre 1., 2. a 3. cyklus;
- prírodovednej gramotnosti pre 2. a 3. cyklus;
- finančnej gramotnosti pre 3. cyklus.

Výsledky testovaní budú poskytovať informácie o úrovni čitateľskej a jazykovej, matematickej a prírodovednej gramotnosti žiakov.

Podľa požiadaviek MŠVVaM sa môžu tieto testovania rozšíriť aj o meranie úrovne digitálnej gramotnosti, prípadne aj globálnych kompetencií. Pojem globálne kompetencie sa používa v súlade s medzinárodnými rámcami hodnotenia a zahŕňa najmä schopnosť analyzovať lokálne a globálne problémy, porozumieť rôznym kultúrnym perspektívam, efektívne komunikovať v rozmanitom prostredí a konať zodpovedne vo vzťahu k udržateľnému rozvoju a občianskej participácii.

Hlavným testovaniam budú vždy predchádzať pilotné merania, v rámci ktorých bude realizované viacúrovňové overovanie testovacích nástrojov z pohľadu kvality položiek a dizajnu testovacieho nástroja.

Výsledky hlavných meraní nebudú slúžiť iba na zisťovanie celkovej úrovne výkonu žiakov. Školám a učiteľom majú poskytovať informácie o tom, ktoré vzdelávacie ciele a výkonové štandardy žiaci zvládajú a ktoré oblasti si vyžadujú dodatočnú pedagogickú podporu. Napríklad, ak výsledky testovania preukážu, že žiaci dosahujú nižšiu úroveň pri interpretácii informácií v texte alebo pri riešení slovných úloh z matematiky, škola môže na tieto oblasti cielene zamerať ďalšie vyučovanie, metodickú podporu učiteľov alebo rozvojové aktivity žiakov. Formatívny potenciál testovania sa tak prejavuje najmä prostredníctvom následnej interpretácie výsledkov a ich využitia pri plánovaní vyučovania a podpore učenia sa.

Tab. 8 Harmonogram pilotných testovaní a hlavných CR-testovaní podľa ročníkov

ročník/pilot	1. pilot	2. pilot	3. pilot	Hlavné meranie
3. ročník	2022/2023	2023/2024	2024/2025	2025/2026
5. ročník	2024/2025	2025/2026	2026/2027	2027/2028
8. ročník	2027/2028	2028/2029	2029/2030	2030/2031

5.2 Hlavné merania

Prvé hlavné meranie sa realizovalo z čitateľskej a jazykovej gramotnosti (ČJG) a matematickej gramotnosti (MG) v období apríl – máj 2026 na vzorke cca 6 000 žiakov 3. ročníka základnej školy, t. j. po ukončení 1. vzdelávacieho cyklu. Hlavné meranie po ukončení 2. cyklu sa bude realizovať okrem ČJG a MG aj z prírodovednej gramotnosti (PG) počnúc školským rokom 2027/2028. V prípade požiadaviek MŠVVaM pribudne v 3. cykle aj testovanie finančnej (FG) a digitálnej gramotnosti (DG), prípadne aj globálnych kompetencií (GK).

Prehľad plánovaného zavádzania jednotlivých oblastí kriteriálneho testovania podľa vzdelávacích cyklov a školských rokov znázorňuje tabuľka 9.

Tab. 9 Časové nastavenie CR-testovania podľa vzdelávacích cyklov a gramotností

Cyklus	1.		2.			3.			
Predmet	ČJG	MG	ČJG	MG	PG	ČJG	MG	PG	FG
2025/2026									
2026/2027									
2027/2028									
2028/2029									
2029/2030									
2030/2031									

Legenda: ČJG = čitateľská a jazyková gramotnosť; MG = matematická gramotnosť; PG = prírodovedná gramotnosť; FG = finančná gramotnosť. Prázdne pole znamená, že v príslušnom školskom roku sa daná oblasť netestuje. Vyfarbené pole označuje plánované hlavné meranie v danej oblasti a vzdelávacom cykle.

Jednotlivé merania sa budú v školách zavádzať postupne a budú realizované na cyklickom princípe z vybraných tém, respektíve oblastí na stratifikovanej vzorke žiakov s presne špecifikovanými znakmi. Systém kritériálnych testovaní bude pozostávať z paralelnej administrácie testov z rôznych oblastí, resp. tém na konci 1., 2. a 3. cyklu vzdelávania na stratifikovanej vzorke žiakov (cca 5 000 žiakov), ktorá bude konštruovaná tak, aby zohľadňovala kľúčové charakteristiky populácie (napr. región, typ školy, veľkosť školy). Vzorke žiakov bude vyberaná tak, aby v rámci daného školského roku bola zapojená každá škola aspoň v jednom CRT. Školy by mali absolvovať hlavné domény cyklicky po 3 až 4 rokoch obdobne, ako je tomu v medzinárodných štúdiách (napr. PISA). Z dlhodobého hľadiska je žiaduce, aby boli po uskutočnení hlavných meraní postupne uvoľňované a školám sprístupňované testové nástroje, ktoré budú môcť školy využiť na vlastné diagnostické a rozvojové účely. Súčasťou zverejnených materiálov by mal byť aj metodický manuál na vyhodnocovanie a interpretáciu výsledkov.

5.3 Dotazníkové prieskumy

Dôležitou súčasťou CR-testovaní na vzorke škôl budú aj dotazníkové prieskumy, ktoré by sa mali zameriavať na monitorovanie sprievodných faktorov vzdelávania vo vzťahu k učiteľom, vedeniu školy, žiakom, rodičom a pod. Dotazníky obsahujúce jasné, jednoznačné a psychometricky overené otázky by mali byť vytvorené s cieľom lepšie interpretovať vzťahy medzi premennými a porozumieť faktorom súvisiacim s dosahovanými výsledkami. Užitočnou inšpiráciou môžu byť aj kvalitne vytvorené dotazníky z medzinárodných meraní (PIRLS, TIMSS, ICILS, PISA), ktoré môžu byť adaptované a rozšírené pre potreby sledovania cieľov a zámerov príslušného CR-testovania. V pravidelných dotazníkových prieskumoch by mali byť vyhodnocované faktory ovplyvňujúce pohodu v školách, interné prostredie vyučovania v triede a v škole, ale aj externé faktory (mimoškolské prostredie) a ich vplyv na vyučovanie v škole.

Osobitne by sa mali podrobne vyhodnocovať aj vplyvy digitálnych technológií a nástrojov umelej inteligencie (AI) na vyučovanie a rozvíjanie zručností žiakov a učiteľov. Z hľadiska dopadov a presahov digitálnych technológií, sociálnych sietí aj AI do každodenného života je vhodné systematicky monitorovať aj digitálnu pohodu žiakov a učiteľov.

Faktory, ktoré ovplyvňujú vyučovací proces, či duševné zdravie žiakov a majú priamy alebo nepriamy vplyv na výsledky ich vzdelávania, sa dynamicky menia a budú sa naďalej meniť, preto je potrebné ich účinne monitorovať, včas vyhodnocovať a komunikovať ako spätnú väzbu aj pre tvorcov kurikula a tvorcov vzdelávacích politík.

Dôležité je, aby bolo možné výsledky žiakov z CRT analyzovať v spojení s údajmi získanými z dotazníkových prieskumov pri zachovaní všetkých požiadaviek na ochranu osobných údajov. Takéto údaje môžu prispieť ku komplexnejšiemu porozumeniu faktorov ovplyvňujúcich vzdelávacie výsledky žiakov, fungovanie tried a procesy vyučovania v škole. Na základe kombinácie výsledkov testovania, dotazníkových údajov a moderných psychometrických postupov, vrátane modelov IRT, bude možné poskytovať presnejšiu spätnú väzbu školám a tvorcom vzdelávacej politiky. Takéto prepojenie výsledkov testovania a kontextových údajov môže významne rozšíriť interpretačné možnosti systému kritériálneho hodnotenia a poskytnúť kvalitnejšie podklady pre rozhodovanie na úrovni školy aj vzdelávacej politiky.

5.4 Elektronické testovania

K efektívnemu zabezpečeniu viacúrovňového pilotného overovania je nevyhnutná administrácia testovaní a dotazníkov v online priestore prostredníctvom vhodného elektronického systému, resp. platformy elektronického testovania (zavedenie e-testovania v Pláne obnovy a odolnosti), kde bude možné efektívne zabezpečiť aj proces vyhodnocovania výsledkov žiakov a poskytovania výstupov školám. Využitie systému e-testovania umožní aj školám, ktoré nebudú zaradené do výberu, sprístupniť testy na dobrovoľnú realizáciu mimo hlavného merania z niektorých obsahových oblastí a vyhodnocovať ich podľa pripravenej metodiky na účely spätnej väzby.

V súčasnosti sa na Slovensku už využíva testovacia platforma TAO (Test Administration Online)¹⁰⁵, adaptovaná na požiadavky MŠVVaM SR a NIVaM, ktorá je rovnako využívaná v mnohých krajinách a bola použitá aj v medzinárodných meraniach vrátane OECD PISA. Licencie na túto platformu umožnia slovenským školám pravidelne sa zapájať do elektronických testovaní a dotazníkových prieskumov pri zachovaní bezpečnosti a ochrany dát a osobných údajov.

Elektronické testovanie prináša výhody nielen z hľadiska administrácie testov, ale aj z pohľadu spracovania a analýzy výsledkov. Umožňuje efektívne zhromažďovanie a uchovávanie dát, automatizované vyhodnocovanie, rýchlejšie poskytovanie spätnej väzby školám a využívanie pokročilých psychometrických analýz. V kombinácii s modelmi teórie odpovede na položku (IRT) vytvára predpoklady na škálovanie výsledkov naprieč rôznymi testovými formami a do budúcnosti aj na postupné zavádzanie adaptívneho testovania, pri ktorom sa náročnosť úloh priebežne prispôsobuje úrovni výkonu žiaka. Zároveň môže elektronické testovanie prispieť aj k rozvoju digitálnych zručností žiakov a učiteľov.

Výsledky CRT z hlavných meraní budú do škôl elektronicky distribuované v podobe výsledkových listín obsahujúcich výsledky žiakov školy z danej testovanej gramotnosti. Zároveň budú z nameraných dát spracované hodnotiace správy za celé Slovensko, ktoré budú poskytovať dôležitú spätnú väzbu pre odbornú verejnosť a tvorcov vzdelávacích politík.

Zavedenie elektronického systému kritériálnych testovaní, prepojeného s modernými psychometrickými postupmi vrátane modelov IRT, vytvorí podmienky na presnejšie meranie výkonu žiakov, kvalitnejšie interpretácie výsledkov a poskytovanie adresnejšej spätnej väzby školám. Významnou perspektívou ďalšieho rozvoja je aj **adaptívne testovanie**, ktoré umožňuje efektívnejšie odhadovať úroveň výkonu žiakov pri menšom počte úloh. Kritériálna interpretácia výsledkov zároveň posilňuje orientáciu systému na dosahovanie vzdelávacích cieľov a podporu učenia.

Preto je dôležité, aby sa systém reportovania výsledkov CRT rozvíjal postupne a v súlade s empirickými poznatkami získavanými z pilotných a hlavných meraní. V počiatočných etapách bude prioritou zabezpečiť spoľahlivé meranie a zrozumiteľné poskytovanie základných výsledkov školám. Následne bude možné postupne rozširovať interpretačnú hodnotu výstupov o podrobnejšie analýzy, opis výkonových profilov

¹⁰⁵ Poznámka: <https://www.taotesting.com/>.

a metodické odporúčania pre pedagogickú prax. Takýto vývoj si vyžaduje dlhodobé budovanie odborných kapacít, databáz úloh, psychometrických modelov a skúseností s interpretáciou výsledkov v podmienkach slovenského vzdelávacieho systému.

5.5 Udržateľnosť meraní, ďalší výskum a vývoj v prepojení s kurikulárnymi zmenami

Dlhodobá implementácia kriteriálneho testovania si vyžaduje zabezpečenie jeho udržateľnosti aj po skončení Plánu obnovy a odolnosti, ako aj systematický výskum a ďalší rozvoj v rezorte školstva.

Výsledky CRT by mali slúžiť nielen ako nástroj spätnej väzby pre školy a učiteľov, ale aj ako **jeden z významných zdrojov údajov pre priebežné precizovanie a revíziu Štátneho vzdelávacieho programu**. Spätá väzba z testovaní má byť systematicky využívaná ako jeden z podkladov pri posudzovaní nastavenia vzdelávacích cieľov, obsahovej náročnosti učiva a očakávaných výstupov učenia v jednotlivých vzdelávacích cykloch.

Výsledky CRT poskytujú informácie o miere dosahovania vzdelávacích cieľov a výkonových štandardov. Nemožno ich však interpretovať ako priamy dôkaz príčin úspešnosti alebo neúspešnosti žiakov, učiteľov, škôl či vzdelávacích politík bez zohľadnenia ďalších kontextových faktorov.

Kľúčovým predpokladom úspechu je **funkčné prepojenie kurikulárneho vývoja a hodnotenia výsledkov vzdelávania**, založené na úzkej spolupráci odborných tímov NIVaM zodpovedných za tvorbu ŠVP a realizáciu a analýzu CRT. Takto vytvorený mechanizmus spätnej väzby umožní systematicky zlepšovať kurikulum, podporovať kvalitu výučby a zvyšovať efektívnosť vzdelávacieho systému.

Z dlhodobého hľadiska je nevyhnutné investovať do odborného rastu a kontinuálneho vzdelávania expertov, pravidelne vyhodnocovať dopady testovania na kvalitu vzdelávania a jasne ukotviť úlohu CRT v kontexte kurikulárnych reforiem. Súčasťou strategického smerovania má byť aj rozvoj medzinárodnej spolupráce, sledovanie inovácií v oblasti psychometrie a formatívneho hodnotenia a prepojenie národného systému CRT s osvedčenými postupmi krajín OECD a medzinárodných konzorcií (napr. CIDREE).

ZÁVER: PODMIENKY A SMEROVANIE KRITERIÁLNEHO TESTOVANIA NA SLOVENSKU

Zavedenie kritériálneho testovania môže významne rozšíriť informačnú hodnotu externého hodnotenia na Slovensku. Jeho prínos však nebude vyplývať zo samotného označenia testov za kritériálne, ale z kvality celého systému: od obsahového vymedzenia a operacionalizácie vzdelávacích štandardov cez tvorbu a pilotné overovanie úloh až po psychometrické vyhodnotenie, stanovovanie výkonových úrovní, reportovanie výsledkov a podporu ich správneho využívania školami.

V kontexte kurikulárnej reformy zavádzanej na Slovensku môže kritériálne testovanie predstavovať významný nástroj monitorovania implementácie zmien vo vzdelávaní. Umožňuje priebežne získavať informácie o tom, do akej miery žiaci po ukončení jednotlivých vzdelávacích cyklov dosahujú očakávané vzdelávacie ciele a výkonové štandardy vymedzené v novom kurikule. Na rozdiel od prístupov zameraných predovšetkým na porovnávanie výsledkov žiakov alebo škôl umožňuje obsahovo interpretovať, ktoré vedomosti, zručnosti a kompetencie žiaci pravdepodobne zvládajú, a v ktorých oblastiach sa prejavuje potreba ďalšej podpory. Výsledky môžu slúžiť ako jeden zo zdrojov spätnej väzby pre školy, tvorcov kurikula a vzdelávaciu politiku a spolu s ďalšími kvalitatívnymi a kvantitatívnymi údajmi podporovať priebežné vyhodnocovanie implementácie reformy. Pri vhodnom nastavení systému tak môže kritériálne testovanie prispieť nielen k monitorovaniu dosahovania cieľov kurikulárnej reformy, ale aj k jej odbornému usmerňovaniu a zdokonaľovaniu na základe dôkazov.

Skúsenosti analyzovaných krajín prinášajú pre Slovensko niekoľko zásadných poučení. Nemecko ukazuje význam odborného konsenzu o vzdelávacích štandardoch, transparentného prepojenia štandardov s testovými úlohami a dlhodobého budovania dôvery v systém hodnotenia. Estónsko zdôrazňuje prepojenie externého hodnotenia s autoevalváciou, kvalitnou spätnou väzbou, profesijnou autonómiou učiteľov a pripravenosťou škôl pracovať s dátami. Portugalská skúsenosť upozorňuje, že zmeny vzdelávacích výsledkov nemožno pripísať samotnému testovaniu, ale súbehu kurikulárnych, pedagogických, personálnych a sociálnych opatrení. Zároveň poukazuje na riziká testovania s vysokými dôsledkami, najmä na tlak na výkon či vyučovanie orientované na testovanie.

Z toho vyplýva, že slovenský model kritériálneho testovania má byť koncipovaný predovšetkým ako testovanie s nízkymi dôsledkami. Jeho výsledky nemajú byť používané na sankcionovanie škôl, hodnotenie učiteľov, vytváranie verejných rebríčkov ani ako jediný podklad na rozhodovanie o vzdelávacej dráhe jednotlivých žiakov. Takéto použitie by presahovalo zamýšľaný účel testov a vyžadovalo by podstatne silnejšie dôkazy o validite interpretácií, presnosti merania, spravodlivosti rozhodnutí a dôsledkoch takéhoto využitia. Už pri návrhu systému je preto potrebné vytvoriť pravidlá, ktoré zabránia zjednodušenému alebo neprimeranému používaniu výsledkov ako nástroja kontroly alebo tlaku na školy.

Osobitnou výzvou slovenského prostredia je obsahová šírka kurikulárnych dokumentov a vzťah medzi štátnym a školským vzdelávacím programom. Ak sú štandardy formulované rámcovo a školy ich konkretizujú rozdielnym spôsobom, môže vzniknúť spor o to, či vybrané testové úlohy dostatočne reprezentujú obsah a náročnosť vzdelávania vo všetkých školách. Kritika výberu položiek môže byť pri kritériálnej interpretácii obzvlášť výrazná, pretože výsledok sa vzťahuje na tvrdenie o dosiahnutí konkrétneho štandardu alebo výkonovej úrovne. Podrobnejšie vymedzenie štandardov môže podporiť obsahový súlad testu s kurikulumom, samo osebe však validitu nezaručuje.

Kľúčovou podmienkou je preto transparentná operacionalizácia vzdelávacích štandardov založená na odbornom konsenze. Musí byť možné preukázať súlad medzi štátnym kurikulumom, teoretickým rámcom testovania, špecifikáciou testu, zastúpením obsahových a kognitívnych oblastí, jednotlivými úlohami a zamýšľanou interpretáciou výsledkov. Do tvorby rámcov a posudzovania úloh je vhodné zapojiť predmetových odborníkov, učiteľov z rozličných typov škôl, odborníkov na pedagogické meranie a ďalších relevantných aktérov. Súčasťou systému má byť aj zverejňovanie rámcov testovania, špecifikácií a primeraného počtu ukážkových úloh.

Psychometrické smerovanie systému musí byť v koncepcii pomenované jednoznačne. Klasická teória testov zostáva využiteľná pri položkovej analýze, pilotnom overovaní a základnom posudzovaní kvality testov. Pre dlhodobé prepájanie testových foriem, tvorbu spoločných škál, vyjadrovanie neistoty odhadu a rozvoj adaptívneho testovania do budúcnosti je však potrebné systematicky budovať podmienky na využívanie IRT. Prínosy IRT nie sú automatické; závisia od kvality položiek a dát, primeranosti zvoleného modelu a splnenia jeho predpokladov.

Udržateľnosť systému si vyžaduje vytvorenie stabilnej domácej odbornej komunity odborníkov na pedagogické meranie – edukometriu a psychometriu. Potrebné je rozvíjať odborné kapacity nielen v NIVaM, ale aj na vysokých školách a výskumných pracoviskách, ktoré sa budú dlhodobo venovať tvorbe testov, modelovaniu dát, validácii interpretácií výsledkov a výskumu dôsledkov testovania. Bez kontinuálneho rozvoja týchto kapacít by bol systém závislý od jednotlivcov alebo krátkodobých projektov.

Súčasne bude potrebné investovať do systematického prenosu zahraničného know-how, metodických partnerstiev a odbornej prípravy domácich tímov, najmä v oblasti IRT, tvorby a správy kalibrovaných bánk úloh, prepájania testových foriem a počítačovo adaptívneho testovania. Adaptívne testovanie si vyžaduje digitálnu administráciu, priebežný odhad úrovne meranej schopnosti, algoritmický výber položiek a dostatočne rozsiahlu, obsahovo vyváženú a kalibrovanú banku úloh. Potrebná je tiež bezpečná technologická infraštruktúra, hoci samotné pripojenie testovania na internet nie je postačujúcou podmienkou jeho adaptivity. Adaptívne testovanie preto treba chápať ako dlhodobý rozvojový cieľ, nie ako automatický dôsledok elektronizácie testov.

Kritériálne výsledky môžu podporovať formatívne hodnotenie iba vtedy, ak sa neobmedzia na celkové skóre alebo zaradenie do výkonovej úrovne. Formatívny prínos vzniká až vtedy, keď výsledky poskytujú primerane spoľahlivé dôkazy o tom, ktoré vzdelávacie ciele žiak pravdepodobne zvláda, kde sa prejavujú ťažkosti, a aké ďalšie pedagogické alebo učebné kroky možno zvážiť. Vyžaduje si to dostatočné zastúpenie jednotlivých obsahových oblastí v teste, zrozumiteľné správy, metodickú podporu škôl, čas na spoločnú interpretáciu výsledkov a prepojenie externých údajov s ďalšími dôkazmi o učení. Externé testovanie nemôže nahradiť odborný úsudok učiteľa ani vnútorné hodnotiace procesy školy.

Riziká je potrebné pomenovať a priebežne monitorovať. Patria medzi ne najmä nedostatočný súlad testov s kurikulumom, spochybňovanie obsahovej reprezentatívnosti vybraných úloh, neprimerane kategorická interpretácia výsledkov, neprimerané prispôbovanie vyučovania obsahu a formátu testov na úkor ďalších vzdelávacích cieľov. Ďalšie riziká predstavujú nerovnaké technologické podmienky škôl, preceňovanie výsledkov pri posudzovaní kvality školy a nedostatočná pripravenosť používateľov pracovať s dátami. Zavádzanie systému preto musí byť postupné, výskumne sprevádzané a pravidelne vyhodnocované. Zistené nedostatky majú viesť predovšetkým k zdokonaľovaniu testovacích nástrojov, spôsobu reportovania a podpory škôl, nie k unáhľeným zmenám kurikula ani k sankcionovaniu škôl.

Zavádzanie kritériálne orientovaného testovania nemožno chápať ako jednorazovú výmenu jedného typu testov za iný. Ide o dlhodobé budovanie systému, ktorý prepája kurikulum, kvalitné meranie, odborne podloženú interpretáciu výsledkov, autoevalváciu škôl, digitálne technológie a profesionálne využívanie dát. Ak budú školy dostávať dostatočne presné a zrozumiteľne interpretované výsledky, ak budú zamýšľané interpretácie a spôsoby ich použitia podložené dôkazmi o validite, a ak školy získajú podporu pri práci s týmito informáciami, externé hodnotenie môže prispieť ku kvalitnejšiemu sebahodnoteniu škôl i žiakov a celkovo k systematickému zlepšovaniu vzdelávania.

ZOZNAM POUŽITEJ LITERATÚRY

- AFONSO, N. a COSTA, E. The influence of the Programme for International Student Assessment (PISA) on policy decision in Portugal: the education policies of the 17th Portuguese Constitutional Government. *Sísifo. Educational Sciences Journal* [online]. 2009, č. 10, s. 53 – 64. Dostupné na: <https://sisifo.fpce.ul.pt>.
- ALLEN, M. J. a YEN, W. M. *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press, 2002. ISBN 978-1-57766-230-3.
- BANDURA, A. *Self-efficacy: The Exercise of Control*. New York: W. H. Freeman, 1997. ISBN 978-0716728504.
- BENNETT, R. E., BRASWELL, J., ORANJE, A., SANDENE, B., KAPLAN, B. a YAN, F. Does it Matter if I Take My Mathematics Test on Computer? A Second Empirical Study of Mode Effects in NAEP. *The Journal of Technology, Learning and Assessment* [online]. 2008, roč. 6, č. 9. Dostupné na: <https://ejournals.bc.edu/index.php/jtla/article/view/1639>.
- BERGBAUER, A. B., HANUSHEK, E. A. a WOESSMANN, L. *Testing*. NBER Working Paper [online]. 2018, č. 24836. Cambridge, MA: National Bureau of Economic Research. Dostupné na: <https://doi.org/10.3386/w24836>.
- BLACK, P. a WILIAM, D. *Assessment and classroom learning*. *Assessment in Education: Principles, Policy & Practice*. 1998, roč. 5, č. 1, s. 7 – 74. ISSN 0969-594X.
- BLACK, P. a WILIAM, D. *Developing the Theory of Formative Assessment*. *Educational Assessment, Evaluation and Accountability* [online]. 2009, roč. 21, č. 1, s. 5 – 31. ISSN 1874-8597. Dostupné na: <https://doi.org/10.1007/s11092-008-9068-5>.
- BLOOM, B. S. *Learning for mastery*. *Evaluation Comment*. 1968, roč. 1, č. 2, s. 1 – 12.
- BREAKSPEAR, S. *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*. *OECD Education Working Papers* [online]. 2012, č. 71, s. 1 – 31. Paris: OECD Publishing. Dostupné na: <https://doi.org/10.1787/5k9dfqffr28-en>.
- BROOKHART, S. M. *How to Give Effective Feedback to Your Students*. Alexandria, VA: ASCD, 2008. ISBN 978-1-4166-0736-6.
- BUSSE, J., BECKER, B., WEIRICH, S. a SCHIPOLOWSKI, S. *IQB Trends in Student Achievement 2018*. *Journal of Open Psychology Data* [online]. 2022, roč. 10, č. 1, čl. 18. Dostupné na: <https://doi.org/10.5334/jopd.71>.
- CARLSON, J. E. a VON DAVIER, M. *Item Response Theory*. In: BENNETT, R. a VON DAVIER, M. (eds.). *Advancing Human Assessment. Methodology of Educational Measurement and Assessment* [online]. Cham: Springer, 2017, s. 127 – 174. Dostupné na: https://doi.org/10.1007/978-3-319-58689-2_5.
- CARVALHO, L. M., COSTA, E. a GONÇALVES, C. *Fifteen years looking at the mirror: On the presence of PISA in education policy processes (Portugal, 2000–2016)*. *European Journal of Education* [online]. 2017, roč. 52, č. 2, s. 154 – 166. Dostupné na: <https://doi.org/10.1111/ejed.12210>.
- CIZEK, G. J. a BUNCH, M. B. *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Thousand Oaks, CA: Sage Publications, 2007. ISBN 978-1-4129-1683-7.

CRATO, N. Curriculum and Educational Reforms in Portugal: An Analysis on Why and How Students' Knowledge and Skills Improved. In: REIMERS, F. M. (ed.). Audacious Education Purposes: How Governments Transform the Goals of Education Systems [online]. Cham: Springer, 2020, s. 209 – 231. Dostupné na: https://doi.org/10.1007/978-3-030-41882-3_8.

CSAPÓ, B., MOLNÁR, G. a NAGY, J. Computer-based assessment of school readiness and early reasoning. Journal of Educational Psychology [online]. 2014, roč. 106, č. 3, s. 639 – 650. Dostupné na: <https://doi.org/10.1037/a0035756>.

DAHAN GOLAN, D., BARZILLAI, M. a KATZIR, T. The effect of presentation mode on children's reading preferences, performance, and self-evaluations. Computers & Education [online]. 2018, roč. 126, s. 346 – 358. Dostupné na: <https://doi.org/10.1016/j.compedu.2018.08.001>.

DE AYALA, R. J. The Theory and Practice of Item Response Theory. New York: The Guilford Press, 2009. ISBN 978-1-59385-869-8.

EDUCATION ESTONIA. Estonian education system [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/about-education-system>.

EDUCATION ESTONIA. How did Estonia become a new role model in digital education? [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/how-did-estonia-become-a-new-role-model-in-digital-education>.

EDUCATION ESTONIA. Voluntary, yet attractive and powerful low-stakes assessment [online]. Tartu: Ministry of Education and Research of Estonia. Dostupné na: <https://www.educationestonia.org/innovation/assessment>.

EJIM, S. An Overview of Computer Based Test [online]. 2017. Dostupné na: <https://doi.org/10.13140/RG.2.2.32040.88326>.

ENO, L. P. Comparing the Reading Performance of High-Achieving Adolescents: Computer-Based Testing versus Paper/Pencil [online]. South Orange: Seton Hall University, 2011. Dizertačná práca. Dostupné na: <https://search.proquest.com/docview/1018431561>.

EURYDICE. Germany: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/germany/overview>.

EURYDICE. Estonia: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/estonia/overview>.

EURYDICE. Portugal: Overview [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/portugal/overview>.

EURYDICE. Portugal: Key features of the education system [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/portugal/overview>.

EURYDICE. Estonia – 4. Single-structure primary and lower secondary education: 4.3 Assessment in single-structure education [online]. Brussels: European Education and Culture Executive Agency (EACEA). Dostupné na: <https://eurydice.eacea.ec.europa.eu/eurypedia/estonia/assessment-single-structure-education>.

EYRE, J. On or off screen reading in a digital world. set: Research Information for Teachers [online]. 2017, č. 1, s. 53 – 58. Dostupné na: <https://doi.org/10.18296/set.0072>.

FERREIRA, A. S., FLORES, I. a CASAS-NOVAS, T. Porque melhoraram os resultados PISA em Portugal? Estudo longitudinal e comparado (2000-2015) [online]. Lisboa: Fundação Francisco Manuel dos Santos, 2017. ISBN 978-989-8863-36-2. Dostupné na: <https://ffms.pt/sites/default/files/2022-08/introducao-ao-estudo-porque-melhoraram-os-resultados-pisa-em-portugal.pdf>.

FICOVÁ, L. a PICHANIČOVÁ, I. Kriteriaálny test ako nástroj na mapovanie absolútneho výkonu žiaka. Pedagogické rozhľady [online]. 2022, č. 3, s. 20 – 26. Banská Bystrica: NIVAM. Dostupné na: <https://nivam.sk/wp-content/uploads/2025/02/pr-3-2022.pdf>.

GLASER, R. Instructional Technology and the Measurement of Learning Outcomes: Some Questions. American Psychologist [online]. 1963, roč. 18, č. 8, s. 519 – 521. Dostupné na: <https://doi.org/10.1037/h0049294>.

GOETHE-INSTITUT. School system – Mein Weg nach Deutschland [online]. 2026. Dostupné na: <https://www.goethe.de/prj/mwd/en/idl/ssy.html>.

GOMES, J. F., HORMIGO, I. a MARÔCO, J. Primeiras lições do PISA: dois países a seguir e Portugal no mundo. Iniciativa Educação [online]. 3. november 2019. Dostupné na: <https://www.iniciativaeducacao.org/pt/ed-on/artigos/estatisticas/primeiras-licoes-do-pisa-dois-paises-a-seguir-e-portugal-no-mundo>.

HAMBLETON, R. K., SWAMINATHAN, H. a ROGERS, H. J. Fundamentals of Item Response Theory. Newbury Park, California: SAGE Publications, 1991. 174 s. Measurement Methods for the Social Sciences, zv. 2. ISBN 978-0-8039-3647-8.

HATTIE, J. a TIMPERLEY, H. The Power of Feedback. Review of Educational Research [online]. 2007, roč. 77, č. 1, s. 81 – 112. Dostupné na: <https://doi.org/10.3102/003465430298487>.

HELMKE, A. Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts [Quality of Classroom Practices and Teachers' Professionalism. Diagnosis, Evaluation and Improvement of Lessons]. 4. vyd. Seelze: Kallmeyer, 2012. ISBN 978-3-7800-1009-4.

HELUS, Z. Dítě v osobnostním pojetí: obrat k dítěti jako výzva a úkol pro učitele i rodiče. 2., přeprac. a rozš. vyd. Praha: Portál, 2009. 286 s. ISBN 978-80-7367-628-5.

INNOVE. Üldhariduskoolide 2018. aasta rahuloluküsitluste tulemused. Aruanne [General Education Schools 2018 Satisfaction Survey Results: Report]. Tallinn: Innove, 2019.

INSTITUTE FOR EDUCATIONAL QUALITY IMPROVEMENT (IQB). Educational Standards. Institute for Educational Quality Improvement (IQB) [online]. Dostupné na: <https://www.iqb.hu-berlin.de/en/bista>.

JERRIM, J., MICKLEWRIGHT, J., HEINE, J.-H., SALZER, C. a McKEOWN, C. PISA 2015: How big is the 'mode effect' and what has been done about it? Oxford Review of Education [online]. 2018, roč. 44, č. 4, s. 476 – 493. Dostupné na: <https://doi.org/10.1080/03054985.2018.1430025>.

JUKK, H. Assessment of mathematical competence using the Estonian national e-tests. Eesti Haridusteaduste Ajakiri. Estonian Journal of Education [online]. 2021, roč. 9, č. 2, s. 110 – 136. Tartu: University of Tartu. Dostupné na: <https://doi.org/10.12697/EHA.2021.9.2.05>.

KANE, M. T. Validating the Interpretations and Uses of Test Scores. Journal of Educational Measurement [online]. 2013, roč. 50, č. 1, s. 1 – 73. ISSN 0022-0655. Dostupné na: <https://doi.org/10.1111/jedm.12000>.

KIKAS, E., MÄDAMÜRK, K., TREIAL, K., LUPTOVA, O., MALLEUS, E., SOODLA, P. a kol. Arvutipõhised hindamisvahendid õpi-, suhtlus- ja enesemääratluspädevuse hindamiseks I ja II kooliastmes [Computer-based Assessment Instruments for Measuring Learning, Communication and Self-regulation Skills in Key Stages I and II]. Tallinn: Tallinn University, 2018.

KULTUSMINISTERKONFERENZ. Bildungsstandards der Kultusministerkonferenz: Erläuterungen zur Konzeption und Entwicklung [online]. München; Neuwied: Wolters Kluwer Deutschland GmbH, 2005. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Konzeption-Entwicklung.pdf.

KULTUSMINISTERKONFERENZ. Die Kultusministerkonferenz als Garant für gemeinsames Handeln in Bildung, Wissenschaft und Kultur [online]. Dostupné na: <https://www.kmk.org/kultusministerkonferenz.html>.

KULTUSMINISTERKONFERENZ. Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring [Strategic Approach of the Standing Conference of the Ministers of Education and Cultural Affairs on Educational Monitoring] [online]. Berlin: Wolters Kluwer Deutschland GmbH, 2016. ISBN 9783556071069. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf.

KULTUSMINISTERKONFERENZ. Vereinbarung zur Weiterentwicklung der Vergleichsarbeiten (VERA): Beschluss der Kultusministerkonferenz vom 08.03.2012 i. d. F. vom 15.03.2018 [online]. Berlin: Kultusministerkonferenz, 2018. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_03_08-Weiterentwicklung-VERA.pdf.

KYLLONEN, P. C. New constructs, methods, & directions for computer-based assessment. In: SCHEUERMANN, F. a BJÖRNSSON, J. (eds.). The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing [online]. 2009, roč. 2009, č. JRC49408, s. 151 – 156. Dostupné na: <https://publications.jrc.ec.europa.eu/repository/handle/JRC49408>.

LEES, M. Estonian Education System 1990-2016: Reforms and Their Impact [online]. 4liberty.eu, 2017. Dostupné na: https://4liberty.eu/wp-content/uploads/2016/08/Estonian-Education-System_1990-2016.pdf.

MAGER, R. F. Preparing Instructional Objectives. Palo Alto, California: Fearon Publishers, 1962. 62 s.

MARÔCO, J. ...E pur si muove! Observador [online]. 11. január 2018. Dostupné na: <https://observador.pt/opiniao/e-pur-si-muove>.

MARÔCO, J. PISA 2018: Portugal é ainda a maior história de sucesso na Europa? Iniciativa Educação [online]. 2. november 2019. Dostupné na: <https://www.iniciativaeducacao.org/pt/ed-on/artigos/estatisticas/pisa-2018>.

MARÔCO, J. International large-scale assessments: Trends and effects on the Portuguese public education system. In: HARJU-LUUKKAINEN, H., McELVANY, N. a STANG, J. (eds.). Monitoring Student Achievement in the 21st Century: European Policy Perspectives and Assessment Strategies [online]. Cham: Springer, 2020, s. 207 – 222. Dostupné na: <https://doi.org/10.1007/978-3-030-38969-7>.

MARÔCO, J. Portugal: The PISA Effects on Education. In: CRATO, N. (ed.). Improving a Country's Education [online]. Cham: Springer, 2021, s. 159 – 174. Dostupné na: https://doi.org/10.1007/978-3-030-59031-4_8.

MINISTERSTVO ŠKOLSTVA, VEDY, VÝSKUMU A ŠPORTU SR. Vzdelávacie štandardy: Vzdelávacia oblasť Matematika a informatika [online]. Dostupné na: https://www.minedu.sk/data/files/11820_matematika-a-informatika.pdf.

MINISTRY OF EDUCATION AND RESEARCH. The Estonian Lifelong Learning Strategy 2020 [online]. Tartu: Republic of Estonia, Ministry of Education and Research, 2014. Dostupné na: https://www.educationestonia.org/wp-content/uploads/2022/12/estonian_lifelong_strategy.pdf.

MINISTRY OF EDUCATION AND RESEARCH. Education Strategy 2021-2035 [online]. Tartu: Republic of Estonia, Ministry of Education and Research, 2021. Dostupné na: https://www.educationestonia.org/wp-content/uploads/2022/12/haridusvaldkonna_arengukava_2035_kinnittaud_vv_eng_0-1.pdf.

NITKO, A. J. a BROOKHART, S. M. Educational Assessment of Students. 7. vyd. Boston: Pearson, 2014. ISBN 978-0-13-343649-5.

NÁRODNÝ ÚSTAV CERTIFIKOVANÝCH MERANÍ VZDELÁVANIA (NÚCEM). Kriteriaálne testy (Program SRSP): Vyvíjanie kriteriaálnych testov pre primárne a nižšie sekundárne vzdelávanie na Slovensku [online]. Bratislava: NÚCEM. Dostupné na: <https://www2.nucem.sk/sk/projekty-a-linky/kriterialne-testy>.

NUNES, L. C., FREITAS, P. a COLAÇO, R. Trends in Student Achievement in Portugal: What Does PISA Tell Us? Porto: Fundação Belmiro de Azevedo – EDULOG, 2024. ISBN 978-989-35739-0-7. Dostupné na: <https://www.edulog.pt/storage/app/uploads/public/666/875/9da/6668759daa0d6352676543.pdf>.

OECD. Annual GDP and consumption per capita, US \$, current prices, current PPPs. OECD Data Explorer [online]. [cit. 2026-08-17]. Dostupné na: <https://data-explorer.oecd.org>.

OECD. Synergies for Better Learning: An International Perspective on Evaluation and Assessment. Paris: OECD Publishing, 2013. ISBN 978-9264190658.

OECD. Estonia Country Note, PISA 2018 Results: Where All Students Can Succeed [online]. Paris: OECD Publishing, 2019. Dostupné na: https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/pisa/publications/national-reports/pisa-2018/featured-country-specific-overviews/PISA2018_CN_EST.pdf.

OECD. PISA 2022 Results: Factsheets – Estonia [online]. Paris: OECD Publishing, 2023. Dostupné na: https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/estonia_dafed886-en.html.

OECD. PISA 2025 Integrated Design [online]. Paris: OECD, 2023 [cit. 2026-08-02]. Dostupné na: <https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/pisa/pisa-database/survey-implementation-tools/pisa-2025/PISA-2025-Integrated-Design.pdf>.

ORBÁNOVÁ, D. Hodnotenie ako súčasť pedagogického procesu. Zborník vedeckých príspevkov [online]. Komárno: Univerzita J. Selyeho, 2019, č. 2, s. 253 – 259. Dostupné na: <https://uk.ujs.sk/dl/3334/22ORBANOVA.pdf>.

- PELLEGRINO, J. W. Important Considerations for Assessment to Function in the Service of Education. Educational Measurement: Issues and Practice [online]. 2020, roč. 39, č. 3, s. 81 – 85. ISSN 1745-3992. Dostupné na: <https://doi.org/10.1111/emip.12372>.
- PHILLIPS, G. W. a JIANG, T. Using PISA as an International Benchmark in Standard Setting. Journal of Applied Measurement [online]. 2015, roč. 16, č. 2, s. 161 – 170. Dostupné na: <https://www.wikidata.org/wiki/Q40834623>.
- PIAGET, J. a INHELDEROVÁ, B. Psychologie dítěte. Prel. E. Vyskočilová. Praha: Portál, 2014. 144 s. ISBN 978-80-262-0691-0.
- POPHAM, W. J. Classroom Assessment: What Teachers Need to Know. 8. vyd. Boston: Pearson, 2017. ISBN 978-0-13-405386-8.
- PORTUGAL. MINISTÉRIO DA EDUCAÇÃO, CIÊNCIA E INOVAÇÃO. Novo modelo de avaliação externa dos alunos a partir de 2024/2025 [online]. 2024. Dostupné na: <https://portugal.gov.pt/gc24/comunicacao/comunicados/novo-modelo-de-avaliacao-externa-dos-alunos-a-partir-de-20242025>.
- PRŮCHA, J. Psychologie učení: teoretické a výzkumné poznatky pro edukační praxi. Praha: Grada, 2020. 272 s. ISBN 978-80-271-2853-2.
- PUPALA, B. Portugalská cesta k úspěchu: Konzervativní oblúk v kurikulárnej politike. Pedagogické rozhľady [online]. 2022, č. 3, s. 27 – 32. Banská Bystrica: NIVAM. Dostupné na: <https://nivam.sk/wp-content/uploads/2025/02/pr-3-2022.pdf>.
- SADLER, D. R. Formative Assessment and the Design of Instructional Systems. Instructional Science [online]. 1989, roč. 18, č. 2, s. 119 – 144. Dostupné na: <https://doi.org/10.1007/BF00117714>.
- SEKRETARIAT DER STÄNDIGEN KONFERENZ DER KULTUSMINISTER DER LÄNDER IN DER BUNDESREPUBLIK DEUTSCHLAND. Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung [online]. Berlin: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2010. ISBN 978-3-556-02609-0. Dostupné na: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_00_00-Konzeption-Bildungsstandards.pdf.
- STØLE, H., MANGEN, A. a SCHWIPPERT, K. Assessing children's reading comprehension on paper and screen: A mode-effect study. Computers & Education [online]. 2020, roč. 151, čl. 103861. Dostupné na: <https://doi.org/10.1016/j.compedu.2020.103861>.
- STATNIPRIJIMACKY.CZ. Přijímačky Cermat: testy, řešení, tipy [online]. Dostupné na: <https://www.statniprijimacky.cz/cermat>.
- SWELLER, J. Cognitive load theory and educational technology. Educational Technology Research and Development [online]. 2020, roč. 68, č. 1, s. 1 – 16. Dostupné na: <https://doi.org/10.1007/s11423-019-09701-3>.
- ŠTUKA, Č. a VEJRAŽKA, M. Testování a hodnocení studentů na VŠ [online]. 1. vyd. Praha: Univerzita Karlova, Nakladatelství Karolinum, 2021. ISBN 978-80-246-5108-8. Dostupné na: <https://doi.org/10.14712/9788024651088>.

THE GLOBAL ECONOMY. Portugal: PISA reading scores [online]. Dostupné na: https://www.theglobaleconomy.com/portugal/pisa_reading_scores/.

THOMPSON, N. A. a WEISS, D. J. Computerized and adaptive testing in educational assessment. In: SCHEUERMANN, F. a BJÖRNSSON, J. (eds.). The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-Scale Testing [online]. 2009, roč. 2009, č. JRC49408, s. 127 – 133. Dostupné na: <https://publications.jrc.ec.europa.eu/repository/handle/JRC49408>.

TIRE, G. Educational Assessment in Estonia. In: HARJU-LUUKKAINEN, H., McELVANY, N. a STANG, J. (eds.). Monitoring Student Achievement in the 21st Century: European Policy Perspectives and Assessment Strategies [online]. Cham: Springer, 2020, s. 119 – 129. Dostupné na: https://doi.org/10.1007/978-3-030-38969-7_10.

ÚRAD VLÁDY SR. Plán obnovy: Domov [online]. 2021. Dostupné na: <https://www.planobnovy.sk>.

VALK, A. Õpilaste ja õpetajate heaolu üle-Eestilises rahuloluküsitluses, lühiülevaade 2017. a tulemustest [Student and Teacher Well-Being in a Pan-Estonian Satisfaction Survey, Overview of Results from 2017]. Tartu: Ministry of Education and Research, 2018.

VYGOTSKIJ, L. S. Psychologie myšlení a řeči. Texty vybral, přel., uspořádal a poznámkami doplnil J. Průcha. 2., uprav. vyd. Praha: Portál, 2017. 152 s. ISBN 978-80-262-1258-4.

WAGNER, I., LOESCHE, P. a BIßANTZ, S. Low-stakes performance testing in Germany by the VERA assessment: analysis of the mode effects between computer-based testing and paper-pencil testing. European Journal of Psychology of Education [online]. 2022, roč. 37, s. 531 – 549. Dostupné na: <https://doi.org/10.1007/s10212-021-00532-6>.

Koncepcia zavádzania kritériálneho testovania na Slovensku

Východiská, podmienky a smerovanie

Autor:

PhDr. Romana Kanovská

Interná recenzia:

PaedDr. Ivana Pichaničová, PhD., PaedDr. Lucia Ficová, PhD.,
Mgr. Jakub Mikuška Ph.D., PaedDr. Janka Kurajová Stopková, PhD.,
Mgr. Martina Pigová, Mgr. Milan Fico, PhD.

Externá recenzia:

doc. RNDr. Juraj Vantuch, CSc.

Jazyková úprava:

Mgr. Alžbeta Palacková

Grafická úprava:

Mgr. Martina Dienešová

Vydal: NIVaM

Vydanie: prvé

Formát: elektronicky

Rok vydania: 2026

ISBN: 978-80-565-1726-0

EAN: 9788056517260

