

OSTRAVSKÁ UNIVERZITA
PEDAGOGICKÁ FAKULTA



UNIVERSITY
OF OSTRAVA
PEDAGOGICAL FACULTY

AKTUÁLNÍ PROBLÉMY PLOŠNÉHO TESTOVÁNÍ

**JOSEF MALACH, MARTIN MALČÍK,
MARTINA ROZSYPALOVÁ (EDS.)**

OSTRAVA 2016

Za obsahovou a jazykovou správnost jednotlivých kapitol odpovídají jejich autoři.

Recenzovala: doc. RNDr. Mária Lucká, PhD.
Slovenská technická univerzita v Bratislave

Název: **Aktuální problémy plošného testování**
Editoři: doc. PhDr. Josef Malach, CSc., doc. RNDr. Martin Malčík, Ph.D.,
Mgr. Bc. et Bc. Martina Rozsypalová
Vydání: první, 2016
Počet stran: 128
Tisk: REPRONIS s. r. o., Ostrava

Tato publikace byla schválena vědeckou redakcí Pedagogické fakulty Ostravské univerzity.

Vydavatel: Ostravská univerzita

© Ostravská univerzita

ISBN 978-80-7464-867-0

OBSAH

EDITORIAL.....	4
<i>Wojciech MAŁECKI, Henryk SZALENIEC</i>	
EWOLUCJA EGZAMINU MATURALNEGO W POLSCE ZAŁOŻENIA I PIERWSZE EFEKTY	7
<i>Lucia FICOVÁ, Ivana PICHANIČOVÁ</i>	
EXTERNÉ TESTOVANIA NA SLOVENSKU – OD PAPIEROVÝCH TESTOV K E-TESTOVANIU.....	21
<i>Josef MALACH, Martin MALČÍK</i>	
VÝZKUM PARAMETRŮ OVLIVŇUJÍCÍCH POČÍTAČEM PODPOROVANOU PEDAGOGICKOU DIAGNOSTIKU	45
<i>Henryk SZALENIEC, Wojciech MAŁECKI</i>	
E-OCENIANIE – OD PROJEKTU DO WDROŻENIA.....	75
<i>Ewa WYSOCKA</i>	
POSITIVE DIAGNOSIS IN A TEACHER’S/EDUKATOR’S WORK.....	91
<i>Maria KOCÓR</i>	
DIAGNOZA I EWALUACJA W POLSKICH SZKOŁACH. CO JEST DOBRE, A NAD CZYM NALEŻY PRACOWAĆ?.....	104
<i>Tomasz M. ZIMNY</i>	
OCENIANIE W KONSTRUKTYWISTYCZNYM MODELU KSZTAŁCENIA	114
RESUMÉ.....	123
SUMMARY	124
REJSTRÍK.....	125

EDITORIAL

Široká odborná i laická veřejnost si již postupně navyká na přísun informací o výsledcích vzdělávání svých žáků či dětí získávaných především z mezinárodních výzkumů PISA, TIMSS, PIRLS, ICILS i PIAAC a dalších, ke kterým mají odborníci někdy oprávněné výhrady. Po mnoha peripetiích se postupně Česká republika blíží k akceptaci nového, často modifikovaného, modelu maturitní zkoušky, který byl v tomto roce již posedmé prakticky uplatněn. Přesto čeká realizátory společné části maturitní zkoušky nelehký úkol, zajistit povinou maturitní zkoušku z matematiky, která má mnoho zastánců s řadou argumentů pro i mnoho odpůrců s nemenší argumentační silou. Tataž veřejnost přijala s pochopením a očekáváním provedení dvou celoplošných generálních zkoušek ověřujících výsledky vzdělávání žáků základních škol v letech 2012 a 2013, aby pak s rozpaky vzala na vědomí jejich nahrazení výběrovým testováním k ověřování výsledků počínaje rokem 2014, které není z důvodu změn v kvantitě výběrových souborů i vybraných předmětů meziročně porovnatelné. Rozdílně přijímanou novinkou tohoto školního roku jsou povinné státní přijímací zkoušky uchazečů o studium na střední škole zakončené maturitní zkouškou, které se skládají z didaktických testů z českého jazyka a matematiky. Tato povinnost se týká všech uchazečů o čtyřleté, šestileté i osmileté studijní obory. Z tohoto stručného přehledu dosud krátké historie národního plošného testování vzdělávacích výsledků žáků základních a středních škol v České republice může vyplynout závěr, že více či méně úspěšné pokusy o získání objektivnějšího a úplnějšího přehledu o výsledcích školní edukace vyžadují hlubší analýzy a výzkumy potřebné k tomu, aby získaly pevné a stabilnější místo ve vzdělávacím systému. Aktivita decizní i exekutivní sféry vedoucí k tomuto cíli jsou uskutečňovány také v okolních zemích – v Polsku i na Slovensku, přičemž řada signálů svědčí o tom, že se tak děje s větší efektivitou, veřejnou i odbornou podporou a s větším přínosem pro žáky i pro rozvoj školského systému. Editoři monografie se tak mohou oprávněně domnívat, že publikované zkušenosti pracovníků polské i slovenské národní agentury pro zjišťování výsledků vzdělávání mohou být jen stěží nahraditelným zdrojem poučení pro budování současné a verifikované teorie plošného testování u nás, která má potenciál k přesvědčivému argumentování pro realizaci plošného testování, z něhož mohou být generovány výstupy v podobě vypočítané přidané hodnoty vzdělávání či stanovení úrovně pokroku žáka v učení nebo jeho silných a slabých stránek či rozdílů mezi jeho učebním potenciálem a jeho reálně dosahovanými učebními výsledky. Pokrok v oblasti edukačních aplikací informačních a komunikačních technologií vytváří nové podmínky pro tvorbu, administraci i vyhodnocování všech variant didaktických testů v podobě e-testování, které se v okolních státech pro své přednosti úspěšně rozvíjí.

Nepochybně mohou vedle hlavních cílů plošného testování vzniknout jeho další záměrné nebo vedlejší efekty v podobě e-autotestování, adaptivního testování nebo formativního testování.

Monografie, která se čtenáři dostává do rukou, obsahuje sedm kapitol. První kapitola věnovaná vývoji maturitní zkoušky v Polsku autorů Wojciecha Małeckého a Henryka Szaleniece popisuje genezi maturitní zkoušky mezi lety 2005 a 2015. Zejména pro čtenáře z jiného státu může být zajímavé srovnání pojetí maturitní zkoušky v Polsku a v jeho mateřské či jiné zemi, jistou paralelu lze bezesporu vysledovat i se současnou organizací

maturitní zkoušky v České republice. Také v Polsku byly po vyhodnocení prvních běhů Nové maturity provedeny změny, které reprezentovaly zásady ne nepodobné principům konstruktivismu (není podstatná pouze věda, oceňuje se také způsob, jakým se student k poznatkům dostává, akcentují se klíčové kompetence, zadání úloh se vyhýbají prvoplánovosti a jednoduchým odpovědím, mají komplexní charakter a iniciují tak holistický způsob hodnocení). V kapitole jsou dále prezentovány první efekty tohoto inovovaného pojetí maturitní zkoušky po dvou letech od jeho zavedení.

Maturitní zkouškou se mimo jiné zabývá také následující kapitola Lucie Ficové a Ivany Pichaničové, které zde popisují přechod od vyplňování testu způsobem papír-tužka k e-testování na Slovensku. Z celého textu, v němž se čtenář seznámí s vývojem procesu slovenské maturitní zkoušky, je patrná významná úloha NÚCEM (Národní ústav certifikovaných meraní vzdelávania). Ten zde má nezastupitelnou roli zejména při zajišťování vadility a reliability hodnotících nástrojů a garantování správnosti statistického vyhodnocení hromadných výsledků. V základních principech se zde v práci s přidanou hodnotou ve vzdělávání rovněž odráží konstruktivistický přístup. Autorky v textu popisují také proces přípravy testovacích nástrojů a pozornost věnují objektivitě administrace testů; neopomíjejí problémy, které při realizaci externích plošných testování nastávají. Čtenář si tak může utvořit poměrně jasnou a komplexní představu o systému e-Test, který slouží jako databáze úloh a testů – různé typy úloh jsou pak ilustrovány názornými příklady.

Přidané hodnotě ve vzdělávání a metodice určení relativního přírůstku znalostí žáků se ve třetí kapitole věnují i Josef Malach a Martin Malčík. V jejich praktičtěji orientovaném textu se čtenář seznámí s diagnostickým systémem TEST, jeho architekturou, včetně různých typů úloh a portálem, který tohoto testu využívá. Autoři prezentují souhrnné výsledky několika testovacích kampaní, které komentují a dávají do souvislosti, resp. je srovnávají s výsledky mezinárodního šetření OECD PISA. Také oni se vedle výhod testovacího systému nevyhýbají jeho záporům.

Henryk Szalaniec a Wojciech Małecki se v další kapitole zamýšlejí nad výsledky projektu, který v posledních osmi letech v Polsku pilotoval způsob elektronického hodnocení otevřených úloh. Autoři popisují průběh projektu, který doplňují příklady konkrétních hodnocení vypracovaných úloh. Metodika využívající názorů více hodnotitelů, kdy je zohledňována možnost efektu mírnosti nebo přísnosti, je v roce 2016 využívána pro gymnaziální zkoušky z matematiky pro celou populaci, čtenář se tedy může seznámit s průběhem vývoje podobného nástroje a utvořit si vlastní názor na možnosti uplatnění takového postupu v pro něj aktuálním edukačním prostředí.

Ewa Wysocka pak v obecněji zaměřené páté kapitole na bázi paradigmatu pozitivismu srovnává pozitivní a negativní diagnostiku, kdy na celý proces diagnostiky nahlíží v souladu s konstruktivistickým pojetím z kognitivní perspektivy. Ačkoli přímo nezmiňuje posun mezi diagnostikou a evaluací, v rámci komplexnosti diagnostického procesu zdůrazňuje úlohu post-diagnostického plánování spolu s komplementaritou diagnostiky a intervence. V této kapitole se čtenář seznámí také s různými diagnostickými nástroji, které jsou užívány v polském prostředí k diagnostice domněnek o vlastní osobnosti a jejím potenciálu.

Analýzu silných a slabých stránek evaluace v polských školách najde čtenář v kapitole Marie Kocór. Autorka zde provedla část SWOT analýzy, kde v tabulce tvořící jádro textu shrnuje činitele podporující či ohrožující diagnostické a evaluační procesy v tomto specifickém edukačním kontextu. Pozornost věnuje nezbytnému vymezení rozdílů mezi diagnostikou a evaluací a také úloze školy, resp. jejího ředitele, v těchto procesech. Pro pojetí této kapitoly je charakteristický citát, který autorka uvádí: „Pomáhám ti tak, aby sis dokázal pomoci sám.“

Problematika je tedy nahlížena spíše z pohledu individuálního rozvoje žáka jako diagnostikovaného či evaluovaného jedince.

Můžeme konstatovat, že pro praktickou výuku na současných školách se velmi široce uplatňovaným paradigmatickým modelem stal konstruktivismus, což se, dá se říci, projevuje ve všech kapitolách této knihy. V poslední kapitole, v níž se čtenář seznámí s teoretickým ukotvením publikace, se proto Tomasz Zimny zabývá tím, co je typické pro evaluaci, která do konstruktivistického modelu zapadá a má zde své nezastupitelné místo. Přestože se analyzovaná problematika vztahuje k edukační realitě polského vzdělávacího systému, jsou autorovy závěry týkající se zejména závěrečných zkoušek užitečné i pro vzdělávací systémy ostatních postkomunistických zemí. S ohledem na probíhající civilizační změny klade autor důraz na žáka jako učící se subjekt, v němž se spojují dva elementy– individualita jeho osobnosti a chápání této osobnosti jako činitele i výsledku působení společnosti. Autor v této souvislosti porovnává myšlenky zakladatele konstruktivismu Jeana Piageta s myšlenkami Lva Semjonoviče Vygotského a v návaznosti na toto srovnání prezentuje možnosti, jak mapovat pokrok žáka v učení.

Monotematicky zaměřená monografie obsahuje odborné texty vycházející z bohatých zkušeností odborníků specializujících se na provádění široce založeného testování. Může být inspirací a poučením pro tvorbu didaktických testů a jejich administraci a vyhodnocování využívající příslušných prostředků informačních a komunikačních technologií, které budou zefektivňovat celou proceduru testování a budou eliminovat chyby vzniklé lidským faktorem.

Monografie je určena odborné pedagogické komunitě, především z řad akademických pracovníků, a zvláště pak těm, kteří se zabývají otázkami plošného testování.

Josef Malach, Martin Malčík, Martina Rozsypalová, editoři

EWOLUCJA EGZAMINU MATURALNEGO W POLSCE ZAŁOŻENIA I PIERWSZE EFEKTY

Wojciech Malecki, Henryk Szaleniec

Abstrakt: Kapitoła popisuje genezi maturitní zkoušky v Polsku mezi lety 2005 a 2015. Také v Polsku byly po vyhodnocení prvních běhů Nové maturity provedeny v jejím obsahu i realizaci změny, které vycházely z jasně formulované ideje, že má obsáhnout nejen vědomosti, ale také metody osvojování vědy a klíčové kompetence. Přibližuje způsoby hodnocení výkonů žáků, monitorování celého procesu a zásady promyšlených a odůvodněných změn v celém konceptu společné státní části maturitní zkoušky

Klíčová slova: státní maturita, koncept maturity, ideje maturity, maturitní zadání, problémy realizace maturity, modifikace maturity

Abstract: The chapter describes the genesis of the graduation exam in Poland between 2005 and 2015. Also in Poland, after evaluating the first runs New Matura exam performed in its content and implementation of the changes, which were based on clearly formulated the idea that he has to encompass not only knowledge but also methods of learning and science core competencies. Describes the methods of assessment of student performance monitoring throughout the process and principles elaborated and justified changes in the whole concept of a common national school-leaving examinations

Key words: Matura state exam, Matura state exam concept, idea of Matura state exam, examination tasks, problems of implementation Matura exam, Matura exam modification

MATURA OD 2005 DO 2015

Zewnętrzne egzaminy w Polsce stanowiły składową pierwszej dużej reformy edukacji w Rzeczypospolitej Polskiej wprowadzonej w życie w 1999 r. Zewnętrzny egzamin maturalny po raz pierwszy odbył się w 2005 roku. Wprowadzenie egzaminu możliwe było dzięki 10 letnim przygotowaniom prowadzonym w ramach programu „Nowa matura” finansowanego przez Ministerstwo Edukacji Narodowej, a od roku 1997 współfinansowanego przez program SMART w ramach PHARE¹.

Egzamin został wdrożony z powodzeniem i znalazł uznanie uczelni. Każdego niemal roku wprowadzano do egzaminu niewielkie zmiany. Najpoważniejszą z nich było (w 2010 roku) uznanie egzaminu z matematyki jako egzaminu obowiązkowego. Oczywiście cały czas w życiu publicznym, najczęściej w mediach, prowadzone były dyskusje o jakości egzaminu, jego strukturze i wartości dydaktycznej. Formułowane były zastrzeżenia dotyczące form egzaminów (także typów zadań), poziomu przygotowania maturzystów do studiów oraz

¹ **Phare** (ang. *Poland and Hungary: Assistance for Restructuring their Economies*) – program Komisji Europejskiej powstały w roku 1989 w celu udzielania materialnej pomocy państwom kandydującym do Wspólnoty Europejskiej.

jakości oceniania. Oprócz głosów krytycznych pojawiały się też, wcale nierzadko, pozytywne opinie.

Od roku 2009 wprowadzono drugą reformę systemu oświaty. Była to reforma znacznie skromniejsza niż poprzednia, w zasadzie korygowała zasadnicze rozstrzygnięcia pierwszej reformy. Sprowadzała się do zmiany podstawy programowej kształcenia ogólnego oraz istotnej zmiany kształcenia zawodowego.

W polskim systemie oświaty podstawa programowa wyznacza cele kształcenia, ustala wspólne dla wszystkich uczących się obszary wiedzy i umiejętności, zapewnia spójność procesu kształcenia i jednocześnie daje możliwość różnorodnych ścieżek kształcenia poprzez dopuszczenie do użytku wielu programów nauczania i wielu podręczników. Podstawa programowa promuje kształcenie dla przyszłości. Stawia cele, których osiągnięcie nie tylko otwiera drogę do dalszego kształcenia, ale umożliwia aktywne uczestniczenie w społeczeństwie obywatelskim.

Egzaminy zewnętrzne (sprawdzian po klasie szóstej, egzamin gimnazjalny, matura) operacjonalizują podstawę programową i mają na celu – między innymi – sprawdzenie stopnia osiągnięcia przez uczniów celów zapisanych w podstawie programowej.

Oczywistym więc było, że zmiana podstawy programowej wymusiła modyfikację egzaminów zewnętrznych, w tym egzaminu maturalnego. Przy okazji zmieniono też – na podstawie wniosków z przebiegu egzaminów we wcześniejszych latach – strukturę i organizację egzaminu. Po raz pierwszy zmodyfikowana matura przeprowadzona została w roku 2015.

1 OPIS EGZAMINU MATURALNEGO OD ROKU 2015

Prawną podstawę modyfikacji egzaminu maturalnego zapisano w ustawie przyjętej przez Sejm oraz w Rozporządzeniu Ministra Edukacji Narodowej. Szczegółowy opis egzaminu maturalnego wraz przykładowymi zadaniami znajduje się w informatorach maturalnych (ogólnym i przedmiotowych) opublikowanych przez Centralną Komisję Egzaminacyjną w roku 2013². Przypomnijmy podstawowe fakty.

Egzamin maturalny w Polsce składa się z dwóch części.

1. Obowiązkowej, którą tworzą trzy egzaminy pisemne na poziomie podstawowym:
 - język polski,
 - matematyka,
 - język obcy nowożytny (jeden z następujących: angielski, niemiecki, rosyjski, francuski, hiszpański, włoski), oraz dwa egzaminy ustne bez określonego poziomu: język polski, język obcy (taki sam jak na egzaminie pisemnym).
2. Dodatkowej, w której należy wybrać przynajmniej jeden (a najwięcej sześć) spośród następujących przedmiotów zdawanych na poziomie rozszerzonym: język polski, matematyka, język obcy nowożytny, język mniejszości narodowej (etnicznej), język regionalny, język łaciński i kultura antyczna, historia muzyki, historia sztuki, filozofia, informatyka, biologia, chemia, fizyka, geografia, historia, wiedza o społeczeństwie.

² <https://www.cke.edu.pl/egzamin-maturalny/egzamin-w-nowej-formule/informatory/>

Warunkiem otrzymania świadectwa maturalnego jest zdanie pięciu egzaminów w części obowiązkowej (jako zdany uznaje się egzamin, w którym zdający uzyska przynajmniej 30% możliwych do zdobycia punktów) oraz przystąpienie do egzaminu z jednego przedmiotu dodatkowego.

Najważniejsze formalne zmiany wprowadzone w roku 2015 to:

- odstąpienie od przygotowania na egzaminie ustnym z języka polskiego prezentacji; zdający odpowiada na pytania zawarte w wylosowanym zestawie,
- ustalenie egzaminów z przedmiotów dodatkowych jedynie na poziomie rozszerzonym,
- wprowadzenie obowiązku przystąpienia do jednego przedmiotu dodatkowego (nie ma tu progu zaliczenia),
- zmiana struktury arkuszy egzaminacyjnych z przedmiotów dodatkowych.

O wizerunku matury zdecydowała modyfikacja celów zapisanych w podstawie programowej kształcenia ogólnego obowiązującej w szkołach ponadgimnazjalnych od roku 2012. Na nowo sformułowane cele stały się punktem wyjścia do sformułowania idei zmodyfikowanej matury.

2 IDEE ZMODYFIKOWANEGO EGZAMINU MATURALNEGO

Podstawa programowa w części odnoszącej się do IV etapu kształcenia (szkoły ponadgimnazjalne) wyraźnie formułuje cele kształcenia i oczekiwane umiejętności uczniów, sugeruje też metodykę pracy nad nimi. Tak cele jak i umiejętności zawierają składowe wspólne dla wszystkich przedmiotów i specyficzne dla każdego z nich. I właśnie ten wspólny obszar celów i umiejętności pozwala wyłonić idee zmodyfikowanej matury. Idee te w skrócie można przedstawić następująco:

- sprawdzamy nie tylko wiedzę
- doceniamy metody zdobywania wiedzy (metodologię nauki)
- akcentujemy kompetencje kluczowe i integrację wiedzy
- preferujemy zadania z „wyposażeniem”
- sprawdzamy więcej umiejętności złożonych
- rozwiązania oceniamy holistycznie

Sformułowane wyżej idee stały się wytycznymi do budowania arkuszy egzaminacyjnych. Zadania obejmowały wiedzę i umiejętności zapisane w podstawie programowej kształcenia ogólnego, a idee podpowiadały dobór typów zadań i sposobów sprawdzania wiedzy i umiejętności. Oczywiście zadbano tu o ciągłość egzaminu i jedynie część zadań maturalnych była nowego typu. W kolejnych latach zakres modyfikacji egzaminu maturalnego będzie stopniowo poszerzany.

Co oznaczają wymienione idee, jak były one sprawdzane na egzaminie maturalnym, jakie są osiągnięcia maturzystów w pierwszym i drugim roku obowiązywania zmodyfikowanych egzaminów? Odpowiemy na te pytania w kolejnych akapitach. Zaczynamy od nadania znaczenia wymienionym wyżej ideom.

Sprawdzamy nie tylko wiedzę

Oznacza to, że na egzaminie poprawna i głęboka wiedza jest niezbędna, ale niewystarczająca. Podstawa programowa wymaga, aby wiedza uczniów była użyteczna. Wiedza ma służyć poznawaniu czyli: zdobywaniu nowej wiedzy bezpośrednio (poprzez zapoznawanie się z tekstami kultury, analizowanie źródeł historycznych, obserwowanie i eksperymentowanie w obszarze przyrody, odkrywanie reguł w matematyce) lub pośrednio (czytanie opracowań i analiz literackich, korzystanie z podręczników, opracowań i monografii, źródeł internetowych, przetworzonych danych w formie tabel, diagramów, wykresów).

Doceniamy metody zdobywania wiedzy (metodologię nauki)

Kształcenie w obszarze tzw. przedmiotów akademickich bazuje na powszechnie uznawanej wiedzy naukowej i metodologii nauk. Wytwory nauki (twierdzenia, prawa, teorie, generalizacje) są ściśle związane z metodami dochodzenia do nich. Możemy powiedzieć, że te pojęcia, wytwór nauki i metoda nauki, są nierozzerwalne. Jednak w szkole przeważa przekaz osiągnięć nauki, prowadzona jest wręcz narracja o prawach, generalizacjach, teoriach, rzadziej o paradygmatach. Metody poznawania nie znajdują należnego miejsca w procesie poznawania kierowanym przez nauczycieli. Tymczasem podstawa programowa bogata jest w zagadnienia metodologiczne. Jednym z celów kształcenia mają być złożone umiejętności opisu (i interpretacji), wyjaśniania, przewidywania (rekonstruowania) oraz bardziej szczegółowe umiejętności metodologiczne stanowiące dla uczniów „narzędzia myślenia”. W zmodyfikowanej maturze, praktycznie w każdym przedmiocie, akcentuje się sprawdzanie przedstawionych tu umiejętności.

Akcentujemy kompetencje kluczowe i integrację wiedzy

Kompetencje kluczowe obecne są w polskiej podstawie programowej już od pierwszej reformy programowej. Nadal jednak nie znajdują należnego im miejsca w kształceniu i ocenianiu. Wiedza i umiejętności często zamykane są w ramach pojedynczych przedmiotów nauczania. Podobnie jest ze stosowaniem wiedzy i umiejętności do złożonych zjawisk i procesów – tak w obszarze przedmiotów humanistycznych jak i matematyczno – przyrodniczych. W podstawie programowej znajdują się umiejętności wspólne dla wielu dyscyplin naukowych (i odpowiadających im przedmiotów nauczania). Do wspomnianych umiejętności należą, między innymi, takie jak: tworzenie i przetwarzanie informacji, dostrzeganie (przedstawianie) zależności – np. związków przyczynowo- skutkowych, projektowanie rozwiązań, krytyczne ocenianie (źródła, rozumowania, wnioski). Do zadań maturalnych w większym niż dotąd stopniu wprowadzono problemy sprawdzające kompetencje kluczowe i umiejętności integrowania wiedzy z różnych obszarów.

Preferujemy zadania z „wyposażeniem”

Koncepcja zmodyfikowanej matury zakłada odejście od prostych pytań i prostych na nie odpowiedzi. Zadania maturalne nawiązują do tekstów kultury, do rzeczywistości przyrodniczej lub społecznej. Trzono zadań są bardziej rozbudowane. Znajdują się w nich: różnego rodzaju źródła, teksty kultury, teksty popularnonaukowe, opisy zjawisk (eksperymentów), zestawienia danych, wykresy i diagramy, mapy, zdjęcia. Dodajmy, że do egzaminu z matematyki przygotowane została broszura „Wybrane wzory matematyczne”³, a do egzaminów z przedmiotów przyrodniczych - broszura „Wybrane wzory i stałe fizykochemiczne na egzamin maturalny z biologii, chemii i fizyki”⁴.

³ https://www.cke.edu.pl/images/_EGZAMIN_MATURALNY_OD_2015/Informatory/2015/MATURA_2015_Wybrane_wzory_matematyczne.pdf

⁴ https://www.cke.edu.pl/images/_EGZAMIN_MATURALNY_OD_2015/Informatory/2015/MATURA_2015_Wybrane_wzory_i_sta%C5%82e_fizykochemiczne.pdf

Sprawdzamy więcej umiejętności złożonych

Umiejętności te charakterystyczne są dla wyższych poziomów poznawania. Wymagają wiązania i integrowania umiejętności prostych. Pozwalają na korzystanie z wiedzy podanej w zadaniu jak i szerokiej wiedzy kontekstowej. Do ich osiągnięcia konieczne są kompetencje pozaprzedmiotowe.

Rozwiązania oceniamy holistycznie

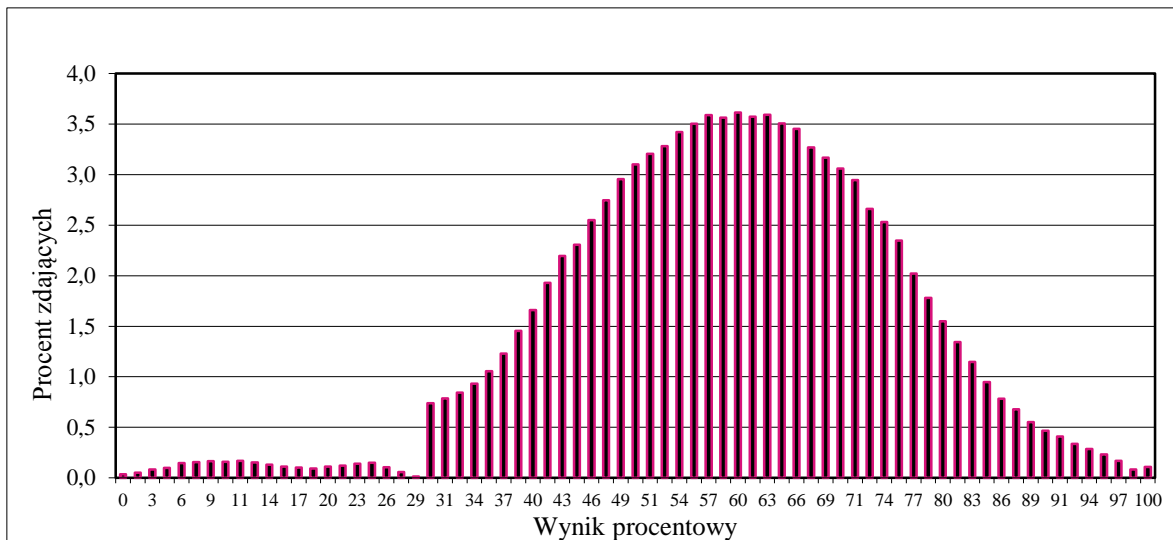
Ocenianie holistyczne zastępuje tak zwane ocenianie analityczne. Tak jedno jak i drugie podejście do oceniania wymaga opracowania stosowania kryteriów oceniania. Przy podejściu holistycznym budujemy kryteria bardziej uniwersalne, staramy się, by spełnienie kolejnych kryteriów oznaczało zbliżenie się do rozwiązania problemu. Schemat oceniania w podejściu holistycznym jest starannie opisany i zawiera przykłady wielu metod rozwiązania, ze wskazaniem punktów krytycznych (pełne rozwianie, pokonanie zasadniczej trudności, brak sukcesu). Wprowadzenie podejścia holistycznego do oceniania zadań maturalnych pozwala uwzględniać stopień opanowania nie tyle pojedynczych (prostych) umiejętności, ale stopień osiągnięcia celu zadania. Wdrożenie oceniania holistycznego poprzedziło staranne przygotowanie egzaminatorów.

3 PIERWSZE EFEKTY ZAREJESTROWANE PO DWÓCH LATACH

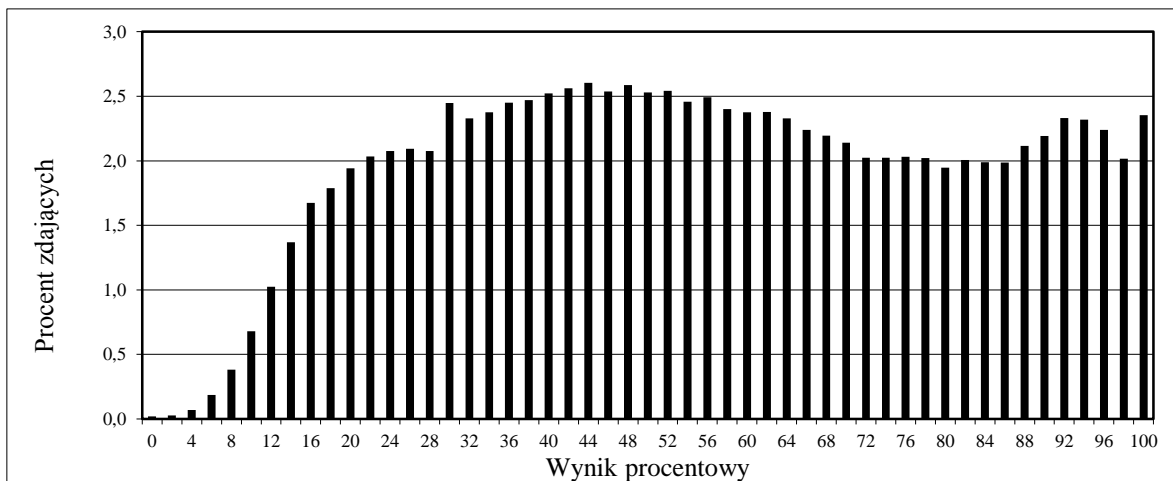
Jak już wspomniano, po raz pierwszy zmodyfikowany egzamin maturalny został przeprowadzony w maju 2015. Przystąpili do niego jedynie absolwenci liceów ogólnokształcących. W maju 2016 nowy egzamin maturalny zdawali także absolwenci techników, czyli cała populacja maturzystów. Poniżej przedstawimy wyniki egzaminu w roku 2016. Ograniczymy się do podania i skomentowania podstawowych parametrów zdawalności egzaminu oraz do analizy stopnia osiągania przez zdających wybranych umiejętności, realizujących omówione wyżej idee zmodyfikowanej matury. Pełne wyniki egzaminu znaleźć można na stronie internetowej Centralnej Komisji Egzaminacyjnej i każdej z ośmiu okręgowych komisji egzaminacyjnych. Tamże opublikowane zostały sprawozdania z każdego egzaminu maturalnego w każdym przedmiocie.

Do egzaminu w roku 2016 przystąpiło 259386 tegorocznych absolwentów (170834 absolwentów liceum ogólnokształcącego (LO), 88564 absolwentów techników (T)). Spośród nich świadectwo maturalne otrzymało (czyli zdało trzy obowiązkowe egzaminy na poziomie podstawowym, dwa egzaminy ustne, przystąpiło do jednego egzaminu z przedmiotu dodatkowego na poziomie rozszerzonym 85% absolwentów (89% w LO, 76% w T). Taki wskaźnik zdawalności niewiele różni się od zdawalności w latach ubiegłych. O wskaźniku zdawalności zdecydowały przede wszystkim rezultaty egzaminu z matematyki – 86% (90% absolwenci LO, 80% - T).

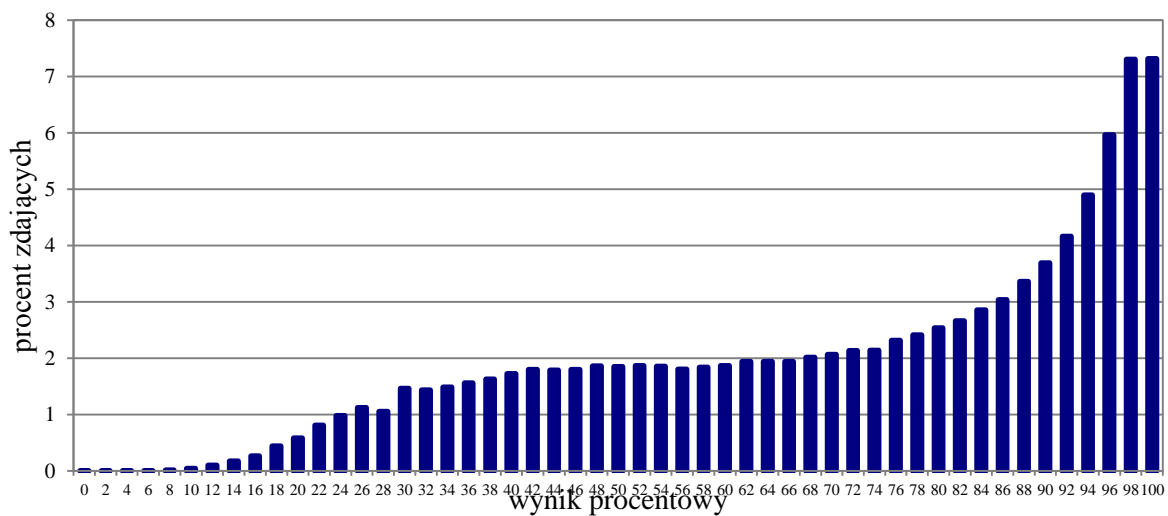
Wskaźniki zdawalności pokazują, że egzamin maturalny w części obowiązkowej nie jest trudnym egzaminem. Świadectwo maturalne uzyskuje zdecydowana większość przystępujących do matury. Egzamin ten jednak silnie różnicuje zdających. Przekonać się o tym można porównując rozkłady wyników dla trzech kluczowych egzaminów pisemnych na poziomie podstawowym:



Rys. 1 Rozkład wyników (poziom podstawowy) z języka polskiego



Rys. 2 Rozkład wyników (poziom podstawowy) z matematyki



Rys. 3 Rozkład wyników (poziom podstawowy) z języka angielskiego

Dla języka polskiego rozkład jest niemal symetryczny między minimalnym rezultatem oznaczającym zdanie egzaminu (30%) a wynikiem maksymalnym (100%). Widać tu efekt egzaminatora na progu zdawalności i naturalne zróżnicowanie zdających. Inaczej jest w przypadku matematyki, gdzie rozkład jest praktycznie płaski. Oznacza to ogromne różnice przygotowania do egzaminu. Mamy niemalą grupę (ponad $\frac{1}{4}$) populacji tworzoną przez osiągających wynik powyżej 80%, dużą grupę spełniających wymagania oraz około $\frac{1}{6}$ populacji nie radzących sobie z matematyką. Sytuacja (z niewielkimi zmianami) powtarza się od kilku lat. Jest to wyzwanie dla szkoły, dla nauczycieli. Z kolei rozkład wyników języka angielskiego uwidacznia wyraźny sukces $\frac{1}{3}$ zdających, mających wynik zbliżony do maksymalnego i niewielką grupę, dla których język angielski nie jest jeszcze językiem komunikowania się.

Jak już wspomniano, egzaminy na poziomie podstawowym niewiele zmieniły się w roku 2015. Jednak efekty zmian są zauważalne i zachęcają do doskonalenia egzaminów zgodnie z przyjętą koncepcją.

Znacznie bardziej złożona jest ocena efektu wprowadzenia zmiany na poziomie rozszerzonym. Decyduje o tym kilka czynników. Najbardziej istotne jest to, że modyfikacja dotyczy koncepcji całego egzaminu w każdym przedmiocie, ale realizowana jest poprzez wymianę jedynie części zadań. Zachowano tu zasadę ciągłości. Uczniowie przygotowywali się do matury biorąc pod uwagę egzaminy z lat poprzednich. Opublikowanie informatorów i przykładowych zadań do nowego egzaminu, przeprowadzanie próbnych matur, szkolenia nauczycieli nie są, niestety, tak skuteczne, jak tego byśmy oczekiwali. Przyzwyczajenie i zachowawczość wpływają na nietrafne przygotowania do nowej formuły. Dopiero pierwsze egzaminy przemawiają do nauczycieli i do zdających. Dlatego efekty nowych egzaminów w pierwszym i drugim roku ich funkcjonowania nie są jeszcze miarodajne.

Przedstawiamy poniżej pierwsze efekty zmiany jako zapowiedź i ilustrację problemów, wobec których stawali zdający.

Efekty przedstawimy przez pryzmat osiągnięć związanych z wiedzą i umiejętnościami zapowiedzianymi w nowej podstawie programowej, przede wszystkim poprzez odniesienie się do omówionych wyżej idei modyfikacji egzaminu na poziomie podstawowym i rozszerzonym w przedmiotach dodatkowych.

Spośród kilkuset zadań wybraliśmy jedynie kilka. Ich krótkie omówienie będzie ilustracją percepcji zmian przez zdających.

W tabeli poniżej przedstawiamy wybrane do analizy zadania z kilku przedmiotów, idee realizowane w wybranych zadaniach, średni wynik za cały arkusz oraz za analizowane zadanie.

egzamin	zadanie	idea						średni wynik [%]	
		1	2	3	4	5	6	zadanie	arkusz
język polski – poziom podstawowy	1.7	x		x	x		x	33	61
matematyka - poziom rozszerzony	8		x				x	20	30
biologia	1		x	x		x		27	34
chemia	32	x				x		16	36
fizyka	15	x	x		x	x		32	37
historia	12	x	x		x			70	40
Wiedza o społeczeństwie	23.1			x				16	26
	23.2	x		x		x		4	

Tab. 1 Idee modyfikacji egzaminów an przykładzie wybranych zadań (1 - sprawdzamy nie tylko wiedzę, 2 - doceniamy metody zdobywania wiedzy (metodologię nauki), 3 - akcentujemy kompetencje kluczowe i integrację wiedzy, 4 - preferujemy zadania z „wyposażeniem”, 5 - sprawdzamy więcej umiejętności złożonych, 6 - rozwiązania oceniamy holistycznie)

Język polski – poziom podstawowy – zadanie 1.7

Przeczytaj uważnie teksty, a następnie wykonaj zadania umieszczone pod nimi. Odpowiadaj **tylko na podstawie tekstów i tylko własnymi słowami** – chyba że w zadaniu polecono inaczej. Udzielaj tyłu odpowiedzi, o ile Cię poproszono.

Zadanie 1.

[O *Lalce*]

Na długiej liście przygnębiających metafor formułowanych przez bohaterów *Lalki* znajdziemy ludzi-marionetki, ludzi-mrówki, nawet – pierwotniaki. Rzecki dodaje do niej

Zadanie 1.7. (0–3)

Napisz streszczenie tekstu liczące 40–60 słów.

Zadanie 1.7. (0–3)

Wymagania ogólne	Wymagania szczegółowe
I. Odbiór wypowiedzi i wykorzystanie zawartych w nich informacji.	Zdający: 1.5) wyróżnia argumenty, kluczowe pojęcia i twierdzenia w tekście argumentacyjnym, dokonuje jego logicznego streszczenia.
III. Tworzenie wypowiedzi.	Zdający: 1.7) wykonuje różne działania na tekście cudzym ([...] streszcza [...]).

Schemat punktowania

- 3 p. – Ze streszczenia wynika, jaki jest temat tekstu i co na ten temat powiedziano w tekście; adekwatny poziom uogólnienia, streszczenie logicznie spójne, właściwa liczba słów.
- 2 p. – Ze streszczenia wynika, jaki jest temat tekstu i co na ten temat powiedziano w tekście ALE zaburzenia dotyczące sposobu uogólnienia LUB logicznej spójności streszczenia.
- 1 p. – Ze streszczenia wynika, jaki jest temat tekstu i co na ten temat powiedziano w tekście; ALE zaburzenia dotyczące poziomu uogólnienia ORAZ logicznej spójności streszczenia.
- 0 p. – odpowiedź błędna LUB brak odpowiedzi.

Zadanie wymagało streszczenia tekstu krytyczno – literackiego liczącego ponad 600 słów. Zdający musieli wykazać się umiejętnością identyfikowania tezy tekstu, argumentowania, uogólniania. Zadanie było trudniejsze niż większość pozostałych. Wielu zdających miało problemy ze zidentyfikowaniem tezy tekstu, zapisaniem jego tematu i przedstawieniem argumentów autora. Rezultat zadania potwierdza hipotezę o niezbyt wysokim poziomie umiejętności czytania ze zrozumieniem i krytycznej analizy czytanego tekstu. Jest to jedna z kompetencji kluczowych. Analiza zadania podpowiada rekomendowanie nauczycielom pracy nad pogłębieniem wskazanych umiejętności. Zadanie oceniane było w podejściu holistycznym. Kolejne punkty przyznawane były za całość odpowiedzi i premiowały spełnianie kluczowych cech streszczenia.

Matematyka poziom rozszerzony – zadanie 8

Zadanie 8. (0–3)

Wykaż, że dla dowolnych dodatnich liczb rzeczywistych x i y takich, że $x^2 + y^2 = 2$, prawdziwa jest nierówność $x + y \leq 2$.

Zadanie wymagało przeprowadzenia dowodu algebraicznego. Absolwenci LO i T dysponują odpowiednimi do rozwiązania problemu narzędziami matematycznymi. Jednak w sytuacji innego niż stereotypowe polecenia mają kłopoty z uzasadnianiem tezy. To typowa trudność metodologiczna. Zadanie sprawiło wiele kłopotów, było jednym z trudniejszych w arkuszu. Zadanie było trudne do oceny przez egzaminatorów. W podejściu holistycznym wyróżniono osiem typów rozwiązania i do każdego adekwatny schemat punktowania uwzględniający postęp zdających w rozwiązaniu zadania.

Biologia zadanie 1.

Zadanie 1. (0–3)

Poniżej wymieniono niektóre właściwości fizykochemiczne wody.

- A. duże napięcie powierzchniowe;
- B. duże ciepło parowania;
- C. maksymalna gęstość w temperaturze $4\text{ }^{\circ}\text{C}$.

Uzupełnij zdania (1.–3.) tak, aby były prawdziwe – wpisz na początku zdania oznaczenie literowe wybranej właściwości wody (A–C), a następnie dokończ zdanie: wyjaśnij, w jaki sposób dana właściwość warunkuje funkcjonowanie wymienionych organizmów.

1. umożliwia poruszanie się niektórych gatunków owadów po powierzchni wody, ponieważ

.....

2. umożliwia przetrwanie ryb słodkowodnych podczas zimy przy dnie zamarzających zbiorników, ponieważ

.....

3. umożliwia pozbywanie się nadmiaru ciepła z organizmu człowieka podczas pocenia się, ponieważ

Zadanie jest interesujące gdyż wiąże wiedzę i umiejętności z biologii i fizyki. Ta integracja stawiała zdających w nietypowej sytuacji. Właściwości wody, ważne w procesach życiowych organizmów, wyjaśniać musieli w języku fizyki. Drugą trudność stanowiła sama procedura wyjaśniania. Złożoność zadania skutkowałą niskim rezultatem.

Chemia zadanie 32

Zadanie 32. (0–1)

Kwas 2-hydroksypropanowy w reakcji ze związkiem X tworzy ester o wzorze sumarycznym $C_5H_8O_4$. Orbitalom walencyjnym każdego z atomów węgla budujących cząsteczkę związku X przypisuje się inny typ hybrydyzacji. Ponadto wiadomo, że w cząsteczce związku X występuje tylko jedna grupa funkcyjna.

Ustal wzór związku X, którego użyto do estryfikacji kwasu 2-hydroksypropanowego, i napisz równanie reakcji otrzymywania opisanego estru. Zastosuj wzory półstrukturalne (grupowe) związków organicznych.

Kluczem do rozwiązania zadania była wnikliwa analiza informacji wprowadzającej. Jedynie poprawna analiza dawała szansę ustalenia wzoru związku X. Warunkiem sukcesu było powiązanie wszystkich informacji, ich kontekstowego zrozumienia i poprawne korzystanie z kilku prostych umiejętności niezbędnych do rozwiązania jednego problemu.

Fizyka zadanie 15

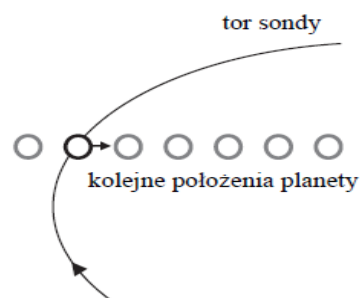
Zadanie 15.

Nadanie obiektowi wystarczająco dużej prędkości początkowej (większej od pewnej granicznej wartości zwanej prędkością ucieczki) pozwala mu uwolnić się od Ziemi i odlecieć na dowolnie dużą odległość. Przyciąganie Ziemi nigdy nie zdoła zawrócić takiego obiektu. Mimo że jest to dosyć duża prędkość, to ludzkość zbudowała pojazdy zdolne ją osiągać. Dużo trudniejsza od ucieczki z Ziemi jest ucieczka od Słońca, którego masa jest prawie 336 tys. razy większa od masy Ziemi. Żaden z pojazdów kosmicznych zbudowanych do tej pory przez człowieka nie był w stanie za pomocą własnych silników nadać sobie energii wystarczającej do ucieczki od Słońca. Wiadomo jednak, że kilka wysłanych z Ziemi sond opuściło Układ Słoneczny – przyciąganie Słońca nigdy ich już nie zawróci. Rozpędzenie sondy uzyskano dzięki sprytnemu zaplanowaniu trajektorii lotu, tak aby w odpowiedni sposób przeleciała w pobliżu planety Układu Słonecznego. [...]

Mechanizm ten nosi nazwę procy grawitacyjnej lub asysty grawitacyjnej i był wykorzystywany do rozpędzania wielu sond, a niektóre z nich wykorzystywały go kilkakrotnie. Rozważmy sytuację, w której sonda obiega Słońce po orbicie w kierunku przeciwnym do kierunku obiegu planety. Wyobraźmy sobie, że sonda zbliża się do planety. W układzie odniesienia związanym z planetą trajektoria sondy ma w przybliżeniu kształt taki jak na rysunku obok. [...]

Składowe prędkości sondy prostopadłe do kierunku ruchu planety się nie zmieniają, natomiast ulegną zmianie składowe równoległe – w fazie zbliżania od wartości składowej równoległej odejmie się prędkość planety, a w fazie oddalania prędkość planety się do niej doda. W konsekwencji, jeżeli porównamy składową równoległą prędkości sondy przed zbliżeniem i po nim, to będą się one różniły o podwojoną wartość prędkości planety. Uzyskamy zatem to, co chcieliśmy – sonda po zbliżeniu będzie się poruszać szybciej!

Na podstawie: Szymon Charzyński, *Proca grawitacyjna*, „Delta”, lipiec 2014.



Zadanie 15.1. (0–3)

Oszacuj wartość prędkości, jaką należy nadać obiektowi znajdującemu się w odległości jednej jednostki astronomicznej (1 j.a. \approx 150 mln km) od Słońca, aby mógł on trwale opuścić Układ Słoneczny.

Zadanie 15.2. (0–1)

Oblicz, jaką maksymalną prędkość względem Słońca mogłaby osiągnąć sonda początkowo poruszająca się względem niego z prędkością 10 km/s, na skutek zastosowania asysty grawitacyjnej Jowisza, krążącego wokół Słońca z prędkością 13 km/s.

Zadanie 15.3. (0–1)

Zaznacz poprawne dokończenie zdania.

Przy zastosowaniu asysty grawitacyjnej energia kinetyczna sondy wzrasta kosztem

- A. energii potencjalnej sondy. B. energii kinetycznej Słońca.
C. energii kinetycznej planety. D. energii potencjalnej planety.

Zadanie 15.4. (0–1)

Oceń prawdziwość poniższych stwierdzeń. Zaznacz P, jeśli zdanie jest prawdziwe, albo F – jeśli jest fałszywe.

Asysta grawitacyjna opisana w tekście

1.	może być wykorzystana do zmniejszenia prędkości sondy.	P	F
2.	nie powoduje zmiany kierunku ruchu sondy.	P	F
3.	wpływa na ruch sondy w podobny sposób, jak zderzenie sprężyste z poruszającym się ciałem o dużej masie.	P	F

Tego typu zadania nie występowały wcześniej na egzaminie maturalnym z fizyki. Podstawą zadania był tekst popularnonaukowy. Uważne przeczytanie tekstu, długiego jak na szkolne i egzaminacyjne warunki, było pierwszym wyzwaniem dla zdających. Drugim było odniesienie do zjawisk nie omawianych w szkole. Do tego należało samodzielnie wyprowadzać wnioski. Wszystko to spowodowało, że zadanie okazało się trudne. Tym niemniej uważamy, że tego typu zadania należy szerzej stosować, stawiają one bowiem zdających w sytuacjach, jakie zdarzać się będą poza szkołą i rozwijają umiejętności ważne w dalszym uczeniu się .

Historia zadanie 12

Zadanie 12. (0–1)

Źródło 1. Fragment źródła z epoki

Monarcha nie potrzebuje przed nikim zdawać sprawy ze swych rozporządzeń [...]. Bez tej władzy absolutnej nie może działać nic dobrego ani niszczyć złego. [...] Odjąć władzę monarsze znaczy podzielić państwo, tj. zakłócić spokój publiczny, stworzyć dwóch panów w przeciwieństwie do słów Pisma Świętego: „Nie można służyć dwóm panom”. Król jest na mocy swej władzy ojcem narodu [...]. Dlatego najlepiej jest oddać całą władzę państwa temu, kto ma największy interes w utrzymaniu i wielkości samego państwa.

Wiek [...] w źródłach. Wybór tekstów źródłowych z propozycjami metodycznymi dla nauczycieli historii i studentów, oprac. M. Sobańska-Bondaruk, S.B. Lenard, Warszawa 1999, s. 318–319.

Źródło 2. Fragment źródła z epoki

Wszystko byłoby stracone, gdyby jeden i ten sam człowiek lub jedno i to samo ciało magnatów albo szlachty, albo ludu sprawowało owe trzy władze: tworzenie praw, wykonywanie publicznych postanowień oraz sądzenie zbrodni lub sporów między obywatelami. [...] Ponieważ w wolnym państwie każdy człowiek [...] winien się rządzić sam, trzeba by, aby cały lud [...] posiadał władzę prawodawczą.

Historia [...]. Wybór tekstów źródłowych dla szkół średnich, oprac. J. Eisler, M. Sobańska-Bondaruk, Warszawa 1995, s. 12.

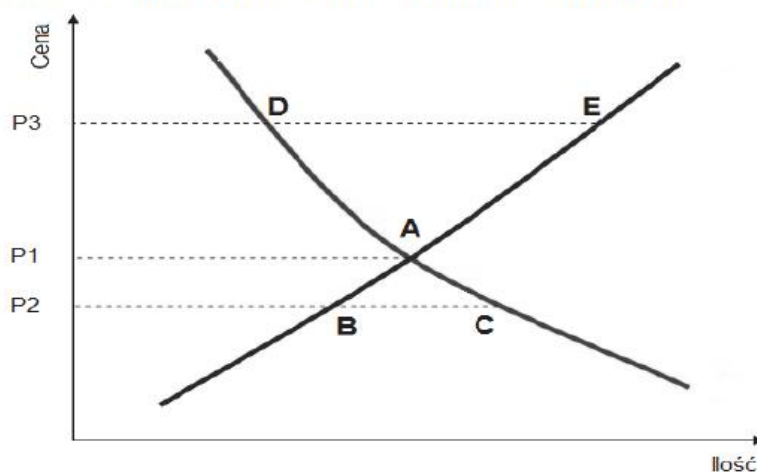
Odwołując się do obu źródeł, wyjaśnij różnicę w poglądach obu autorów na temat istoty władzy w państwie.

Jest to zadanie typowe na egzaminie maturalnym z historii. Wymaga sięgania do wiedzy własnej zdającego oraz korzystania z wielu umiejętności. Rezultat odzwierciedla prace nauczycieli, którzy od kilkunastu lat wzbogacają nauczanie historii, traktując ją nie tylko jako narrację, ale także jako dyscyplinę posługującą się źródłami i zaawansowaną metodologią.

Wiedza i społeczeństwo – zadanie 23

Zadanie 23. (0–2)

Wykres. Mechanizm równowagi rynkowej – przypadek dóbr typowych



Na podstawie: www.nbportal.pl

23.1. Podaj nazwę krzywej, na której umieszczono punkty A, C i D.

.....

23.2. Odnosząc się do danych na wykresie, wyjaśnij działanie ceny jako regulatora rynku.

Wiedza o społeczeństwie jest przedmiotem wiążącym wiedzę z dziedziny politologii, prawa, socjologii, statystyki, ekonomii. Opis i wyjaśnianie procesów społecznych i ekonomicznych możliwy jest dzięki integrowaniu wiedzy z wielu dziedzin. Jest to trudne dla uczniów. Ukazuje to omawiane zadanie, którego rozwiązanie daje obraz niskiej wiedzy z zakresu gospodarki oraz umiejętności interpretowania zjawisk (tutaj mechanizmu równowagi rynkowej). Od kilku lat niepokoją niskie osiągnięcia uczniów z obszaru wiedzy o społeczeństwie.

Zadania realizujące idee zmodyfikowanej matury stanowiły wyzwanie dla zdających. Były na ogół trudniejsze od zadań o właściwościach znanych z poprzednich lat.

Analiza rozwiązań, której daliśmy jedynie przykład, pozwala wyłonić ogólne i szczegółowe umiejętności uczniów, nad którymi warto pracować w szkołach.

Dla autorów zadań i osób odpowiedzialnych za egzamin maturalny analiza jest również bogatym źródłem informacji o właściwościach zadań, stopniu spełnienia ich funkcji. Analiza jest podstawą modyfikowania zadań i prac nad nowymi typami zadań.

Problemy związane z ustawicznym modyfikowaniem matury

W roku 2016 po raz dwunasty przeprowadziliśmy sesję zewnętrznego egzaminu maturalnego, po raz drugi zorganizowaliśmy egzamin zmodyfikowany. Każdego roku egzaminy były starannie monitorowane, zbierane są uwagi zdających, nauczycieli, pracowników uczelni. Po każdej sesji publikowane są szczegółowe wyniki egzaminów oraz szerokie analizy ilościowe i jakościowe egzaminów. Dzięki temu dysponujemy niezwykle bogatymi informacjami o wiedzy i umiejętnościach maturzystów. Znamy też właściwości arkuszy egzaminacyjnych i poszczególnych zadań. Wszystkie zbierane i gromadzone informacje stanowią punkt wyjścia do prac mających na celu doskonalenie matury. Zmiany nie mogą być jednak wprowadzane zbyt szybko i zbyt często. Jest kilka powodów takiego postępowania :

- Konieczność zachowania lojalności wobec uczniów i wobec nauczycieli. Reguły matury wymagają poinformowania o zmianach przynajmniej z dwuletnim wyprzedzeniem.
- Efekty zmian widoczne są najwcześniej po dwóch latach, a nawet po dłuższym czasie od ich wprowadzenia. Dopiero wtedy można zacząć uwzględniać zaobserwowane efekty.
- Aby monitorowanie zmian było skuteczne i przydatne, to wprowadzane zmiany nie powinny obejmować równocześnie zbyt wielu sfer egzaminu, a jedynie jego wybrane aspekty. Dlatego też zadania maturalne tworzące arkusz egzaminacyjny modyfikowane są w kilku kolejnych sesjach egzaminacyjnych, stopniowo wprowadzane jest też ocenianie holistyczne.

Strategia wprowadzania zmiany edukacyjnej poprzez ewolucję egzaminu może dawać dobre efekty. Przykładem jest egzamin maturalny w Polsce.

KONTAKT

Wojciech Malecki

Okręgowa Komisja Egzaminacyjna we Wrocławiu
Polskie Towarzystwo Diagnostyki Edukacyjnej

dr Henryk Szaleniec

Polskie Towarzystwo Diagnostyki Edukacyjnej

EXTERNÉ TESTOVANIA NA SLOVENSKU – OD PAPIEROVÝCH TESTOV K E-TESTOVANIU

Lucia Ficová, Ivana Pichaničová

Abstrakt: Od roku 2000, kedy sa začali realizovať prvé celoplošné monitory na stredných školách v zmysle novej koncepcie maturitnej skúšky, došlo v oblasti externého hodnotenia vo vzdelávaní na Slovensku k významnému progresu. Za obdobie 16 rokov sa externé testovania rozrástli na objektívny systém hodnotenia výsledkov vzdelávania, v rámci ktorého sa realizujú celoplošné merania v kľúčových bodoch vzdelávania, t.j. po ukončení primárneho stupňa vzdelávania (Testovanie 5), po ukončení nižšieho sekundárneho stupňa (Testovanie 9) a na výstupe z vyššieho sekundárneho stupňa (externá časť maturitnej skúšky). V prvej časti článku informujeme o vývoji a súčasnom stave národných (externých) testovaní na Slovensku. Objasňujeme ciele a význam testovaní, ich zakotvenie vo vzdelávacom systéme a poukazujeme na niektoré problémy súvisiace s ich realizáciou a možnosti ich riešenia. Veľkou výzvou v tejto súvislosti je prechod z papierovej na elektronickú formu testovania. Národný ústav certifikovaných meraní vzdelávania v rámci projektu spolufinancovaného z prostriedkov EÚ vytvoril elektronický systém e-Test, ktorý obsahuje nástroje na tvorbu úloh, databázu úloh a umožňuje vykonávať a hodnotiť nielen spomínané národné testovania, ale aj tzv. školské, či učiteľské testovania, elektronicky. V druhej časti článok predstavuje systém e-Test a procesy súvisiace s tvorbou banky úloh. Prezентuje možnosti jeho využitia v pedagogickej praxi. Prináša taktiež ukážky rôznych typov úloh, ktoré sú považované za jeden z benefitov elektronického testovania. V závere sú zrekapitulované dosiahnuté výsledky celoplošných meraní a zamyslenie nad víziou do budúcnosti.

Kľúčová slova: maturitná skúška, testovanie žiakov, elektronické testovanie, e-Test, celoplošné merania, externé testovania, národné testovania, školské testovania, externé hodnotenie vzdelávania, typy úloh

Abstract: Since 2000, when as a result of the new conception of a school-leaving exam (the Maturita exam) the first nationwide pilot testing were administered at secondary schools, there has been a significant progress in the field of external evaluation in Slovak education. During the period of 16 years, national tests have expanded to the objective system of educational quality evaluation. It covers the implementation of nationwide measurements at core levels of education – i.e., after finishing the primary education (Testing 5), after finishing the lower secondary education (Testing 9) and at the outcome of the upper secondary education (external part of school-leaving exam). In the first part of this article, we provide information about development and current situation of national testing in Slovakia. We explain the aims and purposes of testing, its position in the educational system, and simultaneously, we point out some difficulties connected with the implementation of testing as well as some possible ways of solving them. The change from a paper to an electronic form of testing represents a crucial challenge in this field. National Institute for Certified Educational Measurements has, in the frame of the project co-funded by the European Union, developed the electronic system called e-Test, which includes an item bank that enables to perform not only the above mentioned national testing electronically but

also so-called “school testing” and “teacher testing”. The second part of this paper presents the e-Test together with the processes that are closely related to the item and test development. Moreover, it depicts possibilities of its application in educational practice, and introduces samples of different test items which are considered to be one of the benefits of electronic testing. Finally, we summarize the results acquired within the nationwide measurements and revise the vision for the future.

Key words: school- leaving exam, testing of students, electronic testing, e-Test, nationwide testing, national measurements, school testing, external evaluation of education, types of test items

ÚVOD

Kvalita vzdelávania a jej hodnotenie je kľúčovou témou vo všetkých vzdelávacích systémoch. Vo väčšine európskych krajín sa v súčasnosti realizujú národné, externé, celoplošné testovania, ktoré sú vnímané ako objektívny nástroj na meranie a monitorovanie kvality vzdelávania a funkčnosti vzdelávacieho systému. V jednotlivých krajinách sú tieto testovania v rôznej miere zasadené do systému hodnotenia kvality škôl, sú prepájané s autoevalvačnými nástrojmi, a ich význam a dosah pre jednotlivú školu, či žiaka, môže byť rôzny. V niektorých krajinách slúžia výsledky celoplošných (celonárodných) testov, ktoré sú zväčša centrálné organizované a vyhodnocované, ako jedno z kritérií prijímacieho konania na stredné či vysoké školy.

Zavádzanie národných testovaní malo zväčša za cieľ objektivizovať a štandardizovať nástroje hodnotenia úrovne vedomostí a zručností žiakov, na základe ktorých sa rozhoduje o školskej kariére žiaka. V priebehu času sa účel a koncepcia národných testovaní vyvíjali. Menil sa počet ročníkov, v ktorých sa testy administrovali, menila sa štruktúra testov a ich obsah, menili sa účel testovania a význam jeho výsledkov. Zmeny v koncepcii národných meraní v krajinách Európy boli často ovplyvňované aj medzinárodnými štúdiami PISA, TIMSS, PIRLS. Významnou zmenou bolo a je zavádzanie digitálnych technológií do testovania, elektronické administrovanie testov a ich postupná transformácia na elektronické testovanie.

Obdobným vývojom prechádzajú aj externé národné testovania na Slovensku. Ich počiatky úzko súvisia s reformnými zmenami po roku 1989 a potrebou objektivizovať, či modernizovať maturitnú skúšku. Od roku 2000 sa na Slovensku realizujú celoštátne testovania výsledkov vzdelávania na výstupe z vyššieho sekundárneho stupňa vzdelávania. Následne sa od roku 2003 rozšírili aj na úroveň nižšieho sekundárneho vzdelávania a od roku 2012 aj na primárne vzdelávanie. Súčasne od roku 2003 prebiehajú na Slovensku medzinárodné štúdie PISA, TIMSS, PIRLS, a i.

Tieto merania boli spočiatku realizované pracovníkmi Úseku certifikovaných meraní vzdelávania Štátneho pedagogického ústavu (ŠPÚ). Keď slovenské ministerstvo školstva zriadilo ku dňu 1. 9. 2008 Národný ústav certifikovaných meraní vzdelávania (NÚCEM), všetky aktivity spojené s realizáciou národných meraní a medzinárodných meraní prešli do agendy NÚCEM-u.

Od svojho vzniku si NÚCEM vybuďoval pozíciu inštitúcie, ktorá významnou mierou prispieva ku skvalitňovaniu vzdelávania na Slovensku. Jedinečné postavenie v rámci ostatných priamo riadených organizácií ministerstva školstva má NÚCEM najmä v problematike monitorovania a hodnotenia kvality vzdelávania na národnej i medzinárodnej úrovni. Výsledky meraní slúžia ako dôležitá spätná väzba pre slovenský vzdelávací systém aj pre jednotlivé školy a ich zriaďovateľov. Popri organizačnom a odbornom zastrešovaní externých národných meraní prebieha v NÚCEM-e aj výskumná činnosť v oblasti hodnotenia

výsledkov vzdelávania. Výsledky meraní a výskumnej činnosti tvoria východiskovú bázu pre iniciáciu reforiem a zmien v oblasti vzdelávania v SR, na základe výstupov činnosti NÚCEM-u sú inovované pedagogické dokumenty. NÚCEM v období posledných piatich rokov urobil ďalší krok v smere k objektívnejšiemu hodnoteniu kvality škôl a podpore autoevalvácie. V roku 2015 prvýkrát poskytol stredným školám informáciu o ich pridanej hodnote vo vzdelávaní a plánuje tento dôležitý prvok hodnotenia vzdelávania rozšíriť aj na základné školy. V súčasnosti je pre NÚCEM dôležitou výzvou aj elektronická forma testovania.

1 EXTERNÉ HODNOTENIE VO VZDELÁVANÍ V SR

1.1 Východiská a začiatky externého celoplošného testovania – nová koncepcia maturitnej skúšky

Slovenské školstvo, tak ako celá spoločnosť, prešli od roku 1989 významnými zmenami. Vznikli súkromné a cirkevné školy, väčší dôraz sa začal klásť na vyučovanie cudzích jazykov, učitelia dostali väčšiu voľnosť pri napĺňaní obsahu vyučovacích hodín. Jednou z množstva prebiehajúcich koncepčných zmien bola aj inovácia maturitnej skúšky. Z prieskumov a výsledkov diskusií k problematike maturitnej skúšky, ktoré prebiehali medzi odbornou verejnosťou po roku 1990, vyplynulo, že vtedajší model maturitnej skúšky plnil túto funkciu iba čiastočne, maturitná skúška nebola odborne akceptovateľným meradlom vedomostí, zručností a všeobecných kompetencií absolventov stredných škôl. Medzi najzávažnejšie nedostatky maturitnej skúšky sa považovali malá objektivita a nízka validita. Bolo potrebné navrhnúť vhodnú alternatívu maturitnej skúšky, ktorá by spĺňala požadovanú úroveň objektivitu, validity a porovnateľnosti s ďalšími európskymi krajinami.

Problémom ako zobjektívniť maturitnú skúšku, sa začali zaoberať riaditelia mnohých gymnázií a aj učitelia na svojich stretnutiach. Vznikla komisia riaditeľov škôl, učiteľov a ďalších odborníkov, ktorí po preskúmaní situácie v niektorých blízkyh krajinách a na základe vlastných skúseností s medzinárodnými projektmi realizovanými na školách vypracovali návrh projektu experimentálnej maturity. Vyústením týchto aktivít bolo v roku 1998 rozhodnutie Ministerstva školstva SR (MŠ SR) o potrebe inovovať maturitnú skúšku a poverenie Štátneho pedagogického ústavu (ŠPÚ) prípravou novej koncepcie maturitnej skúšky (NKMS) v zmysle schváleného Národného programu výchovy a vzdelávania a projektu Milénium. ŠPÚ zabezpečoval koncepčnú aj organizačnú stránku prípravy a realizácie NKMS. Návrh NKMS bol schválený v novembri 2001.

Súčasťou prípravy NKMS bol prvý ročník celoslovenského testovania maturantov pod názvom MONITOR 1999, ktorý formou externe vytvorených testov overoval stav vedomostí budúcich maturantov z matematiky. Na základe pozitívneho ohlasu sa zrealizovali aj ďalšie monitorovacie testovania (MONITOR 2000 – 2003), pričom v každom roku pribúdali testované predmety a zvyšoval sa počet testovaných žiakov (Obr. 1). Na základe výsledkov monitorovacích testovaní a informácií získaných z učiteľských dotazníkov, ktoré boli pri každom monitore ponúkané, a aj z pripomienok a návrhov učiteľov získaných z iných zdrojov sa postupne skvalitňovali a zefektívňovali jednotlivé testovacie nástroje a činnosti nevyhnutné pre zabezpečenie celoplošného externého testovania.

Medzi základné zmeny, ktoré do maturitnej skúšky priniesla NKMS, patrili povinná externá časť maturitnej skúšky (test) v niektorých predmetoch, externe pripravované a centrálné zadávané témy/zadania písomnej formy internej časti z vyučovacích a cudzích jazykov, centrálné vypracované pokyny na tvorbu zadaní pre ústnu formu internej časti a jej hodnotenie. Generálna skúška NKMS 2004 (GS NKMS 2004) bola prípravným krokom pred oficiálnym začiatkom novej maturity v školskom roku 2004/2005. Uskutočnila

sa z predmetov anglický jazyk, nemecký jazyk a matematika na vzorke 51 965 žiakov stredných škôl. Hlavným cieľom GS NKMS 2004 bolo overiť organizačnú stránku realizácie externej časti MS, jej logistickú, finančnú, personálnu a časovú náročnosť, umožniť všetkým stredným školám s maturitnými ročníkmi vyskúšať si proces administrácie testov externej časti, získať predstavu o náročnosti realizácie externej časti MS, o forme testovacích nástrojov a o spôsobe ich vyhodnocovania, ale aj poskytnúť všetkým maturantom možnosť otestovať si svoje vedomosti z testovaných predmetov a porovnať ich úroveň vedomostí s ostatnými spolužiakmi v rámci celého Slovenska.

1.2 Externá maturita po roku 2004

Od školského roku 2004/2005 boli maturitné skúšky aj iné formy ukončovania stredoškolského štúdia realizované v zmysle platnej legislatívy, teda existujúceho školského zákona a vyhlášok MŠ SR, ktoré upravovali spôsob ukončovania štúdia na stredných školách. Žiaci slovenských gymnázií štandardne povinne maturujú zo 4 predmetov (povinný je slovenský jazyk a jeden cudzí jazyk, ostatné dva predmety sú voliteľné). Žiaci škôl s vyučovacím jazykom maďarským majú povinných 5 predmetov (navyše je maturita z maďarského jazyka). K týmto povinným predmetom si môžu maturanti zvoliť ešte ďalšie dva tzv. dobrovoľné predmety. Od roku 2005 prebieha externá časť a písomná forma internej časti maturitnej skúšky z cudzích jazykov (anglický, nemecký, ruský, francúzsky, španielsky a taliansky jazyk) a matematiky, ktorá je jedným z voliteľných predmetov pre gymnazistov, t.j. je to jediný voliteľný predmet, z ktorého žiak okrem ústnej formy internej časti musí vykonať aj externú časť, test. Žiaci stredných odborných škôl si matematiku môžu voliť iba ako dobrovoľný maturitný predmet. Od roku 2007 je povinná externá časť maturitnej skúšky aj z vyučovacích jazykov (slovenského jazyka a literatúry, maďarského jazyka a literatúry, ukrajinského jazyka a literatúry, slovenského jazyka a slovenskej literatúry pre školy s vyučovacím jazykom maďarským).

Platná legislatíva v rokoch 2005 – 2008 umožňovala žiakom stredných škôl maturovať z väčšiny maturitných predmetov na dvoch úrovniach – základnej (B) a vyššej (A). V prípade cudzích jazykov mohli žiaci v rokoch 2007 a 2008 maturovať na troch úrovniach, pretože na základe požiadavky zo strany stredných odborných škôl a učilíšť bola doplnená úroveň C pre žiakov škôl, ktorých hodinová dotácia vyučovania cudzieho jazyka nespĺňala požadované kritériá.

Významným medzníkom v slovenskom školstve bolo spustenie obsahovej reformy vzdelávania, ktorá bola odštartovaná novým školským zákonom (zákon č. 245/2008 Z. z.). Základným pilierom tejto reformy bol dvojstupňový model tvorby obsahu vzdelávania, t.j. vypracovanie štátnych vzdelávacích programov pre jednotlivé stupne vzdelávania a vzdelávacie oblasti a následne ich rozpracovanie školami do vlastných školských vzdelávacích programov.

Nový školský zákon priniesol veľmi významnú zmenu – zriadenie inštitúcie s celoslovenskou pôsobnosťou (NÚCEM), ktorej hlavnou úlohou je zabezpečovať externé celoplošné testovania a medzinárodné merania na základných a stredných školách. V súčasnosti pri príprave a realizácii externej časti maturitnej skúšky, Testovania 9 a Testovania 5 s NÚCEM-om spolupracujú aj odbory školstva okresných úradov a Centrum vedecko-technických informácií – školské výpočtové strediská, ktorých úlohou je najmä sprostredkovanie priameho kontaktu so školami a zber dát. S NÚCEM-om spolupracuje aj Štátna školská inšpekcia vo forme výkonu kontrolnej činnosti počas testovaní a následného riešenia prípadných podnetov ohľadom podozrení z nedodržania organizačných pokynov počas testovania a neobjektívneho hodnotenia.

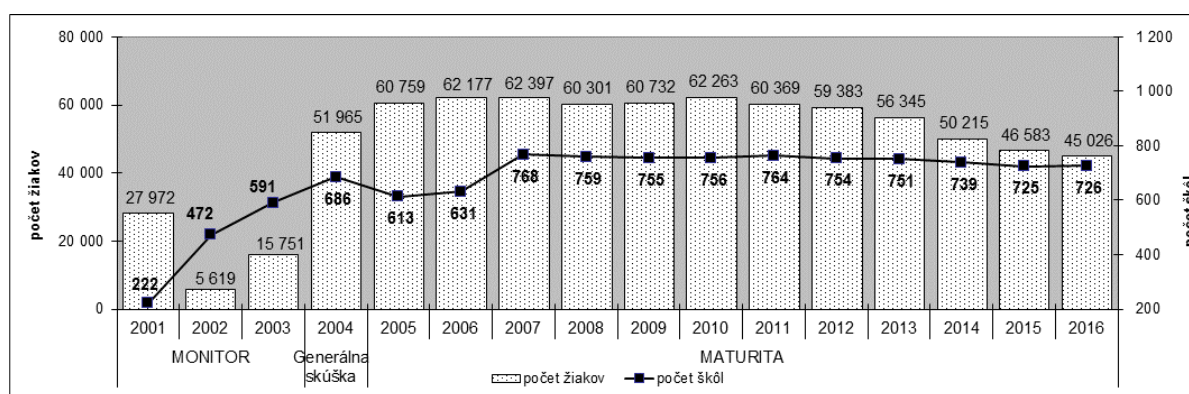
Organizácia maturitných skúšok sa od školského roku 2008/2009 riadi podľa nového školského zákona a príslušných vyhlášok. Jednou zo zmien, ktoré priniesla táto nová legislatíva, bolo zrušenie úrovni vo všetkých predmetoch maturitnej skúšky okrem cudzích jazykov, kde zostali dve úrovne B1 a B2, ktoré vychádzajú zo Spoločného európskeho referenčného rámca pre jazyky. Vyhláška, ktorá od roku 2008 umožňovala uznať vybrané jazykové certifikáty ako náhradu za vykonanie externej časti maturitnej skúšky z príslušného cudzieho jazyka, bola neskôr, v roku 2014, zrušená.

Od roku 2009 môžu žiaci vykonať externú časť maturitnej skúšky iba z jedného cudzieho jazyka. V súčasnosti žiaci gymnázií externe maturujú z cudzieho jazyka povinne na úrovni B2 a žiaci SOŠ a konzervatórií si môžu vybrať medzi úrovňou B1 a B2. Od roku 2017 žiaci bilingválnych tried budú vykonávať maturitnú skúšku z cudzieho jazyka, ktorý je druhým vyučovacím jazykom na úrovni C1 Spoločného európskeho referenčného rámca pre jazyky.

Novelou školského zákona sa od roku 2013 umožnilo žiakom vykonať opravnú skúšku externej časti a písomnej formy internej časti maturitnej skúšky v jesennom termíne (dovtedy bola možná oprava až v nasledujúcom školskom roku pri riadnom termíne externej maturity).

Informácie o počte zúčastnených žiakov a škôl v maturitných MONITOROCH a externých maturitných testoch prezentuje Obr. 1. Z obrázka vyplýva, že počet maturantov mierne klesá, čo je ovplyvnené demografickým vývojom na Slovensku. Z pohľadu počtu maturantov z jednotlivých predmetov pozorujeme, že počet maturantov z matematiky sa iba mierne znižuje, teda zrušením dvoch úrovní nedošlo k významnej zmene záujmu žiakov o maturitnú skúšku z tohto predmetu. Každoročne si maturitu z matematiky zvolí cca 14 % žiakov. Naopak pri cudzích jazykoch vplyvom vyššie uvádzaných legislatívnych zmien došlo k výraznému poklesu maturantov z francúzskeho, španielskeho a talianskeho jazyka a sledujeme vzostup záujmu o maturitu z ruského jazyka.

Taktiež zaznamenávame významný nárast počtu maturantov so zdravotným znevýhodnením, pričom ich výsledky v externej maturite z niektorých predmetov sú porovnateľné s výsledkami intaktných žiakov (AJB2, MAT, NJB1, RJB1) ale v niektorých predmetoch (napr. SJL, AJB1) je úspešnosť žiakov so zdravotným znevýhodnením štatisticky významne nižšia ako u intaktnej populácie.



Obr. 1 Počet zúčastnených žiakov a škôl – MONITOR/Maturita

V priebehu realizácie externých maturít sa menila aj váha výsledku maturitného testu. V prvých rokoch postačovalo, keď mal žiak z ústnej formy internej časti známku 3 alebo lepšiu na to, aby zmaturoval. Od roku 2012 stúpol vplyv výsledku z externej maturity.

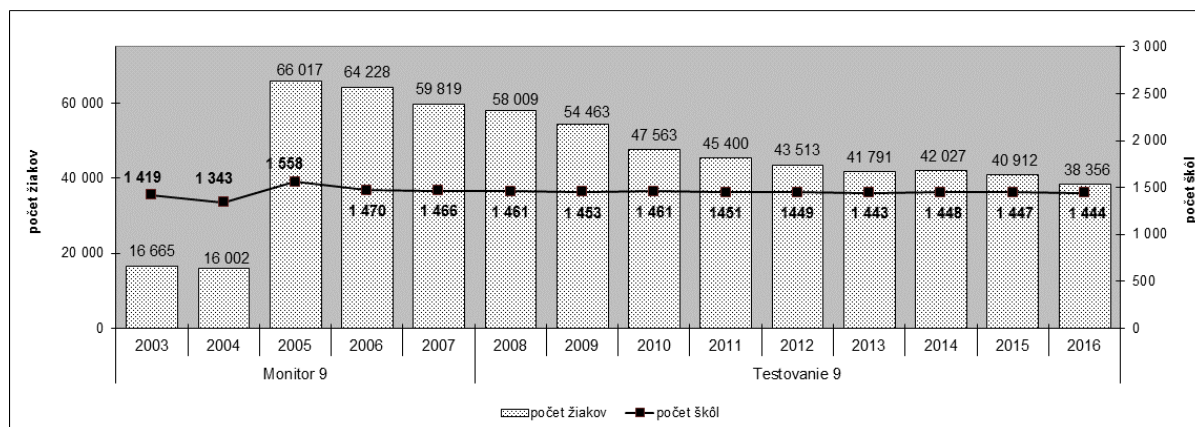
V prípade, že má žiak z ústnej odpovede známku 5, nezmaturoje z daného predmetu, rovnako ako aj v prípade, že z písomnej formy internej časti bude mať úspešnosť 25% alebo menej, resp. z externej časti 33% alebo menej. Vyhláška o ukončovaní štúdia na stredných školách presne stanovuje dolnú hranicu úspešnosti pre jednotlivé predmety v závislosti od známky, ktorú žiak dostane na ústnej časti.

1.3 Externé celoplošné testovania na základných školách

Ministerstvo školstva SR v školskom roku 2002/2003 poverilo vtedajší ŠPÚ odbornou gesciou a organizáciou testovania žiakov 9. ročníka základných škôl. Projekt sa uskutočnil pod názvom MONITOR 9 a zameril sa na testovanie žiaka ako jednotlivca. Účelom bol celoslovenský prieskum individuálnej úrovne vedomostí žiakov na výstupe z ISCED 2. Výber škôl a žiakov bol zámerný, podľa vopred stanoveného kritéria. MONITOR 9 bol určený žiakom 9. ročníka ZŠ s vyučovacím jazykom slovenským a maďarským (okrem žiakov s mentálnym postihnutím a žiakov so špeciálnymi výchovno-vzdelávacími potrebami), ktorí si v školskom roku podali aspoň jednu prihlášku na gymnázium. Žiaci boli testovaní z matematiky a vyučovacieho jazyka. Podobne sa uskutočnilo testovanie aj v školskom roku 2003/2004.

V školskom roku 2004/2005 sa po prvýkrát zapojili do testovania aj žiaci so špeciálnymi výchovno-vzdelávacími potrebami (ŠVVP) okrem žiakov s mentálnym postihnutím. Testovanie sa uskutočnilo celoplošne. Testovacie nástroje boli vytvorené aj pre žiakov s vyučovacím jazykom ukrajinským a od roku 2005 sa testovanie uskutočňuje každoročne na celej populácii žiakov 9. ročníkov ZŠ okrem žiakov s mentálnym postihnutím. Od roku 2009 MONITOR 9, resp. pod novým názvom Testovanie 9, zabezpečuje NÚCEM. Žiaci sú aj naďalej testovaní z matematiky, z vyučovacích jazykov (slovenský jazyk, maďarský jazyk, ukrajinský jazyk) a na školách s vyučovacím jazykom slovenským aj zo slovenského jazyka a slovenskej literatúry.

Prehľad počtu zapojených žiakov a základných škôl do Testovania 9 prináša Obr. 2. Možno tvrdiť, že počet testovaných žiakov kopíruje demografický vývoj na Slovensku, pričom počet základných škôl sa významne nemení. Pozorujeme, že celkové zastúpenie žiakov so zdravotným znevýhodnením na školách rastie, pričom najväčší prírastok tvoria žiaci s vývinovými poruchami učenia. Z analýz vyplýva, že výsledky žiakov so zdravotným znevýhodnením nevykazujú štatisticky významné rozdiely voči skupine intaktných žiakov s rovnakou priemernou známku.



Obr. 2 Počet zúčastnených žiakov a škôl – Monitor 9/Testovanie 9

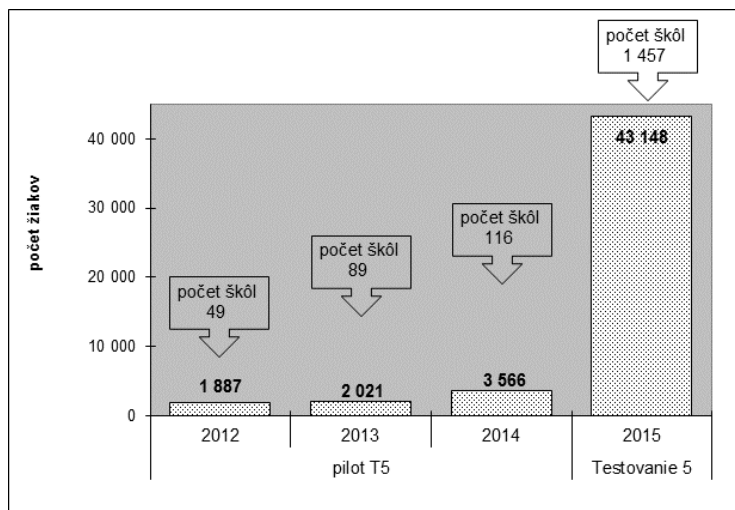
V zmysle platnej školskej legislatívy sú výsledky žiakov v Testovaní 9 zohľadňované pri prijímacom konaní na stredné školy. Váhu výsledku Testovania 9 pri určovaní kritérií výberu uchádzačov na strednú školu určuje riaditeľ školy, avšak školský zákon stanovuje, že v prípade žiakov, ktorí zo všetkých testovaných predmetov súčasne dosiahnu úspešnosť 90% a viac, musí riaditeľ školy rozhodnúť bez prijímacej skúšky.

NÚCEM sa hneď od svojho vzniku začal venovať aj hodnoteniu matematickej a čitateľskej gramotnosti žiakov na národnej úrovni. Podnetom na to boli práve alarmujúce zistenia medzinárodného výskumu OECD PISA. V školskom roku 2007/2008 sa po prvýkrát uskutočnilo testovanie matematickej a čitateľskej gramotnosti. Testovanie matematickej a čitateľskej gramotnosti žiakov 9. ročníka ZŠ, ktoré realizoval NÚCEM na národnej úrovni v rokoch 2009 - 2011, sa na Slovensku celkovo konalo štyrikrát na reprezentatívnom výbere základných škôl s vyučovacím jazykom slovenským a maďarským. Od roku 2012 sa testovanie matematickej a čitateľskej gramotnosti zakomponovalo priamo do testov Testovania 9.

V období rokov 2005 až 2012 teda NÚCEM (do roku 2008 ŠPÚ) realizoval národné merania na úrovni ISCED 2 a ISCED 3 na celej populácii žiakov základných a stredných škôl v zmysle platnej legislatívy a príslušných pedagogických dokumentov. Tieto merania boli nastavené na koncepciu vzdelávania a štandardy pred spustením reformy vzdelávania, ktorú inicioval nový školský zákon z roku 2008. NÚCEM v rámci realizácie projektu spolufinancovaného z EÚ „Hodnotenie kvality vzdelávania na základných a stredných školách SR v kontexte prebiehajúcej obsahovej reformy vzdelávania“ inovoval testovacie nástroje pre tieto merania po formálnej a obsahovej stránke v súlade s obsahovou reformou. Súčasne navrhol a overoval systém monitorovania a hodnotenia výsledkov vzdelávania na národnej úrovni pre vzdelávacie stupne ISCED 1 – ISCED 3, v rámci ktorého je možné sledovať, do akej miery obsahová reforma školstva v SR napĺňa stanovené ciele na uvedených stupňoch vzdelávacej sústavy v SR. V tomto kontexte projekt zahŕňal aj prípravu a realizáciu doposiaľ v slovenskom vzdelávacom systéme absentujúcich meraní výsledkov vzdelávania na výstupe z úrovne ISCED 1.

Počas spomínaného projektu v rokoch 2012 až 2013 NÚCEM realizoval pilotné overovanie testovacích nástrojov v rámci príprav na generálnu skúšku testovania žiakov 5. ročníka ZŠ v školskom roku 2014. Cieľ tohto pilotného testovania bolo overiť testovacie nástroje z matematiky a vyučovacích jazykov (slovenského jazyka a literatúry a maďarského jazyka a literatúry) vytvorené už v zmysle nových štátnych vzdelávacích programov pre ISCED 1.

V roku 2014 sa v rámci príprav na celoplošné testovanie žiakov 5. ročníka základných škôl uskutočnila jeho generálna skúška, ktorá overila aj organizačné zabezpečenie a logistiku testovania. V roku 2015 sa po prvýkrát zrealizovalo celoplošné Testovanie 5, do ktorého sa zapojili všetci žiaci 5. ročníka základných škôl okrem vybraných skupín žiakov so ŠVVP a mentálnym postihnutím. Žiaci boli testovaní z matematiky, zo slovenského jazyka a literatúry a z maďarského jazyka a literatúry. Výsledky žiakov NÚCEM zasielal nielen na základné školy, na ktorých testovaní žiaci aktuálne navštevovali 5. ročník, ale aj na ich bývalé základné školy, na ktorých absolvovali 4. ročník so zámerom podporiť autoevalvačné procesy na školách. Počet zapojených žiakov a škôl do Testovania 5 prezentuje Obr. 3.



Obr. 3 Počet zúčastnených žiakov a škôl – Testovanie 5

1.4 Testovacie nástroje - validita a objektivita meraní

Pri tvorbe testovacích nástrojov vo všetkých realizovaných externých meraniach na Slovensku (externá časť maturitnej skúšky, Testovanie 9, Testovanie 5) sa vychádza z platných pedagogických dokumentov, ktoré pripravujú predmetové komisie pri ŠPÚ. Pre maturitnú skúšku sú centrálné vypracované cieľové požiadavky, ktoré presnejšie vymedzujú nároky na maturantov v jednotlivých predmetoch. Pre Testovanie 5 a Testovanie 9 je štátny vzdelávací program záväzný dokument, ktorý stanovuje všeobecné ciele vzdelávania a kľúčové kompetencie pre jednotlivé vzdelávacie stupne, ku ktorým má vzdelávanie v jednotlivých vzdelávacích oblastiach a predmetoch smerovať. Štátny vzdelávací program vymedzuje aj rámcový obsah vzdelávania. Je východiskom pre tvorbu školského vzdelávacieho programu, v ktorom sa zohľadňujú aj špecifické podmienky a potreby regiónu. Všetky celoplošné externé testy, ktoré NÚCEM realizuje, sú plne v súlade so štátnym vzdelávacím programom pre daný stupeň vzdelávania (Testovanie 5 a Testovanie 9) a cieľovými požiadavkami (externá časť maturitnej skúšky a písomná forma internej časti maturitnej skúšky).

Vysoká validita a objektivita externých testov je zabezpečená zadávaním štandardizovaných, externe pripravovaných a utajených testov, ktoré pripravujú tímy odborníkov na základe platných cieľových požiadaviek, resp. štátnych vzdelávacích programov. Všetky úlohy a testy prechádzajú pilotážou a recenzným posudzovaním. NÚCEM pred každým testovaním zverejňuje špecifikácie testov. Štruktúru testov popisuje Tab. 1.

Tab. 1 Štruktúra testov

predmet	Maturita		Testovanie 9	Testovanie 5
	PFIČ	EČ		
SLJ, MJL	150 min	100 min (40 úvo + 24 úko)	60 min (10 úvo + 10 úko)	60 min (20 úvo + 10 úko)
SJSL, UJL	150 min	100 min (40 úvo + 24 úko)	50 min (10 úvo + 10 úko)	-
cudzie jazyky - C1	90 min	150 min (64 úvo + 26 úko)	-	-

cudzie jazyky - B2	60 min	120 min (46 úvo + 34 úko)	-	-
cudzie jazyky - B1	60 min	100 min (36 úvo + 24 úko)	-	-
matematika	-	150 min (20 úvo + 10 úko)	60 min (10 úvo + 10 úko)	60 min (10 úvo + 20 úko)

Legenda: PFIC – písomná forma internej časti, EČ – externá časť, úvo – úlohy s výberom odpovede, úko – úlohy s krátkou odpoveďou.

Úlohy v testoch sa zameriavajú na logické myslenie, porozumenie textu, overovanie hĺbky vedomostí a zručností, aplikáciu poznatkov v praktických súvislostiach. V testoch z vyučovacích jazykov majú už niekoľko rokov svoje stále zastúpenie aj úlohy na čítanie s porozumením, v ktorých žiaci pracujú s ukážkami súvislých a nesúvislých textov. V testoch z cudzích jazykov sa testujú zručnosti a kľúčové kompetencie ako počúvanie s porozumením, čítanie s porozumením, gramatika v kontexte, schopnosť prezentovať vlastný prejav písomnou formou (v cudzích aj vyučovacích jazykoch). Testy z matematiky preverujú nielen vedomosti a zručnosti žiakov, ale sú zamerané aj na matematickú gramotnosť, obsahujú aj úlohy začlenené do kontextu praktického života, úlohy obsahujú grafy, tabuľky a pod. (Žiaci majú k dispozícii prehľad matematických vzťahov – Testovanie 9 a externá maturita.)

Procesy v priebehu prípravy, tlače a balenia testov podliehajú prísnyim bezpečnostným opatreniam. Tlač, balenie a distribúciu testovacích nástrojov v prípade Testovania 5 a Testovania 9 zabezpečuje externý dodávateľ, ktorý musí dodržiavať prísne zmluvne stanovené bezpečnostné pravidlá. NÚCEM naopak zabezpečuje kompletnú logistiku, tlač a balenie testovacích nástrojov pre externú časť a písomnú formu internej časti maturitnej skúšky taktiež za dodržania prísnych bezpečnostných podmienok. V logistickom zabezpečení externých maturitných testovaní sa každoročne optimalizujú procesy tak, aby sa zvýšila bezpečnosť manipulácie s testovacími nástrojmi a zavádzajú sa opatrenia, ktoré pôsobia preventívne voči nežiaducemu úniku informácií a prispievajú k objektívnejšiemu hodnoteniu úloh s krátkou odpoveďou.

Objektivita administrácie testovania je zaručená tým, že sa testovanie realizuje v tom istom čase na všetkých školách pri zachovaní rovnakých podmienok a harmonogramu počas administrácie testov, pričom všetky činnosti sú vykonávané podľa presne vypracovaných pokynov. Kontrolu objektivity testovania na školách zabezpečuje externý dozor – učiteľ z inej školy, v prípade maturitnej skúšky v pozícii predsedu predmetovej, resp. školskej maturitnej komisie. Externý dozor je na škole prítomný už od rozbalenia zásielky s testami ráno až do zabalenia odpoveďových hárkov do bezpečnostných obálok po ukončení testovania popoludní. Kontrolu regulárnosti priebehu testovaní každoročne vykonáva aj Štátna školská inšpekcia.

NÚCEM pripravuje testy vo väčšine predmetov v dvoch formách líšiacich sa poradím úloh a distraktorov v prípade úloh s výberom odpovedí. Testy z matematiky a pokyny k testom z cudzích jazykov sa pripravujú aj v maďarskom preklade pre žiakov zo škôl s vyučovacím jazykom maďarským. K testom z cudzích jazykov sú vytvorené zvukové nahrávky.

Z cudzích a vyučovacích jazykov v prípade maturitnej skúšky NÚCEM vytvára zadania a témy písomnej formy maturitnej skúšky. Pre vyučovacie jazyky sú vytvorené tri štvorice tém, z ktorých sa následne vylosuje jedna štvorica a zverejňuje sa vo vysielaní Slovenského rozhlasu a tiež sa zverejňuje na webových stránkach. Pre cudzie jazyky sú vytvárané tri

štruktúrované zadania pre každú úroveň (pre všetky jazyky ide o jednotnú tému), z ktorých sa losuje jedno a následne zverejňuje na webových stránkach.

Pre žiakov so zdravotným znevýhodnením sú testy, resp. zvukové nahrávky, upravené na základe individuálnych potrieb žiaka. Upravujú sa hlavne po formálnej stránke, overujú rovnaké špecifické ciele ako u žiakov intaktnej populácie. Takto upravených testov sa ročne pripravuje vyše 200 rôznych druhov, zväčša ide o zmenu veľkosti a zvýraznenie písma, prípadne o nahradenie položiek testu, ak zdravotné znevýhodnenie bráni porozumeniu či riešeniu pôvodnej úlohy. Testy sa tiež upravujú do elektronickej formy na CD a niektoré sú upravené do Braillovo písma. V prípade nepočujúcich žiakov sa v testoch z cudzích jazykov vynecháva časť počúvanie s porozumením. Pre niektoré skupiny postihnutia sa iba predlžuje čas na vykonanie testu.

Objektivita hodnotenia testovania je zabezpečená automatizovaným vyhodnocovaním odpovedových hárkov, do ktorých žiaci zaznačujú svoje odpovede. Testy Testovania 9 obsahujú iba úlohy s výberom odpovede, v prípade matematiky aj s tvorenou číselnou odpoveďou, teda je možné ich kompletne vyhodnocovať strojovo skenovaním. Testy Testovania 5 obsahujú aj krátke slovné odpovede, ktoré sú hodnotené centrálné v hodnotiacom centre a následne automatizovane spracovávané a vyhodnotené. Úlohy s krátkou odpoveďou v testoch externej časti maturitnej skúšky (okrem matematiky) sú najprv hodnotené na školách hodnotiteľmi podľa stanovených kritérií a pod dozorom predsedov predmetových maturitných komisií ihneď po ukončení testovania. Následne sú odpovedové hárky externej časti maturitnej skúšky zasielané na spracovanie – skenujú sa a automatizovane vyhodnocujú.

Práce žiakov písomnej formy internej časti sú hodnotené na školách, ich hodnotenie je plne v kompetencii učiteľov školy, ktorí sú povinní dodržiavať pokyny a kritériá pre hodnotenie vypracované NÚCEM-om. Kontrolu tohto hodnotenia vykonáva neskôr predseda predmetovej komisie, prípadne pracovník Štátnej školskej inšpekcie.

S cieľom odhaliť prípadné porušenie pravidiel objektivity testovania sa v NÚCEM-e po spracovaní a vyhodnotení odpovedových hárkov vykonáva spätná kontrola dokumentov a analýza odpovedí žiakov. Na základe zistení z tejto kontroly NÚCEM podáva podnety na prešetrenie podozrenia z neobjektívne vykonávanej administrácie testov alebo hodnotenia Štátnej školskej inšpekcií. Na základe takýchto podnetov boli potvrdené viaceré podozrenia a prijaté opatrenia v súlade s platnou legislatívou.

1.5 Ciele a význam externého plošného testovania, pridaná hodnota vo vzdelávaní

Spoločným cieľom externých testovaní na Slovensku je porovnať výkony žiakov danej cieľovej skupiny v testovaných predmetoch a na základe ich výsledkov poskytnúť spätnú väzbu školám o úrovni vedomostí ich žiakov aj v porovnaní s ostatnými školami na Slovensku.

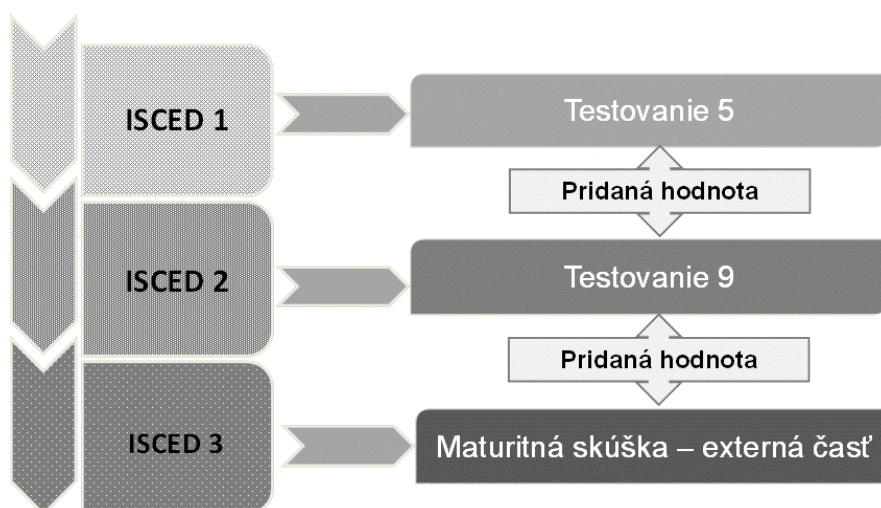
Externé testovania predstavujú objektívny nástroj na hodnotenie výsledkov vzdelávania. Vo všetkých troch prípadoch (externá časť maturitnej skúšky, Testovanie 9 a Testovanie 5) ide o sumatívne hodnotenie žiakov na výstupe z príslušného vzdelávacieho stupňa. Ich realizáciou sa získavajú objektívne informácie o výkone žiakov a na základe ich analýzy je možné poskytnúť školám, decíznej sfére, širokej odbornej verejnosti spätnú väzbu a komplexnejší obraz o vedomostiach a zručnostiach žiakov z testovaných predmetov, ktorý môže pomôcť pri zvyšovaní kvality vzdelávania. Výsledky testovania, ktoré rozlišujú žiakov

podľa ich výkonov v Testovaní 9 a externej maturite, môžu a v súčasnosti aj slúžia ako jedno z kritérií prijatia žiakov na stredné, resp. vysoké školy.

Externé testovania Testovanie 9 a externá maturita sa za viac ako 10 rokov svojej existencie, napriek diskusiám, ktoré prebiehajú medzi odborníkmi aj širokou verejnosťou, stali dôležitým prvkom školského systému na Slovensku. NÚCEM každoročne zverejňuje na svojej webovej stránke necelý mesiac po testovaní informáciu o výsledkoch na celoslovenskej úrovni. Žiaci aj učitelia netrpezlivo očakávajú výsledkové listiny, ktoré NÚCEM distribuuje do škôl zároveň so zverejnením celoštátnych výsledkov. Aj odborná verejnosť i médiá venujú zverejneným výsledkom a analýzám testovaní veľkú pozornosť. Verejnosť, zriaďovatelia aj samotné školy veľmi oceňujú portál Výsledky testovaní (<http://vysledky.nucem.sk/>), ktorý NÚCEM prevádzkuje od roku 2012. Na tomto portáli je možné prehľadným spôsobom prezerat' výsledky jednotlivých škôl a filtrovat' ich podľa požadovaných kritérií.

Zverejnené výsledky vždy vyvolávajú vlnu pozitívnych, ale aj negatívnych reakcií, prebiehajú verejné i odborné diskusie na tému významu testovania, obsahu, či kvality testovacích nástrojov. V diskusiách súvisiacich s mierou akceptácie a pripisovania váhy výsledkom externých testovaní pri posudzovaní kvality škôl NÚCEM neustále zdôrazňuje, že tieto údaje treba vnímať iba ako jeden z viacerých ukazovateľov kvality škôl. Médiá totiž veľmi radi výsledky testovaní spracovávajú do tzv. rebríčkov, v niektorých prípadoch nevhodne, či nekorektne interpretujú výsledky. NÚCEM zastáva názor, že výsledky škôl je nutné zverejňovat', aby školy mali spätnú väzbu a zároveň cítili zodpovednosť za kvalitu vzdelávania, ktoré poskytujú pod dohľadom verejnosti. Zároveň však tieto výsledky musia školy interne seriózne vyhodnotiť s prihliadnutím na špecifiká, ktoré každá škola má. NÚCEM pokladá za veľmi dôležité byť v neustálom kontakte s riaditeľmi škôl a ich zriaďovateľmi. Na pravidelných stretnutiach so zriaďovateľmi škôl zamestnanci NÚCEM-u prezentujú analýzy výsledkov škôl pre jednotlivé kraje, vysvetľujú, ako výsledky interpretovat' a akú mieru významu im treba pripisovat'.

Vzhľadom na to, že výsledky celoplošných testov majú svoje obmedzenia a nemožno ich vnímať ako jediné a správne kritérium na posúdenie kvality školy, NÚCEM sa už niekoľko rokov v rámci svojej výskumnej činnosti zaoberá problematikou ďalších merateľných indikátorov kvality vzdelávania. Jedným z takýchto indikátorov je tzv. pridaná hodnota vzdelávania, ktorej koncept a metodiku NÚCEM overuje od roku 2011, kedy v rámci národného projektu spolufinancovaného z prostriedkov EÚ „Hodnotenie kvality vzdelávania na základných a stredných školách SR v kontexte prebiehajúcej obsahovej reformy vzdelávania“ navrhol a overoval model systému monitorovania a hodnotenia výsledkov vzdelávania na národnej úrovni pre vzdelávacie stupne ISCED 1 – ISCED 3 v SR. V rámci tohto systému bude možné prepájať výsledky testov medzi dvoma stupňami vzdelávania a zisťovat' tak pridanú hodnotu vzdelávania tak pre základné ako aj pre stredné školy. Schematické znázornenie tohto konceptu prezentuje Obr. 4.



Obr. 4 Systém monitorovania a hodnotenia výsledkov vzdelávania na Slovensku

V septembri 2016 zaslal NÚCEM už po druhýkrát stredným školám informáciu o ich pridanej hodnote vo vzdelávaní zo slovenského jazyka a literatúry. Okrem výsledkov z externej Maturity 2016 tak riaditelia stredných škôl získali ďalšie externé hodnotenie, ktoré predstavuje príspevok školy k vzdelaniu ich žiakov z tohto povinného maturitného predmetu. Školy sa prostredníctvom tohto údaja dozvedeli, či ich príspevok ku vzdelávaniu žiakov bol počas štyroch rokov štúdia nad úrovňou očakávania, v súlade s očakávaním alebo, naopak, pod úrovňou očakávania.

Pri výpočte pridanej hodnoty vo vzdelávaní NÚCEM spracováva úspešnosť žiakov danej školy zo slovenského jazyka a literatúry v Testovaní 9 a následne úspešnosť tých istých žiakov v externej časti maturitnej skúšky. Veľkosť pridanej hodnoty predstavuje, rozdiel medzi reálnou a očakávanou priemernou úspešnosťou školy v externej časti maturity. Je v nej teda zohľadnené, s akým vedomostným základom boli žiaci prijatí na strednú školu. Väčšiu pridanú hodnotu môže vykazovať napr. aj škola, ktorá nedosahuje najlepšie výsledky v externých maturitách, pretože prijala žiakov s nižšou úspešnosťou v teste zo slovenského jazyka a literatúry v Testovaní 9. Výsledok nad úrovňou očakávania zdokladuje, že práve táto škola dokázala nadpriemerne zlepšiť študijné výsledky žiakov počas štyroch rokov.

Pri tvorbe modelu a vyhodnocovaní pridanej hodnoty sú odlišované štvorročné gymnáziá od ostatných stredných škôl so študijnými odbormi. Výskumníci NÚCEM-u spracovávajú pri výpočte archívne dáta za tri ročníky absolventov, ktorí ukončili strednú školu v danom školskom roku a dva roky dozadu. Školy teda dostávajú informáciu o ich pridanej hodnote vo vzdelávaní za 3-ročné obdobie.

Pridaná hodnota vo vzdelávaní je považovaná za ďalšie nezávislé a objektívne hodnotenie, ktoré presnejšie vypovedá o kvalite školami poskytovaného vzdelávania. Pridaná hodnota by mala slúžiť riaditeľom škôl najmä na autoevalvačné účely. Školy si môžu samé vyhodnotiť, či vytvorili študentom dostatočné podmienky pre zlepšovanie, t.j. či bola výstupná úroveň žiakov na očakávanej úrovni vzhľadom na ich vstupné výsledky, či naopak sa škole nepodarilo dostatočne rozvinúť vstupný potenciál žiakov. Výpočet pridanej hodnoty sa na stredných školách vzťahuje len na slovenský jazyk a literatúru, keďže len v tomto predmete sa na Slovensku robia paralelné externé testovania deviatakov a zároveň všetkých maturantov. Informáciu o pridanej hodnote v tomto roku dostane 535 stredných škôl (z toho 165 gymnázií a 370 stredných odborných škôl), čo tvorí vyše 70% celkového počtu škôl s maturitnými ročníkmi. Kvôli korektnosti štatistického spracovania výskumníci zaraďujú

do výpočtu pridanej hodnoty iba školy s minimálne 11 maturantmi, ktorí zároveň absolvovali Testovanie 9 pred 4 rokmi a neboli zdravotne znevýhodnení v čase písania oboch testov.

Vyššie spomínaný systém hodnotenia vzdelávania na troch stupňoch vzdelávania predpokladá výpočet pridanej hodnoty vo vzdelávaní aj pre základné školy. NÚCEM pripravuje model pre tento výstup a v roku 2020, keď sa po prvýkrát Testovania 9 zúčastnia žiaci, ktorí v novembri 2015 absolvovali Testovanie 5, bude môcť základným školám zaslať informácie o ich pridanej hodnote vo vzdelávaní aj z slovenského jazyka a literatúry aj z matematiky. Tým sa naplnia dlhodobé ciele NÚCEM-u v problematike poskytovania objektívnej spätnej väzby o stave vzdelávania na Slovensku.

1.6 Problémy, ktoré súvisia s realizáciou externých plošných testovaní

V súvislosti s realizáciou centrálnych, tzv. certifikačných meraní je nutné konštatovať, že testovania takéhoto rozsahu vyžadujú vysoké nasadenie tak zo strany realizátorov a spolupracujúcich organizácií, ako aj zo strany škôl, a pre ich bezproblémový priebeh musia byť vytvorené vhodné legislatívne, personálne a ekonomické podmienky.

Ako sa vyvíjali a menili projekty jednotlivých plošných testovaní na Slovensku za obdobie od roku 2000, menil sa aj charakter s nimi súvisiacich problémov. Spočiatku každé plošné testovanie bolo vnímané ako zásah do kompetencií školy a kontrola jej činnosti, bola tu, do istej miery opodstatnená, obava, že za zlé výsledky bude škola kritizovaná, prípadne budú voči vedeniu školy, prípadne učiteľom vyvodzované dôsledky. Postupom času však školy a učitelia pochopili význam testovaní ako objektívnej spätnej väzby a v súčasnosti pozitívne prijímajú možnosť porovnať vedomosti svojich žiakov so žiakmi iných škôl. Na základe výsledkov externých testovaní tak školy môžu prehodnotiť svoje vzdelávacie ciele a metódy, analyzovať príčiny prípadného zlyhania žiakov v niektorých testovaných kompetenciách či tematických oblastiach a skvalitňovať vzdelávací proces na škole. Aj v súčasnosti sú testovania vnímané ako kontrolný nástroj, ktorého výstupy môžu slúžiť decíznej sfére pri rozhodovaniach, napr. ohľadom regulácie siete škôl. Avšak prínos, ktorý testovania majú pre školy, vo významnej miere zmenšil obavy z ich realizácie a stali sa bežnou súčasťou školskej praxe.

Vzhľadom na vyššie spomínaný význam testovaní a obavy z nich musel NÚCEM čeliť viacerým pokusom o neoprávnené získanie obsahu testov pred začiatkom samotného testovania. Na základe týchto skúseností NÚCEM postupne inovoval a upravoval procesy prípravy testov a manipulácie s nimi, aby bola v maximálnej miere zabezpečená ich utajenosť. Veľmi dôležitým opatrením bolo rozhodnutie, že maturitné testy budú preberať riaditelia škôl vždy ráno pred samotným testovaním z tzv. distribučného miest.

Ďalšou výzvou pre NÚCEM je zabezpečenie objektivity administrácie testov. Testy na školách v súčasnosti administrujú pedagogickí zamestnanci týchto škôl. Pre zvýšenie objektivity však v každej triede (v prípade Testovania 5 a Testovania 9) je prítomný aj externý dozor – učiteľ inej školy. Napriek tomu každoročne NÚCEM spracováva podnety o neobjektívnej administrácii testov na školách, najmä (ale nie len) počas externej maturity. V prípade externej maturity sú za externý dozor vnímaní predsedovia predmetovej a školskej maturitnej komisie, ktorí však nevedia reálne vykonávať dozor v každej testovacej skupine. Preto NÚCEM žiada legislatívne zakotvenie povinnosti externého dozoru v každej triede.

Rieši sa aj problém zabezpečenia vyššej objektivity hodnotenia externej časti maturitnej skúšky. Nutné by bolo centrálné hodnotenie úloh s krátkou odpoveďou, čo však bude vyžadovať navýšenie rozpočtu na projekt externej maturity. Súčasne je potrebné prehodnotiť legislatívny rámec externej maturitnej skúšky – počet predmetov s externou časťou, úroveň,

termíny realizácie a spracovania výsledkov, dobrovoľná maturitná skúška a pod. Medzi problémy, ktoré sú pre NÚCEM výzvou, patrí organizačno-technické zabezpečenie elektronickej maturity vo väčšom rozsahu.

Aj vzhľadom na vyššie spomínané problémy, na Slovensku pretrváva nedostatočná akceptácia výsledkov externej maturity vysokými školami pri prijímacom konaní.

Problémom iného charakteru, ale nie menej závažným, v súvislosti s externým testovaním a meraním v oblasti vzdelávania, je slabá podpora zo strany vysokých škôl. Mnohé pracoviská pripravujúce budúcich učiteľov, až na niekoľko výnimiek, tejto problematike nevenujú podľa nás dostatočnú pozornosť. V ich študijných programoch nie sú zahrnuté semináre, prednášky alebo cvičenia zaoberajúce sa edukometriou, mnohí absolventi opúšťajú pôdu vysokej školy bez dostatočnej informácie o význame a súčasnom stave merania v oblasti vzdelávania na Slovensku. NÚCEM vyvíja rôzne aktivity za účelom informovanosti a publicity ohľadom externých testovaní, prezentuje svoje výstupy na webovej stránke, ale aj na rôznych odborných aj verejných podujatiach. Tiež apeluje na kompetentných zamestnancov vysokých škôl, aby svojich študentov pripravovali na ich budúce povolanie zohľadňujúc súčasný stav a trendy v slovenskom školstve a v kontexte realizovaných národných a medzinárodných meraní.

2 ELEKTRONICKÉ TESTOVANIE NA SLOVENSKU, SYSTÉM E-TEST

2.1 Systém e-Test

V rámci projektu spolufinancovaného z prostriedkov EÚ „Zvyšovanie kvality vzdelávania na základných a stredných školách s využitím elektronického testovania“ (projekt E-test) bola vytvorená elektronická databáza úloh a testov, ktorá disponuje úlohami a testami zo všeobecnovzdelávacích predmetov v súlade so Štátnym vzdelávacím programom a tiež úlohami a testami z gramotností (čitateľskej, matematickej, prírodovednej, finančnej, štatistickej). Vytvorila sa tak možnosť realizovať jednak sumatívne hodnotenie úrovne vzdelávania na výstupe pre vzdelávací stupeň ISCED 2 a ISCED 3, ako aj formatívne hodnotenie úrovne vedomostí žiakov v priebehu školského roka. Do pedagogickej praxe bol zavedený systém elektronického testovania, ktorý slúži na monitorovanie úrovne vedomostí, zručností a kompetencií žiakov, využívaný priamo vo vyučovacom procese a poskytuje priamu spätnú väzbu učiteľom a decíznej sfére.

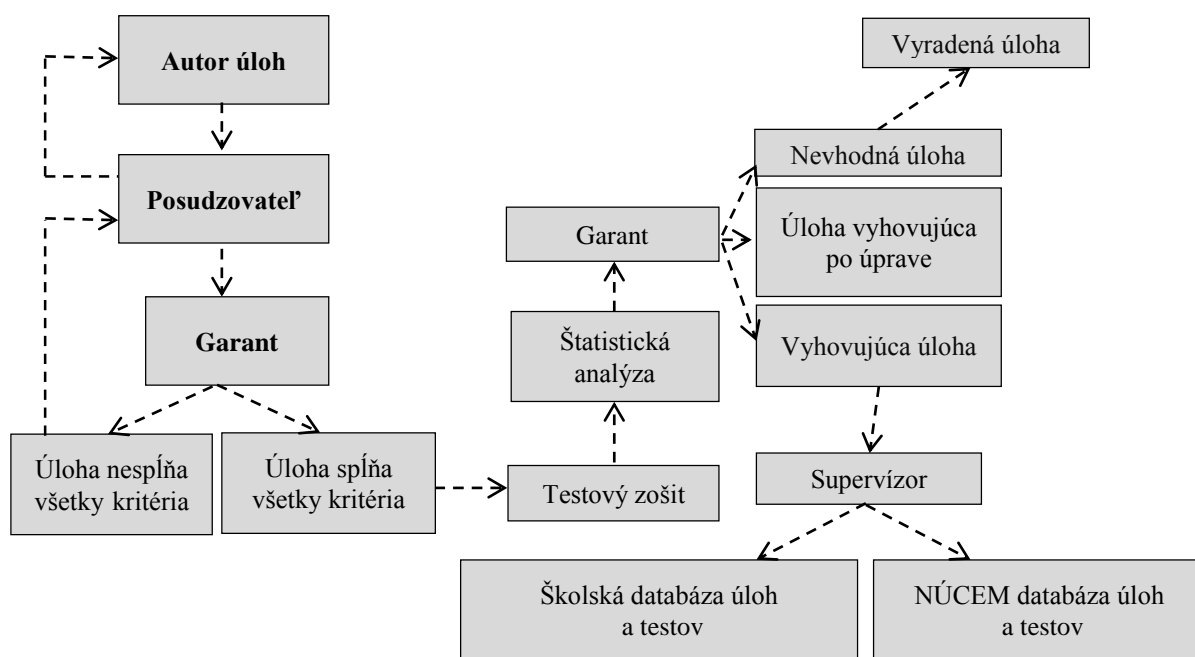
Naplnenie cieľov projektu bolo realizované prostredníctvom dvoch aktivít. V rámci Aktivity 1. bola realizovaná tvorba testovacích nástrojov, zabezpečenie pilotných testovaní, recenzie, hodnotenie a štatistické spracovanie výsledkov, reporty školám, vzdelávacie aktivity. V rámci Aktivity 2. bol vytvorený nástroj, organizačné, kvalifikačné a technické podmienky pre zavedenie elektronického testovania, vrátane overenia elektronického certifikačného testovania, bol zrealizovaný nákup licencií pre zapojené školy, žiakov a používateľov systému, zabezpečené technické vybavenie zapojených škôl, vyškolenie školského personálu.

Vytvorený elektronický systém e-Test disponuje interným portálom, ktorý je primárne určený pre autorov úloh a testov a taktiež pre personál, ktorý zabezpečuje technickú podporu e- testovania a externým portálom, ktorý je určený pre potreby učiteľov, riaditeľov, školských administrátorov, žiakov a rodičov. Dáva možnosť otestovať veľké množstvo žiakov; zabezpečuje efektívnu komunikáciu so školami pred realizáciou testovania; zabezpečuje rozmanitosť vytváraných testových úloh; možnosť vytvárania rôznorodých a moderných testovacích nástrojov; možnosť zistiť množstvo dát o testovaných žiakoch; možnosť efektívneho spracovania výsledkov testovania pre výskum. Významným benefitom systému

je okamžitá informácia o predbežnom výsledku testu zobrazená žiakovi ihneď po ukončení testovania.

V rámci systému e-Test je možné realizovať Tvorbu úloh, ktorá zahŕňa samotnú tvorbu úlohy, jej overenie a poskytnutie reportu o procese tvorby úlohy. Banka úloh predstavuje súbor úloh, ktoré prešli vyššie uvedenými krokmi a boli zaradené do školskej banky (úlohy dostupné v systéme pre učiteľa) alebo do NÚCEM banky (úlohy pripravené na certifikačné testovanie). Systém umožňuje zostavenie testov, ktoré môžu mať školský alebo certifikačný charakter (líšia sa technickými prostriedkami administrácie a niektorými procesmi počas tvorby úloh a testov). Realizácia testovania je možná prostredníctvom školy, kde si testy zostavuje samotný učiteľ (učiteľské testovanie) alebo NÚCEM-om (školské testovanie), kde sa dajú zostaviť testy, testové zošity a dotazníky. Významnú súčasť systému tvorí modul Vyhodnocovanie testov, ktorý poskytuje štatistické spracovanie a vyhodnotenie príslušných parametrov úlohy klasickými štatistickými postupmi (CTT) aj IRT metódou, štatistické výstupy (vytvorenie a zasielanie reportov školám do konta riaditeľa školy) a spätnú väzbu testovaným žiakom a učiteľom (do ich osobného konta na portáli).

Proces tvorby úloh, od vytvorenia úlohy až po jej záverečné zaradenie do niektorej z príslušných bánk, znázorňuje Obr. 5.



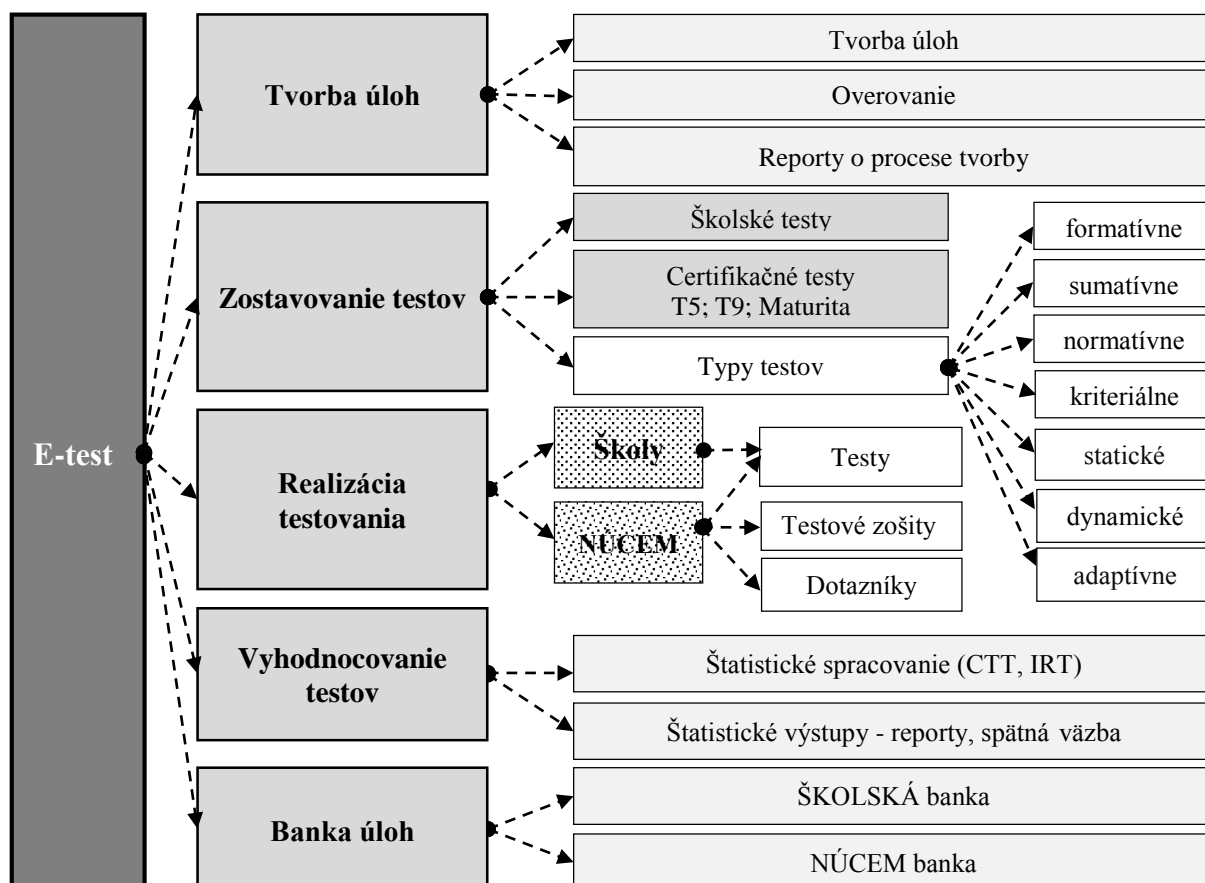
Obr. 5 Schéma procesu tvorby úloh.

Z úloh sú vytvárané súbory úloh, ktoré sú zaradované do testových zošitov, ktoré sa dostávajú priamo k žiakom, a na základe vyhodnotenia úspešnosti riešenia úloh v testových zošitoch možno každej úlohe priradiť štatistické parametre, pomocou ktorých možno identifikovať kvalitu úlohy z pohľadu rôznych štatistických parametrov. Okrem pilotovania úloh sa v projekte realizovala aj tvorba testov, ktoré boli následne štatisticky vyhodnocované a posudzované ako kompaktný celok.

V rámci projektu bolo v e-Teste vytvorených 44 246 úloh a 72 testov z rôznych všeobecno-vzdelávacích predmetov na nižšom a vyššom sekundárnom stupni vzdelávania.

Vytvorené úlohy boli použité na pilotovanie položiek a testy na školské testovanie. V projekte boli s využitím elektronického nástroja realizované aj celoslovenské, tzv. certifikačné merania piatakov, deviatakov a maturantov. V rámci projektu bolo tiež realizovaných päť vln testovaní z predmetov matematika, slovenský jazyka a literatúra, slovenský jazyka a slovenská literatúra a maďarský jazyk a literatúry.

Schematické znázornenie využiteľnosti potenciálu vytvoreného elektronického systému prezentuje Obr. 6.



Obr. 6 Schéma využiteľnosti potenciálu elektronického systému e-Test.

Systém e-Test umožňuje vytvárať úlohy otvorené a uzavreté. Otvorené úlohy sú charakteristické krátkou formou odpovede a uzavreté úlohy sú charakteristické výberom z možností. V elektronickom systéme e-Test pri úlohách s krátkou odpoveďou možno pracovať s tromi typmi úloh a to - úlohy s krátkou odpoveďou (Fill); úlohy s doplnením odpovede/odpovedí do textu (Custom fill); úlohy s odovzdaním riešenia v súbore (File). Pri uzavretých úlohách ide o ponuku deviatich typov úloh a to - úlohy s výberom jednej správnej odpovede z ponúkaných možností (Single choice); úlohy s výberom viacerých správnych odpovedí z ponúkaných možností (Multiple choice); úlohy s výberom jednej správnej odpovede v riadku (Single matrix); úlohy s výberom viacerých správnych odpovedí v riadku (Multiple matrix); úlohy zoradovacie (Ordering); úlohy dichotomické (True/False); úlohy s označením odpovede v texte (Marking text); úlohy umiestňovacie (Drag and Drop); úlohy s označením odpovede v objekte (Hotspot). V nasledujúcej časti uvádzame ukážky jednotlivých typov úloh.

Marek a Tomáš bežia na súťaži tratí „Päťdesiatka pre zdravie“. Marek sa aktuálne nachádza na prvej zastávke so zdravotnou službou. Tomáš je na štvrtej zastávke s občerstvením od štartu.

Profil trate - „Päťdesiatka pre zdravie“

Nadmožná výška (m)

Vzdialenosť (km)

LEGENDA: ■ občerstvenie ● zdravotná služba

2. Na ktorú zastávku sa dostane Marek, ak zabehne ešte ďalších 15 kilometrov?

na piatu zastávku s občerstvením
 na druhú zastávku s občerstvením
 na štvrtú zastávku s občerstvením
 na tretiu zastávku s občerstvením

Obr. 7 Úloha s výberom jednej správnej odpovede z ponúkaných možností (Single choice)

1. Z uvedených typov trojuholníkov označte všetky, v ktorých platí kosínusová veta $a^2 = b^2 + c^2 - 2 \cdot b \cdot c \cdot \cos \alpha$.

pravouhlý trojuholník rovnostranný trojuholník tupouhlý trojuholník rovnoramenný trojuholník

Obr. 8 Úloha s výberom viacerých správnych odpovedí z ponúkaných možností (Multiple choice)

7. Čísla v trojuholníkoch na obrázku označujú obsahy týchto trojuholníkov. Vypočítajte obsah trojuholníka, ktorý je na obrázku označený ako §.

Odpoveď:

Obr. 9 Úloha s krátkou odpoveďou (Fill)

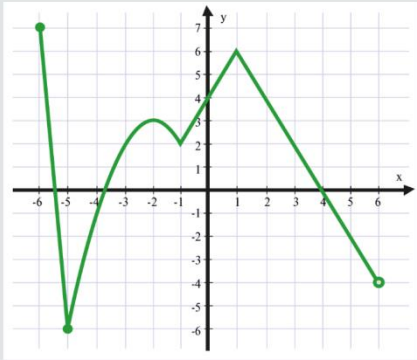
6. Vo vrecku je 13 bielych, 8 modrých a 9 červených guľôčok. Pravdepodobnosť vytiahnutia bielej guľôčky označme p_1 , pravdepodobnosť vytiahnutia modrej guľôčky označme p_2 a pravdepodobnosť vytiahnutia červenej guľôčky označme p_3 .

Ako treba zmeniť počty guľôčok jednotlivých farieb tak, aby sa celkový počet guľôčok nezmenil a aby platilo $p_1 : p_2 : p_3 = 3 : 2 : 1$?

Vo vrecku by malo byť bielych guľôčok, modrých guľôčok
a červených guľôčok.

Obr. 10 Úloha s doplnením odpovede/odpovedí do textu (Custom fill)

1. Pomocou grafu funkcie f zobrazenej na obrázku označte k tvrdeniam 1 až 4 správnu číselnú hodnotu spomedzi uvedených možností, tak aby všetky tvrdenia boli pravdivé.



	3	-5	-4	7	2	1
3) Hodnota funkcie f v bode -3 sa rovná číslu.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2) Hodnota $x = -2$ prislúcha funkčnej hodnote rovnjej číslu.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1) Funkčná hodnota rovná číslu 6 prislúcha x -ovej hodnote.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4) Hodnota funkcie f rovná číslu -6 prislúcha x -ovej hodnote.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Obr. 11 Úloha s výberom jednej správnej odpovede v riadku (Single matrix)

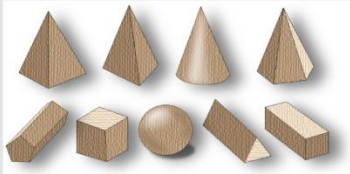
1. Daný je štatistický súbor s hodnotami: 2, 2, 2, 2, 3, 3, 3, 4, 4, 5.

Označte hodnotu modusu a mediánu daného súboru.

	2	2,5	3	3,5	4	4,5	5
modus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
medián	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Obr. 12 Úloha s výberom viacerých správnych odpovedí v riadku (Multiple matrix)

7. Janko sa rozhodol, že telesá v detskej stavebnici správne pomenuje a zoradí ich podľa počtu hrán. Usporiadajte dané telesá vzostupne od telesa s najmenším počtom hrán po teleso s najväčším počtom hrán.



- Kužeľ
- 5-boký ihlan
- Kocka
- 3-boký ihlan
- 3-boký hranol
- 5-boký hranol

Na úlohu nechcem odpovedať

Obr. 13 Úloha zorad'ovacia (Ordering)

1. Máme dané 4 trojuholníky, trojuholník ABC , trojuholník KLM , trojuholník PQR a trojuholník STU . V každom z trojuholníkov poznáme veľkosti jeho dvoch vnútorných uhlov.

- ΔABC : 55° ; 33°
- ΔKLM : 33° ; 85°
- ΔPQR : 92° ; 33°
- ΔSTU : 33° ; 70°

Spomedzi daných trojuholníkov označte všetky tie, ktoré nie sú podobné so žiadnym zo zvyšných trojuholníkov.

ΔABC ΔKLM ΔPQR ΔSTU

Obr. 14 Úloha s označením odpovede v texte (Marking text)

10. Priradte počet percent k obrázku, na ktorom je práve takáto časť vyznačená červenou farbou.

Priestor pre nepoužitú odpoveď

Obr. 15 Úloha umiestňovacia (Drag and Drop)

1. Na hranici dedín A a B plánujú postaviť spoločné termálne kúpalisko. V referende zisťovali postoj obyvateľov oboch dedín. V dedine A s počtom obyvateľov 2 400, oprávnení voliči tvoria 75 % obyvateľstva. V dedine B s počtom obyvateľov 3500, oprávnení voliči tvoria 80 % obyvateľstva. Výsledok referenda vyjadrujú kruhové diagramy:

Dedina A : Dedina B :

súhlasí
 je mu jedno
 nezúčastnil sa
 nesúhlasí

Určte pravdivosť nasledovných výrokov:

1. S výstavbou termálneho kúpaliska nesúhlasí celkovo 18 % oprávnených voličov.
2. Z dediny A viac oprávnených voličov sa nezúčastnilo referenda ako z dediny B .

S výstavbou termálneho kúpaliska nesúhlasí celkovo 18 % oprávnených voličov.
 Z dediny A viac oprávnených voličov sa nezúčastnilo referenda ako z dediny B .

Obr. 16 Úloha dichotomická/Výber jednej správnej z dvoch ponúkaných možností/tvrdení (True/False)

1. V živote sa na každom kroku stretávame s reklamou, ktorá sa nás snaží upútať. Aj banky prezentujú svoje produkty prostredníctvom reklamy. Ponúkajú nám letáčky so základnými informáciami o svojich produktoch, kde sú samé priaznivé informácie. My však musíme vedieť tieto informácie vyhodnotiť. Spoločnosť s názvom Finančná sporiteľňa ponúka pre klientov úver. Označte v texte jej reklamy informáciu, ktorá pre klienta nie je prínosná - podstatná.

Obr. 17 Úloha s označením odpovede v objekte (Hotspot)

Výstupy projektu E-test mali rozmanitý charakter v súlade s podrobným opisom projektu a stanovenými cieľmi. Všetky dosiahnuté výstupy projektu boli podrobne prezentované v rámci záverečnej konferencie pod názvom „Elektronické testovanie – skúsenosti a výzvy“, ktorá sa konala v Bratislave 20. – 21. októbra 2015. Medzi tie najpodstatnejšie patrí banka úloha a testov, počet zrealizovaných e-testovaní, počet škôl zapojených do realizácie aktivít projektu, publikácie, počet realizovaných vzdelávaní a pod. Podrobný prehľad výstupov projektu ponúka nasledujúca Tab.2.

Tab. 2 Prehľad výstupov projektu E-test

Tvorba úloh a testov (pilotné testovania a školské testovania) 2013 – 2015		
Vzdelávacia oblasť	Počet vytvorených úloh	Počet vytvorených testov
Jazyk a komunikácia (SJL, MJL, SJSL, AJ, NJ, RJ)	20 363	22
Matematika a práca s informáciami (MAT, INF)	7 100	6
Človek a príroda (BIO, CHE, FYZ)	4 295	12
Človek a spoločnosť (DEJ, GEO, OBN)	5 292	3
Gramotnosti (ČG, MG, FG, ŠG, PG)	7 196	29
Spolu	44 246	72
Vzdelávania 2013 - 2015		
Školenia		56
Workshopy (Letná škola)		6
Regionálne kontaktné centrá		41
Výsledky projektu 2013 - 2015		
Počet žiakov/študentov zapojených do realizácie aktivít projektu		142 660
Počet vykonaných testovaní		341 048
Počet škôl zapojených do realizácie aktivít projektu		1 649
Počet publikácií		11

Vytvorené publikácie sú dostupné v elektronickej forme na <http://www.etest.sk/publikacie/>.

2.2 Elektronická forma maturity a e-testovania na základných školách

Obrovský rozmach digitálnych technológií zasahuje vo významnej miere aj do procesu vzdelávania. V súčasnosti na všetkých slovenských školách sa v rozličnej miere využívajú počítače, laptopy a v poslednom období aj mobilné technológie nielen pri vzájomnej komunikácii medzi žiakmi, rodičmi a učiteľmi. Do škôl pribudli digitálne tabule a všetky tieto technológie učitelia priamo využívajú v rámci prípravy na hodinu, ale aj počas samotného vyučovania. Prirodzene sa elektronizácii nevyhne ani hodnotenie žiakov a testovanie.

Podobne, ako aj v mnohých iných krajinách Európy, sa na Slovensku už v roku 2006 začalo diskutovať o možnosti elektronického (on-line) testovania. V roku 2008 bol spustený testovací model on-line maturity na vybraných 11 školách, pričom sa otestovalo 215 žiakov. V roku 2009 prebehla maturita on-line z predmetov anglický jazyk a matematika na vzorke 791 žiakov. V školskom roku 2009/2010 sa realizovala on-line maturita z matematiky a anglického jazyka na 37 školách a otestovaných bolo 1 452 žiakov. Popri testoch sa k projektu on-line testovania vyjadrili učitelia a žiaci formou dotazníkov. Z dotazníkov vyplynulo, že žiaci i učitelia pozitívne hodnotia on-line formu testovania a tiež aj nižšiu logistickú náročnosť tejto formy testovania, zo strany realizátora bola ako pozitívum vnímaná väčšia objektivita a efektivita vyhodnocovania testov. Medzi negatívne stránky tohto projektu boli zaradené vysoká finančná náročnosť bezpečnostného riešenia on-line testovania, nedostatočné technické vybavenie a nevhodné internetové pripojenie na školách. Aj vzhľadom na uvedené technické obmedzenia a vysokú finančnú náročnosť sa v tomto projekte ďalej nepokračovalo. NÚCEM však vnímal veľké výhody využitia digitálnych technológií v procesoch tvorby testovacích nástrojov, samotného testovania a hodnotenia. V rámci projektu „Hodnotenie kvality vzdelávania na základných a stredných školách SR v kontexte prebiehajúcej obsahovej reformy vzdelávania“ boli vytvorené základy elektronickej banky úloh, monitorovali a analyzovali sa možnosti ďalšieho využitia technológií v činnosti NÚCEM-u.

Dôležitým podnetom pre obnovenie konceptu elektronického testovania bol plánovaný a neskôr aj zrealizovaný prechod na elektronickú formu testovania v medzinárodnej štúdii PISA 2015. NÚCEM obnovil úmysel realizovať centrálné administrované plošné testy elektronicke a pretavil ho do zámeru národného projektu E-Test koncom roku 2012. Následne, ako bolo vyššie uvedené, sa v rámci projektu E-Test vytvoril, overoval a neskôr aj v praxi začal využívať elektronický testovací systém e-Test.

Ako uvádzajú Tab.2 a Tab. 3, počas realizácie projektu sa uskutočnilo veľké množstvo pilotných a školských testovaní. Zároveň sa v prostredí e-Testu zrealizovali aj národné tzv. certifikačné testovania a ich generálne skúšky E-Testovanie 5-2014, E-Maturita 2015, E-Testovanie 9-2015 a E-Testovanie 5-2015. Do týchto aktivít sa zapojilo celkom 1 100 základných a stredných škôl. Záujem o elektronickú formu testovania prejavujú vo väčšej miere stredné školy. Zo všetkých stredných škôl s maturitnými ročníkmi sa riadneho termínu E-Maturity 2016 zúčastnilo vyše 15 % škôl. Elektronickú formu maturity z vyučovacích jazykov si volilo približne 4% žiakov, oveľa vyšší záujem bol o elektronickú formu maturity z matematiky (12 % všetkých maturantov z matematiky). Záujem základných škôl o E-Testovanie 9-2016 bol na úrovni 5 %. Vzhľadom na to, že išlo o novú formu testovania a účasť žiakov na e-forme testovania bola dobrovoľná, je počet zapojených škôl a žiakov dobrým znamením, že moderné formy hodnotenia výsledkov vzdelávania majú na Slovensku perspektívu.

Toto tvrdenie potvrdzuje aj fakt, že aj po ukončení projektu E-test naďalej pretrváva záujem o jeho využitie zo strany škôl. V roku 2016 sa uskutočnil ďalší ročník E-Maturity a E-Testovania 9. Počet žiakov a škôl zapojených do jednotlivých certifikačných e-testovaní

prezentuje Tab. 3. Okrem certifikačných testovaní NÚCEM ponúkol v máji a júni 2016 školám školské e-testovania z rôznych predmetov. Túto možnosť otestovať sa využilo 346 škôl a otestovalo sa 16 141 žiakov, niektorí aj vo viacerých testoch, čo predstavovalo 34 069 zrealizovaných e-testov. NÚCEM už od septembra ponúka v systéme e-test ďalšie školské testy, celkom je ich pre ISCED 2 a ISCED 3 pripravených 29. V novembri NÚCEM ponúkne základným školám popri papierovej forme aj elektronickú formu Testovania 5.

Tab. 3 Prehľad certifikačných testovaní v systéme e-Test

Certifikačné testovania	Počet testovaných žiakov		Počet zapojených škôl	
	2015	2016	2015	2016
E-Testovanie 5	2 232	-	111	-
E-Testovanie 9	1 700	1 209	111	75
E-Maturita	3 495	2 317	176	115

ZÁVER

V súčasnej dobe je význam celonárodného merania vo vzdelávaní nespochybniteľný. Národné testovania žiakov ako prostriedok na meranie a monitorovanie vzdelávania sa stali na Slovensku dôležitou súčasťou vzdelávacej politiky a tvoria pomocný, ale zároveň aj veľmi účinným nástrojom na skvalitňovanie vzdelávania. Nezastupiteľný význam celoplošných testovaní možno vnímať z niekoľkých pohľadov. V prvom rade garantujú validitu a objektivitu hodnotenia dosiahnutých výsledkov vzdelávania – zadávaním štandardizovaných, externe pripravovaných a vyhodnocovaných testov, administrovaných v tom istom čase na všetkých školách pri zachovaní rovnakých podmienok a harmonogramu testovania. Zároveň štandardizujú a zjednocujú požiadavky na vedomosti a zručnosti absolventov základných a stredných škôl. Zvyšujú zodpovednosť škôl za výsledky ich vzdelávania a tým prispievajú k skvalitňovaniu výchovno-vzdelávacieho procesu. Externá maturita overuje a hodnotí tie vedomosti a zručnosti žiakov, ktoré nie je možné overiť v dostatočnej miere v ústnej forme internej časti – zručnosti a kľúčové kompetencie ako počúvanie s porozumením (v cudzích jazykoch), čítanie s porozumením, gramatika v kontexte, schopnosť prezentovať vlastný prejav písomnou formou (v cudzích aj vyučovacích jazykoch), práca s grafmi, tabuľkami, aplikácia poznatkov v reálnom kontexte a matematická gramotnosť (matematika).

Možno teda konštatovať, že celonárodné externé testovania si obhájili svoje miesto vo vzdelávacom systéme SR. V priebehu ostatných rokov sa vyladili jednotlivé kroky organizačno-technického zabezpečenia týchto celonárodných projektov. Na základe skúseností získaných niekoľkoročnou realizáciou externého testovania žiakov na výstupe z ISCED 2 a ISCED 3 NÚCEM pripravil a začal realizovať projekt monitorovania vzdelávacích výsledkov žiakov na výstupe z ISCED 1. V minulom školskom roku sa úspešne uskutočnilo celoslovenské testovanie žiakov 5. ročníka základných škôl a 23. novembra tohto roku sa po druhýkrát na Slovensku uskutoční na všetkých základných školách.

Celonárodné testovania kladú vysoké nároky na organizačné zabezpečenie a sú finančne náročné. Výskum v oblasti hodnotenia vo vzdelávaní si vyžaduje neustálu finančnú podporu. Preto NÚCEM využil možnosti financovania prostredníctvom fondov EÚ a vyššie spomínané národné projekty napomohli k významnému posunu v oblasti hodnotenia vo vzdelávaní na Slovensku. V prvom projekte NÚCEM overil a následne odštartoval trojstupňový model národných testov na výstupe z troch stupňov vzdelávania a hodnotenie rozšíril o ďalšiu objektívnu informáciu o kvalite vzdelávania vo forme pridanej hodnoty vo vzdelávaní.

Projekt E-Test na Slovensku urobil dôležitý krok smerom k modernizácii školstva. Vytvorený elektronický testovací systém e-Test môže napomôcť k väčšej motivácii žiakov, keďže už aj proces hodnotenia a skúšania vie byť realizovaný v im blízkej a z bežného života známej elektronickej forme. Rovnako významným krokom v modernizácii je vytvorenie elektronickej banky úloh zameraných na aplikáciu a prepájanie získaných vedomostí z celého spektra vyučovacích predmetov. Zo strany učiteľov je po nich veľký záujem, pretože úloh z čitateľskej, finančnej, matematickej, štatistickej či prírodovednej gramotnosti nie je dostatok. Za nezanedbateľný prínos oboch projektov je možno považovať aj vyškolenie tisícok učiteľov na tvorbu úloh z rôznych predmetov a gramotností a tiež aj tvorbu úloh a prípravu testov v elektronickom prostredí e-Testu, ktorými určite obohatia vyučovací proces na svojej škole aj v neskoršom období.

V súčasnom období, po skončení projektu, je pozornosť NÚCEM-u zameraná na ďalšie využitie jeho výstupov v pedagogickej praxi. V systéme e-Test sa aj dnes realizujú školské a certifikačné testovania v rôznych fázach školského roka, čo dokazuje potrebu centrálného testovania na rôznych úrovniach vzdelávania. Je potrebné zabezpečiť, aby sa tento vytvorený živý organizmus ďalej vyvíjal a stal sa pre učiteľov novým, moderným prostriedkom na získanie objektívnych informácií o efektívite vzdelávacieho procesu.

Cieľom NÚCEM-u je neustále sledovať trendy v oblasti hodnotenia vo vzdelávaní, realizovať výskum, ktorého výstupom sú nové postupy a metódy vedúce k zefektívňovaniu jednotlivých organizačno-technických fáz realizácie národných testovaní a k skvalitňovaniu testovacích nástrojov. Priestor pre ďalší vývoj a skvalitňovanie je aj v oblasti centralizovaného hodnotenia otvorených úloh, čím by sa významne zvýšila objektivita výsledkov testovaní. Perspektívu vidí NÚCEM v rozšírení systému externých meraní na základných a stredných školách o ďalšie predmety. Na základných školách, či už na výstupe z primárneho stupňa vzdelávania, alebo z nižšieho sekundárneho vzdelávania, absentuje monitorovanie úrovne vzdelávania v cudzích jazykoch. Zároveň vidíme potrebu sledovať úroveň a dosahovanie cieľov vzdelávania z prírodovedných predmetov na výstupe z ISCED 2 a ISCED 3. V rámci svojich výskumných aktivít NÚCEM pokladá za dôležité venovať sa analýze a sledovaniu rôznych faktorov ovplyvňujúcich kvalitu škôl a pokračuje v inovácii modelu pridanej hodnoty vzdelávania, ktorá prepája výsledky testovania v kľúčových bodoch vzdelávania. Rozšírený systém meraní doplnený o ďalšie objektívne informácie o úrovni vzdelávania na školách umožní vo väčšej miere a objektívnejšie hodnotiť kvalitu a monitorovať vývoj vzdelávania na základných a stredných školách a ovplyvňovať strategické rozhodnutia v oblasti vzdelávacej politiky. V neposlednej miere sú analýzy výsledkov externých meraní a všetky ďalšie výstupy NÚCEM-u pre školy významným zdrojom nezávislých informácií nevyhnutných pre ich autoevalváciu a napomôžu k zvyšovaniu kvality vzdelávania na slovenských školách.

LITERATURA

EACEA, Eurydice. *Národné testovanie žiakov v Európe: Ciele, organizácia a využívanie výsledkov (2009)*. Brusel: Eurydice . [online] Dostupné z WWW:

http://eacea.ec.europa.eu/education/eurydice./documents/thematic_reports/109SK.pdf

Kanovská, R. & Pichaničová, I. (2009). Päť rokov externej maturitnej skúšky na Slovensku. In *Vývoj a perspektívy maturitnej skúšky v Slovenskej republike. Zborník z odborného seminára*. Bratislava: NÚCEM – MPC – ŠPÚ.

NÚCEM (2012). *Podrobný opis projektu*. Bratislava: NÚCEM.

Pichaničová, I. (2010). Vývoj a význam externej maturitnej skúšky v Slovenskej republike. In *Pedagogická veda a školská prax v historickom kontexte. Zborník z medzinárodnej vedeckej konferencie*. Trnava: Univerzita sv. Cyrila a Metoda.

Pichaničová, I. (2011). Celonárodné meranie vzdelávania v Slovenskej republike. In *Meranie vedomostí ako súčasť zvyšovania kvality vzdelávania. Zborník príspevkov z konferencie s medzinárodnou účasťou*. Trnava: TYPI UNIVERSITATIS TYRNAVIENSIS.

KONTAKT

PaedDr. Lucia Ficová, PhD.

Národný ústav certifikovaných meraní vzdelávania

Žehrianska 9, 851 07 Bratislava 5

Telefon: +421-2-68 26 03 27

e-mail: lucia.ficova@nucem.sk

PaedDr. Ivana Pichaničová, PhD.

Národný ústav certifikovaných meraní vzdelávania

Žehrianska 9, 851 07 Bratislava 5

Telefon: +421-2-68 26 03 01

e-mail: ivana.pichanicova@nucem.sk

VÝZKUM PARAMETRŮ OVLIVŇUJÍCÍCH POČÍTAČEM PODPOROVANOU PEDAGOGICKOU DIAGNOSTIKU

Josef Malach, Martin Malčík

Abstrakt: Autoři kapitoly se zabývají konceptuálními otázkami přidané hodnoty ve vzdělávání a metodice určení relativního přírůstku znalostí a dovedností žáků. Text přináší přístupy k měření výsledků vzdělávání a jeho očekávané výstupy v podobě osvojených předmětových dovedností. Je představena architektura diagnostického systému TEST, s různými typy úloh a portálem, který tohoto testu využívá. Autoři prezentují souhrnné výsledky několika testovacích kampaní, které komentují a dávají do souvislosti, resp. je srovnávají s výsledky mezinárodního šetření OECD PISA. Vedle výhod testovacího systému poukazují i na jeho zjištěné slabší stránky, které jsou další optimalizací procesu testování postupně eliminovány.

Klíčová slova: testování znalostí a dovedností žáků, přidaná hodnota vzdělávání, relativní přírůstek znalostí, diagnostický systém TEST

Abstract: The authors of the chapter are dealing with conceptual issues of added value in education and methodology of determining the relative gain of students' knowledge and skills. The text brings approaches to measurement of the results of education and its expected outcomes in the form of skills acquired in specific subjects. The architecture of the diagnostic system TEST, with different types of tasks and a portal that uses this test, is presented. The authors present the summary results of several test campaigns that are commented and put into context, respectively it is compared with the results of an international campaign OECD PISA. Apart from the advantages of the test system, also the weaker points, that have become the objective of further optimizing of the testing process and are being gradually eliminated, are identified.

Key words: testing of knowledge and skills, added value of education, relative gain of knowledge, diagnostic system TEST

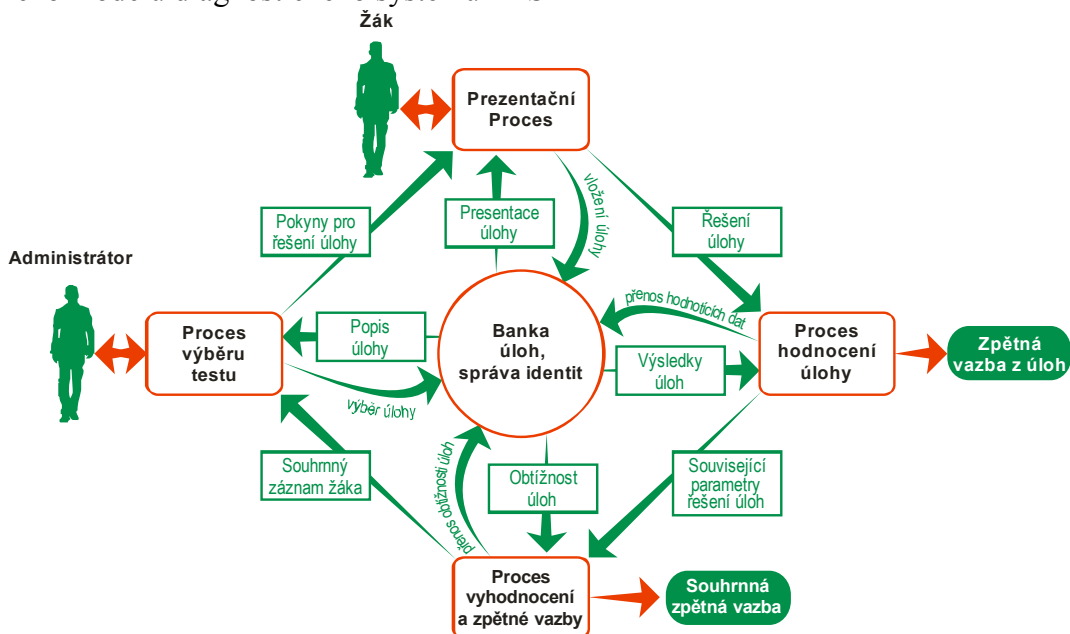
ÚVOD

V této kapitole jsou analyzovány klíčové parametry počítačem podporovaného diagnostického systému a metodiky relativního přírůstku znalostí žáků. Nejprve jsou charakterizovány soubory žáků, které byly pomocí diagnostického systému testovány v letech 2007 až 2011, následně jsou prezentovány metody statistické analýzy a vytěžování dat, které byly použity jak při vývoji metody relativního přírůstku znalostí, tak při vývoji informačního systému. Současně jsou pro porovnání pořadí škol podle velikosti přidané hodnoty ve vzdělávání navrženy koeficienty celkové a absolutní shody.

1 POČÍTAČEM PODPOROVANÝ DIAGNOSTICKÝ SYSTÉM

1.1 Logický model počítačem podporované diagnostiky ve vzdělávání „TEST“

Hodnocení ve vzdělávání se všemi ovlivňujícími faktory představuje složitý systém, který je obtížné modelovat s cílem zkoumání vzájemných vztahů, závislostí, posouzení významnosti a síly faktorů a pro modelování výstupních charakteristik. Obrázek 1 představuje návrh logického modelu diagnostického systému TEST



Obr. 1 Logický model elektronického diagnostického systému TEST

Každý diagnostický systém musí mít – ve své základní podobě – nejméně čtyři různé procesy: proces výběru testu a úloh, prezentační proces, proces testování žáka a proces vyhodnocení a zpětné vazby.

1. *Proces výběru testu* je zodpovědný za výběr úloh z banky úloh. To mohou být úlohy diagnostikující výsledky učení se žáka v kognitivní, psychomotorické i afektivní doméně podle Bloomovy taxonomie. Podle mezinárodních výzkumů představuje pořadí úloh důležitou charakteristiku ovlivňující úspěšnost výsledku žáků, proto certifikované testy jsou sestavovány a pilotovány jako celek a pro zajištění shodných testovacích podmínek v rámci standardizovaného testování by nemělo docházet k „rotaci“ úloh. Výjimku z tohoto pravidla tvoří testy, zpracované pomocí teorie odpovědi na položku (IRT – Item Response Theory) metodikou CAT (Computer Adaptive Testing).
2. *Prezentační proces* zajišťuje zobrazení úloh žákům. *Prezentační proces* musí zabezpečit správné zobrazení úloh různých formátů, včetně obrázků, zvukových záznamů, animací, appletů a dalších vzdělávacích objektů. Když žák vyřeší úlohu, *prezentační proces* zachytí žakovu reakci a pošle informaci jako artefakt dále v rámci komponenty *řešení úlohy*.
3. *Proces hodnocení úlohy* je prvním krokem v procesu hodnocení. Tento proces vyhodnocuje všechny parametry konkrétní odpovědi žáka, které poskytují aktuální informace o vědomostech, dovednostech a schopnostech žáka.

4. *Proces vyhodnocení a zpětné vazby* realizuje souhrnné hodnocení žáka v procesu testování. Proces využívá komponent *řešení úlohy* a *související parametry řešení úloh* k celkovému hodnocení žáka.

Mezi jednotlivými procesy a katalogem úloh existuje komunikace, která zahrnuje celý cyklus hodnocení žáka a ústí v komplexní záznam výsledků žáka. V rámci komunikace mezi procesy jsou přenášeny komponenty: *pokyny pro řešení úloh*, *řešení úlohy*, *související parametry řešení úloh* a *souhrnný záznam žáka*.

1. *Pokyny pro řešení úloh* jsou pokyny zaslané *procesem výběru testu prezentačnímu procesu*. Jedná se o parametry, které jsou součástí úlohy i metadata o úloze, a také identifikátor další úlohy.
2. *Řešení úlohy* jsou různé formy odpovědi žáka v závislosti na úloze. Jedná se o celou škálu možných odpovědí od jednoduchých reakcí na multiple-choice úlohy až po vytvořené artefakty a jiné produkty.
3. *Související parametry řešení úloh* popisují kvalitu hlavních rysů *řešení úlohy*, jako správnost, přesnost, „uměleckost“, a současně, za jakých podmínek byla daná kvalita dosažena, včetně času a doby řešení úlohy.
4. *Souhrnný záznam žáka* jsou kumulované výsledky žákovy způsobilosti v oblasti vědomostí, dovedností a schopností. Všechna tato data dávají dohromady *komplexní záznam žáka*, který obsahuje úlohy, které žák řešil, jaké hodnocení získal a za jakých podmínek úlohy řešil. Současně obsahují základní statistiky výsledného stavu pro *souhrnnou zprávu žáka*. *Souhrnný záznam třídy* a *souhrnný záznam školy* jsou generovány ze *souhrnných záznamů žáků*.

Centrálním prvkem diagnostického systému je *banka úloh*. Zde jsou umístěny kalibrované úlohy (úlohy se známými parametry) s metadaty, jako např. stupeň obtížnosti úlohy, testovaný předmět a vhodný věk žáků, kdo úlohu připravil, pilotoval a kdo ji garantuje, označení kampaně testování, ve kterých byla úloha použita, obsažené dovednosti atd. Z těchto úloh jsou vytvářeny různé varianty testu jak pro klasické testování v rámci teorie CTT, tak pro adaptivní testování v rámci teorie IRT.

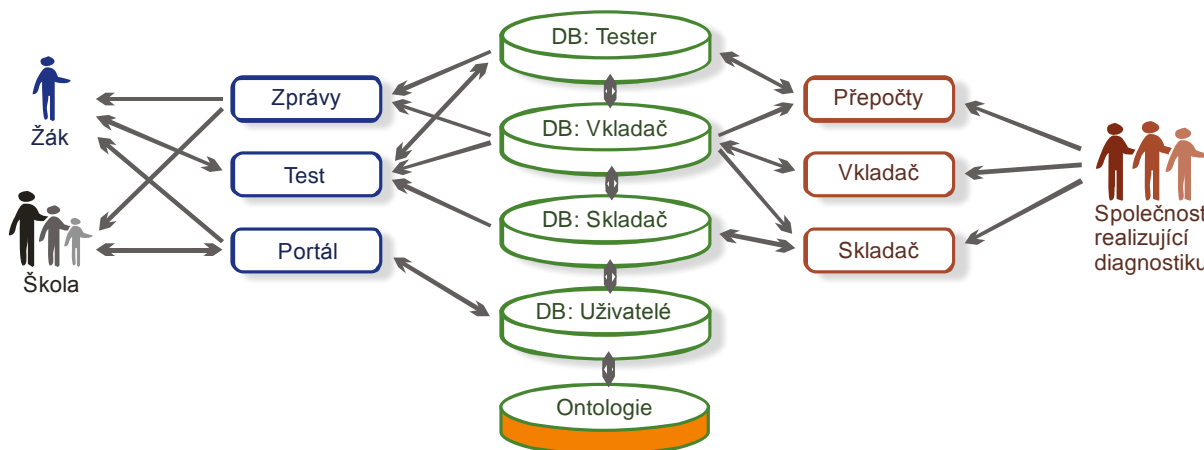
Cíl metodicky správného vytvoření banky úloh spočívá také v nastavení zadání úloh z různých testů na jednu stupnici. Tímto způsobem *banka úloh* představuje množinu testových úloh, jejichž parametry jsou umístěny na obecnou stupnici tak, že různé podmnožiny těchto zadání mají stejný stupeň obtížnosti pro diagnostikovaného žáka (Kardanova, 2008).

1.2 Architektura diagnostického systému – test

Architektura počítačem podporovaného diagnostického systému je znázorněna na obrázku 2 konceptuálním modelem a je kompletně zrealizována prostřednictvím webových technologií a prostředků ve formě lehkého klienta. Z internetového prohlížeče jsou tedy přístupné všechny jeho části ve všech rolích, které v systému existují. Základem diagnostického systému je jeho databáze. V databázi se nachází úlohy, seznamy žáků, výsledky z testování, analýzy, zprávy a různá provozní data diagnostického systému. Aplikační část řešení je rozdělena do dvou bloků. Obslužná část slouží pro správu a údržbu a obvykle ji využívají administrátoři a obsahoví analytici. Zde je možno editovat objekty a taxonomie ontologie, definovat nové prvky a úlohy a nastavovat různé další provozní údaje a parametry. V uživatelské části aplikace se prezentují výstupy všech shromážděných informací a znalostí na základě individuálních potřeb každého uživatele. Vůči těmto prvkům se pak obsahově

porovnávají všechny informace, žáci, úlohy, dovednosti a další informace uložené v datovém úložišti.

Obecný modul je postaven na abstraktní vrstvě (tedy abstraktním ontologickém modelu) a měl by tudíž mít schopnost podporovat a integrovat dohromady rozličné ontologické znalosti reprezentované v různých jazycích. Tento přístup zajistí, že systém bude ještě flexibilnější nejen vůči změnám, ale také v budoucnosti vůči nově přicházejícím standardům v reprezentaci znalostí. Navíc, s ohledem na cílovou funkcionalitu informačního systému, by měl IS také poskytovat odvozování pro abstraktní ontologický model.



Obr. 2 Konceptuální model elektronického diagnostického systému TEST

Testovací systém TEST je složen z portálu a tří dílčích subsystémů – dvou interních (určených pro administrátory, programátory a metodiky projektu) a jednoho externího (testovací rozhraní samotné). Do testovacího systému se nepočítají generátory zpráv, ty jsou samostatnými systémy.

VKLADAČ – Interní systém pro vkládání, editaci a mazání:

- předmětů (a přiřazování otázek do nich);
- otázek (dva typy – běžná a vyškrtávací);
- podotázek otázek (čtyři typy – checkbox, radio, select a input);
- skupin podotázek;
- odpovědí (a jejich správností);
- dovedností (a přiřazování vah dovedností u jednotlivých otázek);
- dalších globálních metadat přiřazených k danému testování.

SKLADAČ

Jedná se o interní systém pro skládání testů, využívá informace z VKLADAČE a slouží pro vytváření souboru konfigurace bloků, částí a dalších metadat, která následně využívá TESTER pro zobrazení testu. Hlavní funkčnosti tohoto subsystému jsou:

- vytváření variant (a přiřazování otázek z vkladače do nich);
- vytváření částí (bloků) testu (např. „Český jazyk – část 1.“);

- každý ročník má své vlastní bloky a varianty;
- vytváření pořadí bloků (jak se zobrazí v testu);
- editace dalších parametrů a metadat, která TESTER využívá pro tvorbu testu;
- časové limity na jednotlivé části testu a velikosti úlev pro žáky se sníženou schopností;
- zapínání/vypínání a editace: úvodů, závěrů bloků, cvičných otázek, upozornění na možnost přerušení testu, zobrazování předběžné vyhodnocovací tabulky a okamžité zprávy, odhady úspěšnosti žáka v předmětech;
- rotování otázek, rotování odpovědí;
- texty nadpisů, úvodu/závěru testování, popisy jednotlivých částí atd.

TESTER

Externí systém, je hlavní aplikací testování, po přijetí požadovaných parametrů složí test pro daného žáka z informací a konfigurací, které čerpá z VKLADAČE a SKLADAČE. Vytváří posloupnost částí (včetně rozhodnutí, zda žákovi zobrazit AJ nebo NJ) a navigaci v testu, vybírá žákovi variantu ze seznamu variant, dále ukládá odpovědi žáka, vyhodnocuje jejich správnost, odhady úspěšnosti, časy strávené v jednotlivých částech a celkový stav (pozici) vyplnění testu (umožňuje tedy zavřít prohlížeč, spustit znova test z portálu a pokračovat od poslední uložené části). TESTER ukládá např. čas poslední aktivity v testu, jméno prohlížeče uživatele a tyto další parametry:

- příjem parametrů;
- výběr variant;
- udržování stavu;
- zobrazování částí a otázek;
- zobrazování podotázek;
- ukládání odpovědí;
- vyhodnocování správnosti odpovědí.

Zobrazení testu žákovi se řídí mezinárodně platnými standardy (Standardy pro pedagogické a psychologické testování, 2001, s. 146 – 151).

Úlohy v testu jsou prezentovány po tematických celcích (sémantický modus), což se dle provedených výzkumů (Norman et al., 2001), ukázalo jako efektivnější jak z hlediska času řešení testu, tak z hlediska úspěšnosti žáků v testu. Řazení úloh podle obtížnosti a podle tematické návaznosti jednotlivých diagnostických znalostí zvyšuje formativní vlastnosti diagnostické metody. Žáci se v sémantickém modu zobrazování úloh mohou po jednotlivých úlohách pohybovat, mohou z řešení jednodušších úloh vyvozovat řešení složitějších úloh, mohou si upevňovat a systematizovat zapamatování jednotlivých pojmů tematického celku.

Portál IS Diagnostic pro systém TEST

Portál je základním vstupním místem do systému TEST. Skládá se z více částí a zpřístupňuje další moduly systému. Přístup k obsahu portálu je řízen rolí uživatele a úrovní autentizace. Pro nepřihlášené uživatele je na portále dostupný pouze základní informační servis a stránka pro registraci nového uživatele. Přístup k ostatním částem portálu je možný pouze

pro registrované uživatele a po přihlášení. Základní registrace je možná přímo na portále a vyžaduje pouze ověření zadané e-mailové adresy. Takto registrovaný uživatel má roli „Veřejnost – registrovaný uživatel“.

Portál IS Diagnostic je prostředí, které slouží jako komunikační rozhraní mezi klienty diagnostického systému a samotným testovacím systémem. Součástí modulu Portál je také jednotná infrastruktura pro autentizaci, autorizaci, monitoring a auditování systému. Toto uživatelské rozhraní lze využívat s registrací a přihlášením nebo i bez něho. Portál vykonává zejména tyto funkcionality:

- vstup pro uživatele bez registrace;
- přihlášení uživatele;
- vstup po registraci a přihlášení.

Registrace školy se realizuje přes PSC, typ školy, kdy jsou na zadaný email z centrální databáze škol zaslány přihlašovací údaje. Pokud email odpovídá, registraci potvrdí, pokud neodpovídá, pak požádá o změnu emailu telefonicky či elektronicky administrativní podporu systému.

Grafické uživatelské rozhraní (GUI) portálu se řídí Pravidly tvorby přístupného webu⁵ a také WCAG (Web Content Accessibility Guidelines).⁶

1.3 Banka úloh

Systém banka úloh je hlavním úložištěm úloh pro celý systém. Databáze s citlivými údaji jsou šifrovány pomocí technologie MS SQL Server Transparent Data Encryption, čímž je zajištěna ochrana před zcizením dat jak zkopírováním z primárního úložiště, tak ze záložních médií. Práva provozních administrátorů databází jsou přidělována na základě pravidla „least privilege“ tak, že pro zajištění provozu databází nebudou nikdy využívány účty s nejvyšším oprávněním. Veškeré bezpečnostně citlivé operace prováděné v databázích (změny v přidělení práv a rolí, vytváření uživatelských účtů, exporty dat, změny systémových parametrů, změny v nastavení šifrování), jsou auditovány a ukládány v odděleném zabezpečeném úložišti pomocí technologie MS SQL Server Auditing.

2 CHARAKTERISTIKA PŘEDMĚTU ZKOUMÁNÍ

Elektronický diagnostický systém byl opakovaně a dlouhodobě používán v následujících hromadných testováních:

1. Testování přidané hodnoty u žáků 1. a 3. ročníků středních škol v Moravskoslezském kraji;
2. Testování žáků základních a středních škol v česko-slovenské příhraniční oblasti;
3. Diagnostika stavu znalostí a dovedností žáků 5. a 9. ročníků v České republice se zaměřením na jejich rozvoj.
4. Rozvoj a podpora kvality ve vzdělávání ve Zlínském kraji 2013 – 2015.

⁵ Pravidla tvorby přístupného webu. [online]. 2011. [cit. 20. 1. 2011] <http://pristupnost.nawebu.cz/texty/pravidla-standardy.php>

⁶ Web Content Accessibility Guidelines. [online]. W3C: New York, 2008 [cit. 30. 1. 2012] <http://www.w3.org/TR/WCAG20/>

V rámci těchto hromadných testování byly ověřovány parametry systému z různých hledisek a různými metodami statistiky a dolování znalostí.

2.1 Testování přidané hodnoty žáků 1. a 3. ročníků středních škol v Moravskoslezském kraji

V Moravskoslezském kraji je od roku 2007 realizován projekt Kvalita 2007 až 2020. Jeho hlavním cílem je testování žáků prvních a třetích ročníků maturitních oborů středních škol v matematice, českém jazyce, anglickém jazyce a německém jazyce metodou relativního přírůstku znalostí.

Data využívaná v této práci byla získána v rámci projektu „*Testování žáků 1. ročníků oborů vzdělávání poskytující střední vzdělávání s maturitní zkouškou v letech 2007 až 2013!* A „*Testování žáků 3. ročníku oborů vzdělání poskytujících střední vzdělání s maturitní zkouškou a zjištění relativního přírůstku znalostí na všech středních školách v Moravskoslezském kraji*“. Sběr dat v 1. ročníku SŠ, viz tabulka 1, probíhá na začátku studia, na podzim, tedy pár týdnů po nástupu žáků na střední školu. Sběr dat v 3. ročníku SŠ, viz tabulka 2, probíhá na jaře konkrétního roku, obvykle na přelomu měsíce května a června. V rámci testování je na školách prováděno i dotazníkové šetření. Dotazník je zaměřen na základní socioekonomické faktory a také spokojenost žáků ve škole. Testování probíhá v elektronickém diagnostickém systému TEST.

Tab. 1 Celkové počty testovaných žáků 1. ročníků, tříd a škol – vstupní testování

Rok	2007	2008	2009	2010	2011	2012	2013	2014	2015
Počet testovaných žáků	7961	7865	7332	6202	6651	6233	5899	5756	6021
Počet testovaných tříd	319	323	313	280	281	258	237	235	236
Počet testovaných škol	86	90	92	86	88	86	79	79	80

Tab. 2 Celkové počty testovaných žáků 3. ročníků, tříd a škol – výstupní testování

Rok	2010	2011
Počet testovaných žáků	6180	6118
Počet testovaných tříd	279	290
Počet testovaných škol	85	81

2.2 Testování žáků základních a středních škol v česko-slovenské příhraniční oblasti.

Data byla získána v rámci projektu „*Diagnostika stavu znalostí a dovedností žáků v česko-slovenské příhraniční oblasti se zaměřením na jejich rozvoj*“ v letech 2009 až 2011 v Moravskoslezském kraji a také v Trnavském, Žilinském a Trenčianském kraji. Sběr dat v roce 2009 (vstupní testování) 3., 5., 7. ročníku ZŠ a 1. ročníku SŠ, viz tabulka 3, probíhal na konci školního roku, v měsíci květnu a červnu, žáci daných ročníků měli tedy učivo probráno, testy byly sestaveny s ohledem na dobu testování. Sběr dat v roce 2011 5., 7., 9. ročníků ZŠ probíhal na jaře daného roku, od druhé poloviny měsíce března do poloviny měsíce května, také v tomto výstupním testování obsah testů rovněž vycházel

z již probraného učiva. V rámci testování bylo na školách provedeno i dotazníkové šetření. Dotazník byl zaměřen na základní socioekonomické faktory a také spokojenost žáků ve škole. V této práci nebyly analyzovány výsledky škol ze Slovenska.

Tab. 3 Celkové počty testovaných žáků, tříd a škol v Moravskoslezském kraji

Rok	2009	2009	2011	2009	2011	2011	2009
Ročník	3. roč.	5. roč.	5. roč.	7. roč.	7. roč.	9. roč.	1. roč.
Počet testovaných žáků	2776	2743	1560	2877	1502	1422	372
Počet testovaných tříd	152	139	78	159	82	78	16
Počet testovaných škol	86	87	43	89	43	42	5

2.3 Diagnostika stavu znalostí a dovedností žáků 5. a 9. tříd se zaměřením na jejich rozvoj

Tato data byla získána v rámci projektu „*Diagnostika stavu znalostí a dovedností žáků se zaměřením na jejich rozvoj – testování žáků 5. a 9. ročníků základních škol a odpovídajících ročníků víceletých gymnázií*“ v celé ČR. Sběr dat probíhal v měsících duben, květen a červen 2011 elektronickou formou. Součástí testů bylo rovněž dotazníkové šetření. Dotazník byl zaměřen na základní socioekonomické faktory a také spokojenost žáků ve škole. Testování bylo navrženo pro žáky 5. a 9. ročníků a zahrnovalo následující vzdělávací obory a oblasti:

5. ročník:

Český jazyk + Člověk a jeho svět – 30 minut + 20 minut

Matematika + Anglický jazyk – 30 minut + 20 minut

celkem 100 minut testování ve dvou blocích

9. ročník:

Český jazyk + Člověk a příroda – 40 minut + 30 minut

Matematika + Cizí jazyk dle výběru, Anglický jazyk, Německý jazyk – 40 minut + 30 minut

celkem 140 minut testování ve dvou blocích

Testování i vyhodnocování výsledků probíhalo elektronickou formou. Každý test obsahoval úvodní instrukce, cvičné úlohy a časový limit pro zpracování testu. V tabulce 4 jsou obsaženy celkové počty testovaných žáků v roce 2011 v celé ČR.

Tab. 4 Celkové počty testovaných žáků, tříd a škol – vstupní testování

Rok	2011	
Ročník	5. roč.	9. roč.
Počet testovaných žáků	18131	15580
Počet testovaných tříd	963	830
Počet testovaných škol	621	

2.4 Rozvoj a podpora kvality ve vzdělávání ve Zlínském kraji 2013 – 2015

Společnost pro kvalitu školy, o.s. (dále jen SPKŠ) realizovala testování žáků metodou Relativního přírůstku žáka ve školách Zlínského kraje v rámci projektu Podpora kvality ve vzdělávání realizátora Krajského centra dalšího profesního vzdělávání a Centra uznávání a celoživotního učení, o. p. s. V tabulce 7 je možno vidět souhrnné počty zapojených žáků, tříd a škol.

Tab. 5 Celkové počty testovaných žáků, tříd a škol

ZŠ a SŠ	Školy	Třídy	Žáci
2013	87	348	6739
2014	98	444	8378
2015	109	521	9547

Hlavním cílem projektu bylo zavedení komplexního systému zvyšování kvality ve vzdělávání pomocí využití evaluačních nástrojů s propojením se systémem podpory zvyšování kvality ve vzdělávání formou metodických kabinetů a prostřednictvím zavádění inovativních forem učení v ZŠ a SŠ ve Zlínském kraji. První klíčová aktivita „Diagnostika výsledků vzdělávání psychodiagnostickými metodami a adaptivním testováním a měřením přidané hodnoty vzdělání žáků“ zahrnuje celokrajské elektronické testování žáků metodou relativního přírůstku znalostí.

3 MĚŘENÍ PŘIDANÉ HODNOTY VZDĚLÁVÁNÍ

V České republice byly sice před léty publikovány dvě teoretické studie věnované měření přidané hodnoty vzdělávání (Zjišťování přidané hodnoty, 2008, Ryška 2009) i výstupy z reálného měření v Moravskoslezském kraji (Malach, Malčík, 2010 a Malach, Malčík, Krpec, 2011), ale v procesu zavádění plošného testování (které po dvou letech generálních zkoušek v rocích 2012 a 2013 přešlo do režimu tzv. výběrových testování) se s jejím určováním nijak nepočítalo, ani není její stanovování zatím uvažováno.

3.1 Metodika určení relativního přírůstku znalostí žáků

Studie OECD (Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools, 2008) vymezuje *modely přidané hodnoty* tak, že se vždy porovnávají dva časové řezy, v nichž jsou zjišťovány výsledky vzdělávání. Modely přidané hodnoty měří příspěvek školy k pokroku žáka vůči předem určeným školním vzdělávacím cílům. Příspěvek je zhruba očištěná hodnota od jiných faktorů, které také přispívají k pokroku žáka v učení. Tento přístup jasně vymezuje potřebu měřit pokrok žáka, tj. časovou změnu, v níž se odráží působení školy na vzdělávaného jedince (Ryška, 2009).

V jiném pojetí jsou modely zjišťování přidané hodnoty třídou statistických modelů, které se užívají k určení odhadu příspěvku školy k žákovu učení měřenému pomocí trajektorií testových skóre. Tato definice akcentuje statistické vnímání modelů přidané hodnoty, do níž se sbírají data nejčastěji pomocí testových měření.

„Přidaná hodnota je zlepšení, kterého žák dosáhl (v podobě znalostí, dovedností a dalších vlastností) jako výsledku jeho učení se v rámci vzdělávání ve škole“ (Harvey, 2011).

Ukazatele a modely přidané hodnoty

Jednou z hlavních oblastí využití modelů přidané hodnoty je zlepšování kvality vzdělávání na úrovni jednotlivé školy. Jedná se o to, že produkce výkonových indikátorů a informací, které podstatně zpřesní identifikaci výchozího stavu na dané škole, se může stát velice efektivním podkladovým materiálem pro zaměření cílů školy a pro cílené zásahy do naplňování všech funkcí školy. Analýzy dat vytvářejí informace o vztazích mezi školními vstupy a výstupy, což dále zpřesňuje výběr takových efektivních strategií, které upravují potřebné postupy včetně změn ve využití dostupných lidských a finančních zdrojů. Je tak zřejmé, že vedení školy získává k rozhodovací činnosti významná data. To posiluje současný trend, v němž se akcentuje informované rozhodování na základě dostatečného množství informací.

Ukazatele přidané hodnoty lze rozdělit na ukazatele absolutní a relativní (Dolata, 2006; Niemierko, 2009).

Absolutní ukazatel je hodnotnější typ ukazatele. Vyžaduje však to, aby měření výsledků na různých etapách vzdělávání bylo realizováno pomocí stejných měřítek. Například by musel být k dispozici nástroj měřící porozumění textu od 6. ročníku základní školy až po 3. ročník střední školy. Měřící nástroj by musel mít obsahovou interpretaci, což by umožnilo nejen měřit postup, ale také interpretovat získané výsledky z hlediska obsahových očekávání. Takové nástroje zatím nejsou k dispozici, ale je žádoucí zahájit práce na jejich tvorbě.

Relativní ukazatel je méně hodnotným ukazatelem, ale jeho použití vychází z daného stavu řešení problematiky. Jelikož měření není realizováno shodnými nástroji, pak musíme přijmout řešení, že výsledek měření na nižší úrovni je udáván jako celková míra vzdělávacího potenciálu. Většina modelů přidané hodnoty stanovuje relativní ukazatel. Konkrétní modely se pak liší mírou akcentace některých postupů a proměnných. Každý z modelů má své výhody a nevýhody a nelze některý z nich jednoznačně doporučit k používání. Výběr modelu mohou ovlivňovat politické cíle a omezení, za kterých se provádí (Zjišťování přidané hodnoty, 2008). Modely se vyvíjejí, vzájemně kombinují a optimalizují. V zemích OECD jsou vytvořeny a využívány zejména následující modely: lineárně regresní model přidané hodnoty, model přidané hodnoty s pevnými efekty (fixed-effects model), model s rozptylovými složkami (variance component model), model s více závislými proměnnými (multivariate response model), modely s kříženými efekty (cross-classified model) a další (Ryška, 2008).

Klíčové faktory modelu přidané hodnoty

Pokud má být modelování a výpočty přidané hodnoty vzdělání žáka objektivní, vyžadují splnění několika technických a organizačních podmínek.

1. Kvalita údajů o hodnocení žáků a výsledků testů

Vzhledem k tomu, že modely přidané hodnoty jsou postaveny na údajích získaných prostřednictvím hodnocení žáků, je třeba věnovat maximální pozornost charakteru a kvalitě těchto dat. Významnou roli hraje splnění zejména těchto podmínek:

- Lze na základě testů získat údaje vztahující se ke všem (či alespoň těm nejdůležitějším) cílům vzdělávacích programů?
- Absolvují všichni žáci testy za srovnatelných podmínek?
- Jsou výsledky testů dostatečně přesné, aby z nich bylo možno vyvozovat závěry?
- Jsou výsledky testů oproštěny od nepatřičných vlivů a/nebo korupce?

2. Integrita dat a datové pokrytí

Je třeba pečlivě posoudit postupy pro převádění hrubých výsledků testů do použitelných datových souborů, včetně faktoru, jako je úplnost dat. Z obecného hlediska jsou pro modelování přidané hodnoty nezbytné zejména záznamy o žácích za dvě a více let.

3. Filosofie úprav

Modely přidané hodnoty se liší také tím, do jaké míry zahrnují úpravy charakteristik žáků. U některých kategorií modelů jsou takové úpravy hlavním východiskem pro posuzování odhadů jako ukazatelů vlivem příspěvků škol.

4. Technická složitost

Modelů přidané hodnoty je dnes celá řada - od poměrně jednoduchých regresních modelů, až po nesmírně sofistikované modely, které vyžadují bohaté datové základny a nejmodernější výpočetní postupy.

5. Transparentnost

Přestože myšlenka měření přidané hodnoty je intuitivně lákavá, ve školním prostředí může vyvolávat protichůdné názory, zejména pokud jsou motivy pro zavádění takového měření některými zainteresovanými aktéry sledovány s určitým podezřením.

3.2 Metoda relativního přírůstku znalostí

Dostatečně výkonný, efektivní a operativní diagnostický systém musí splňovat některé zásadní technické, odborné a logistické podmínky:

- Ukazatele musí být podloženy dostatečně velkými soubory testovaných jedinců, musí vykazovat dostatečnou úroveň reliability, musí být dostatečně validní.
- Systém musí být operativní s vysokou mírou reliability.
- Systém musí být používán s dostatečnou důvěrou směrem k pozitivní změně.

Hodnocení vstupní úrovně by mělo co nejvěrohodněji odpovídat žakovým úspěchům v pozdějších letech, jinými slovy musí co nejtěsněji korelovat s pozdějšími výsledky v různých předmětech. Míra korelace mezi vstupní úrovní a přidanou hodnotou musí být co největší (predikční validita), hodnocení musí být trvale a konzistentně realizováno u všech učitelů a u všech škol (reliabilita).

Pro zjišťování přidané hodnoty ve vzdělávání byla vyvinuta metoda relativního přírůstku znalostí (Malčík, Krpec, 2010; Malčík, Kubincová, 2008). Hlavní ideou metody relativního přírůstku znalostí je zkoumání změn u žáků v čase s tím, že se porovnávají zejména žáci, kteří měli podobnou výchozí vstupní úroveň. Termín relativní přírůstek znalostí může být interpretován jako relativní pokrok žáků, neboli jak dobře žáci ovládají jistou znalost ve srovnání s ostatními žáky se stejnou vstupní úrovní této znalosti. Jde tedy o vztah mezi statisticky předpokládaným výsledkem testu (skóru) u každého žáka a jím dosaženým aktuálním výsledkem testu.

Relativní přírůstek znalostí je metoda výpočtu přidané hodnoty vzdělání žáka mezi dvěma jeho uzlovými body vzdělávání. Aby toto sledování mělo dostatečnou validitu, je zapotřebí žáka testovat v průběhu celého procesu vzdělávání a za stálých podmínek. Stálými podmínkami se myslí především časová pravidelnost testování, sledování stejných znalostí a dovedností u všech testů v čase, stejná metodika vyhodnocování testů. Metoda relativního přírůstku znalostí je založena na počítačem podporované pedagogické diagnostice. Elektronické didaktické testy umožní okamžité rozšířené vyhodnocení výsledků a individuální

trvalý přístup k výsledkům dílčích oblastí testu a k doporučením pro další práci. Žák může porovnávat své výsledky vzhledem k ostatním žákům ve škole a vzhledem k ostatním žákům stejné skupiny oborů vzdělávání.

Metoda relativního přírůstku znalostí, která byla vyvinuta na základě výsledků v pedagogické diagnostice, zkušeností a teoretických analýz, se skládá ze čtyř navazujících etap:

1. *Hodnocení vstupní úrovně znalostí žáka*
2. *Hodnocení výstupní úrovně znalostí žáka*
3. *Výpočet relativního přírůstku znalostí žáka*
4. *Zpracování relativního přírůstku znalostí a prezentace výsledků a výstupů*

Celá metodika je certifikována ve společnosti Společnost pro kvalitu školy, o. s. podle ČSN EN ISO 9001:2009 v oblasti: „Metodická a odborná činnost v oblasti hodnocení škol”.

Testování a výpočet relativního přírůstku znalostí žáka je navrženo jak v uzlových bodech vzdělávání žáka, tedy v 5., 9. ročníku ZŠ a ve 4. ročníku SŠ, tak také mezi uzlovými body, tedy v 3., 4. a 7. ročníku základní školy a ve 3. ročníku střední školy.

V *první a druhé etapě* jsou žáci elektronicky testováni pomocí náhodně vybíraných variantních úloh podobné obtížnosti, přičemž výsledky jsou harmonizovány. Jsou testovány nejen vědomosti, ale také dovednosti potřebné pro úspěšné řešení úloh. Vyhodnocení testu je okamžité. Žákům je také nabídnuta možnost procvičit si učivo dalšími úlohami. Ze vstupního i výstupního testování jsou žákům poskytnuty komplexní zprávy. Výstupní testování se realizuje po dvou letech.

Ve třetí etapě jsou realizovány zejména tyto aktivity:

- párování dat ze vstupního testování s výstupním testováním na úrovni základních identifikátorů žáka;
- určení regresní přímky po oborech vzdělání a po předmětech;
- určení relativního přírůstku znalostí žáka, třídy, školy.

Pro párování žáků z různých testů byly připraveny speciální algoritmy. Přestože v České republice podle Zákona o ochraně osobních údajů č. 101/2000 Sb. takřka není možno použít pro párování žákovo ID (rodné číslo), byla úspěšnost párování žáků v projektu velmi dobrá na úrovni 80 % všech testovaných žáků.

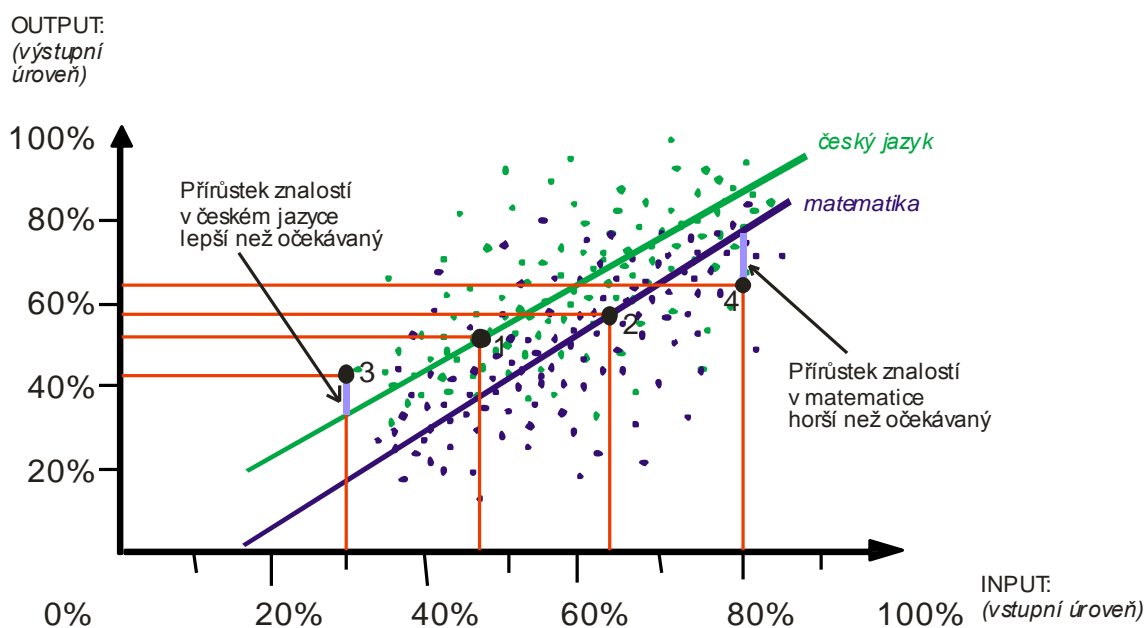
Pro získání relativního přírůstku znalostí žáka potřebujeme znát informaci o tom, do jaké míry žák pracoval lépe či hůře, než by se dalo předpokládat. To, „co by se dalo předpokládat“, jsou výsledky shodné s výsledky podobných žáků na jiných školách. Žáci by si měli být podobní především tím, co se týká předchozích výsledků, protože právě ty jsou nejlepším předpokladem pro výsledky následující (Malčík, Krpec, 2010).

Ve čtvrté etapě je zpracován a interpretován relativní přírůstek žáků. Jeho číselné vyjádření v daném předmětu určuje velikost osově vzdálenosti výsledku žáků od předpokládaného výsledku, kterého měl žák dosáhnout. Tato hodnota může nabývat kladných i záporných hodnot. Kladná číselná hodnota – žák za sledované období dosáhl v průměru lepších výsledků, než se očekávalo. Znamená to, že jeho výsledky jsou lepší než u žáků, kteří při vstupním testu byli na stejné úrovni jako on. Záporná hodnota – žák za sledované období dosáhl horších výsledků, než se očekávalo. Znamená to, že jeho výsledky jsou v průměru horší než u žáků, kteří při vstupním testu byli na stejné úrovni jako on. Čím větší je samotná absolutní hodnota udávajícího relativní přírůstek žáka, tím větší je rozdíl mezi očekávanými znalostmi žáka a měřenými znalostmi žáka. Pro snadnější interpretaci zjištěného relativního

přírůstku žáka, třídy a školy jsou výsledky rozděleny do několika stupňů úspěšnosti rozvoje znalostí a dovedností a tyto stupně jsou slovně okomentovány. Zpracování výsledků a výstupů relativního přírůstku znalostí probíhá takto:

1. Rozdělení – přiřazení relativního přírůstku znalostí žáka do pěti stupňů.
2. Rozdělení – přiřazení relativního přírůstku škol do čtyř stupňů po kvartilech.

Relativní přírůstek byl počítán v rámci oboru vzdělání žáků a daného testovaného předmětu (Schratz, 1997).



Obr. 3 Výpočet relativního přírůstku znalostí žáka

Pro výpočet relativního přírůstku znalostí žáka je použit lineárně regresní model, který je založen na měření znalostí žáka ve dvou časových řezech mezi předchozími a následujícími výsledky (Sanders, Horn, 1994). Obrázek 3 znázorňuje proces určení předpokládaného přírůstku znalostí žáka ve dvou předmětech, českém jazyce a matematice. Na vodorovnou osu jsou vyneseny vstupní procentuální úspěšnosti žáků, výstupní úspěšnosti jsou na svislé ose. Z datového pole výsledků žáků je vypočtena regresní přímka, která je po předmětech a oborech vypočtena podle vztahu:

$$y_{ij} = a + a_1x_{ij} + b_1X_{1ij} + \dots + b_pX_{pij} + \varepsilon_{ij} \quad (1)$$

kde i – označení žáka v rámci j -té školy, y_{ij} – výsledek výstupního testu žáka, x_{ij} – výsledek předchozího testování žáka, $\{X\}$ – charakteristiky žáka nebo jeho/její rodiny, $a_0, a_1, b_1, \dots, b_p$ – regresní koeficienty, ε_{ij} – náhodná chyba s normálním rozdělením, pro jednotlivé žáky nezávislá a se shodným rozptylem.

Regresní přímku pro daný předmět lze chápat jako očekávaný výsledek žáků daného souboru. Regresní přímka zhruba znázorňuje průměrné, střední výsledky žáků, jejichž předchozí výsledky je umístily na určitou pozici na ose vstupních výsledků. Z obr. 3 můžeme vyčíst, že žák 2, jehož úspěšnost při vstupním testování z matematiky je 64 %, by měl dosáhnout úspěšnosti ve výstupním testování úspěšnost 52 %. Tato úspěšnost je předpokládáno

výstupní úspěšností žáka. Pokud žák dosáhne lepších výsledků, než se předpokládalo, žák má „pozitivní reziduál“. Reziduál odpovídá rozdílu mezi měřenou a odhadovanou úspěšností žáka, přičemž odhad je představován regresní přímkou. Někteří žáci získali větší přidanou hodnotu než jiní, jak ukazují jejich reziduály (viz obr. 3) jsou to všichni žáci zobrazení nad regresní přímkou. Číselné vyjádření relativního přírůstku znalostí žáka určuje velikost osově vzdálenosti výsledku žáka na ose výstupů od předpokládaného výsledku, kterého měl žák dosáhnout. Pokud žák dosáhne horšího výsledku, než se předpokládalo, potom má žák „negativní reziduál“; viz žák 4 v obr. 3. Jak již bylo uvedeno, reziduály jsou často označovány jako měřítko přidané hodnoty. Bude ale určitě přesnější, pokud je budeme označovat jako *relativní* přidanou hodnotu.

4 VÝZKUM NĚKTERÝCH PARAMETRŮ METODIKY RELATIVNÍHO PŘÍRŮSTKU ZNALOSTÍ ŽÁKŮ A DIAGNOSTICKÉHO SYSTÉMU TEST

V rámci výzkumu byly pomocí metod matematické statistiky a metod dolování dat analyzovány tyto vlastnosti diagnostického systému TEST a metody relativního přírůstku znalostí žáků:

- Rozšíření testování znalostí o testování dovedností.
- Predikční validita výsledků testu ve spojení s výstupy z paralelního dotazníkového šetření.
- Hledání meziročních závislostí u měření relativního přírůstku znalostí žáků, tříd a škol.

Tento výzkum byl realizován na různě velkých datových souborech, viz část 2.1, které byly v průběhu let 2007 až 2015 naměřeny počítačem podporovaným diagnostickým systémem TEST.

4.1 Rozšíření testování znalostí o testování dovedností

Dovednosti jsou učením získané předpoklady pro vykonávání určité činnosti nebo její části – postup, strategie určité činnosti. Jsou to také dispozice pro užití vědomostí pro řešení problémů, vykonávání činností určitého druhu.

V rámci jednotlivých předmětů je vytvořen seznam dovedností, které jsou v rámci testování sledovány. V současnosti jsou zpracovány a sledovány dovednosti v osmi předmětech a šesti ročnících základních a středních škol, viz příloha 1. Seznam dovedností byl vytvořen experty podle metodiky, která se skládá z těchto postupných kroků:

- analýza očekávaných výstupů a učiva v rámcových vzdělávacích programech základního vzdělávání a gymnázií a výsledků vzdělávání a učiva v rámcových vzdělávacích programech pro střední odborné vzdělávání;
- posouzení a výběr dovedností didaktiky jednotlivých předmětů;
- posouzení a výběr dovedností zkušenými učiteli z praxe.

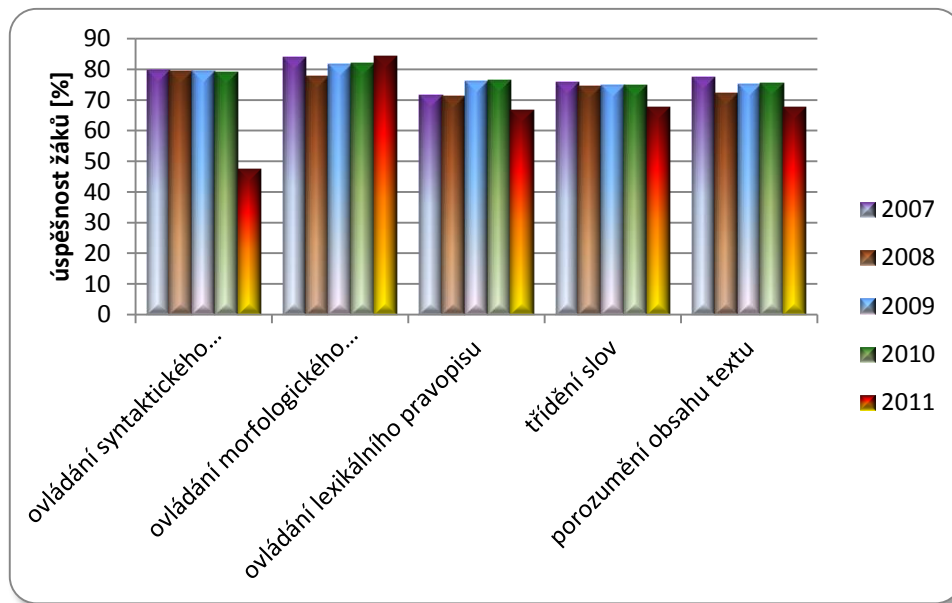
Výsledkem realizované metodiky je seznam dovedností, které jsou podrobně definovány v Číselníku dovedností s jejich popisy. Protože se tento seznam váže na metodiku tvorby úloh pro testování a některé moduly diagnostického systému TEST, podléhá tento číselník změnovému řízení, což je v tomto případě nastavený proces, v rámci kterého jsou na základě definovaných pravidel prováděny úpravy dané metodiky.

V rámci přípravy úloh (viz dále) jsou autory úloh definovány nejen úlohy samotné, ale také dovednosti ze seznamu dovedností, které jsou potřebné pro úspěšné řešení připravených úloh. Při vyhodnocení testů jsou potom u každého žáka vyhodnoceny nejen celkové výsledky v předmětu, ale také informace o míře naplnění sledovaných dovedností v jednotlivých předmětech. Žákův rozvoj v jednotlivých dovednostech je dále zajišťován formou doporučených úloh pro další procvičování.

Tab. 6 Počty dovedností testovaných v jednotlivých předmětech v daném ročníku

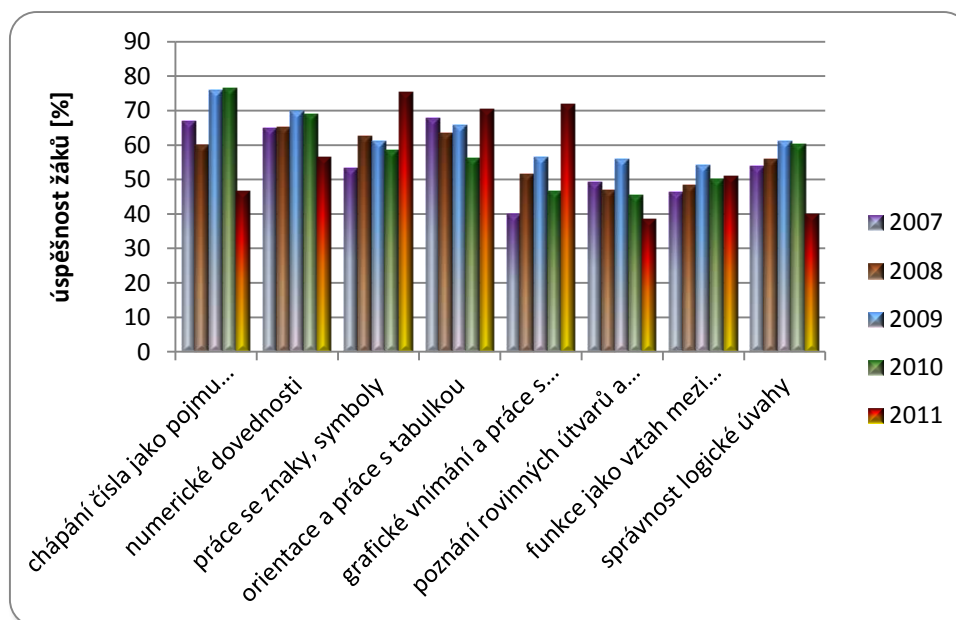
Předmět/Ročník	3. ročník	5. ročník	7. ročník	9. ročník	1. ročník SŠ	3. ročník SŠ
Anglický jazyk	6	6	9	8	8	8
Německý jazyk	5	5	7	7	7	7
Český jazyk	3	3	5	5	5	5
Matematika	8	8	8	8	8	8
Přírodopis/Biologie	3	3	6	5	5	5
Fyzika			5	5	5	5
Chemie				5	5	5

Vyhodnocování dovedností v rámci realizace testů přináší cenné informace pro žáky, učitele, vedení školy, ale také pro zřizovatele a tvůrce školské politiky. V následujících grafech, které vyhodnocují výsledky z testování žáků 1. ročníků středních škol v Moravskoslezském kraji, viz část 2.1 je možno sledovat, jak se vyvíjí úroveň znalostí a dovedností vždy u asi 8000 žáků prvních ročníků v letech 2007 až 2015. V letech 2007 až 2010 byly v těchto vstupních testováních použity stejné úlohy pro všechny žáky, v roce 2011 došlo ke změně úloh i systému, kdy bylo použito variantního testování.



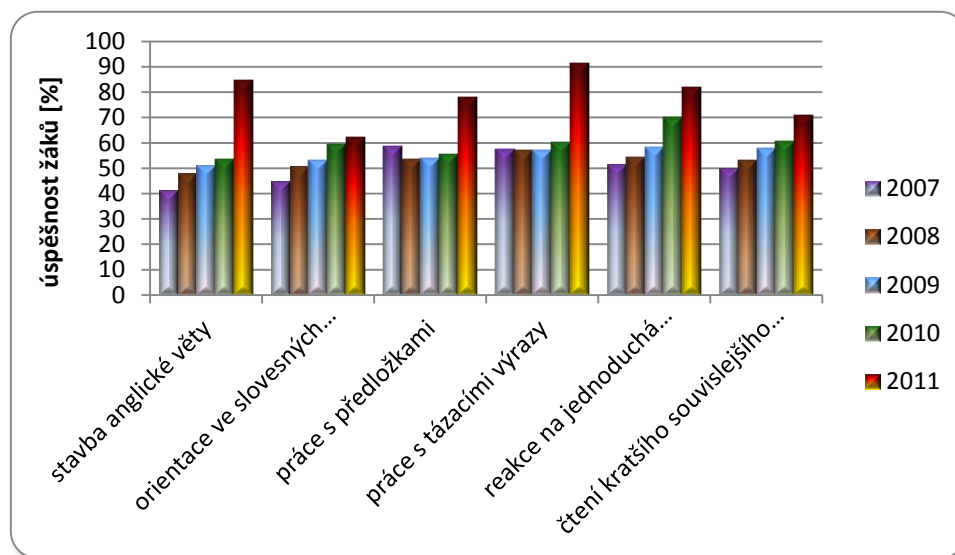
Graf 1 Úspěšnost žáků v jednotlivých dovednostech v českém jazyce v testování 1. ročníků středních škol v Moravskoslezském kraji v letech 2007 – 2011

V grafu 1 je možno sledovat vývoj úspěšnosti žáků 1. ročníků středních škol v Moravskoslezském kraji v jednotlivých dovednostech v českém jazyce. Z grafu lze vysledovat nižší a ještě klesající úspěšnost v dovednosti *porozumění obsahu textu*, což je sledovaná dovednost, která nejvíce odpovídá čtenářské gramotnosti žáků. Výraznější změny v roce 2011 mohou být způsobeny jinými úlohami a mírně odlišnou metodikou testu.



Graf 2 Úspěšnost žáků v jednotlivých dovednostech v matematice v testování 1. ročníků středních škol v Moravskoslezském kraji v letech 2007 – 2011

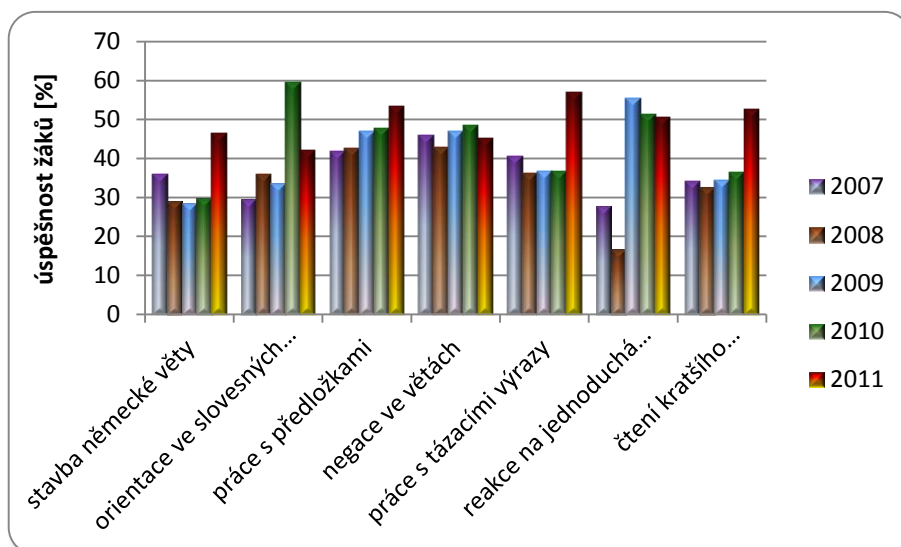
V grafu 2 jsou znázorněny úspěšnosti žáků 1. ročníků středních škol v Moravskoslezském kraji v matematice. Stoupající úspěšnost je možno vidět u dovednosti *správnost logické úvahy*. Velmi nízké úspěšnosti žáci dosahovali v dovednostech *poznávání rovinných útvarů a práce s nimi* a *grafické vnímání a práce s textem*. Tento fakt souvisí také se snižující se úrovní přírodovědné gramotnosti u 15 letých žáků v ČR (PISA, 2009).



Graf 3 Úspěšnost žáků v jednotlivých dovednostech v anglickém jazyce v testování 1. ročníků středních škol v Moravskoslezském kraji v letech 2007 – 2011

Graf 3 ukazuje vývoj úspěšností žáků 1. ročníků středních škol v Moravskoslezském kraji v dovednostech v anglickém jazyce. Nejnižší úspěšnost je možno vysledovat v dovednosti *stavba anglické věty*, výrazně zlepšující úspěšnost vykazují dovednosti *reakce na jednoduchá sdělení* a *čtení kratšího souvislejšího textu s porozuměním*.

Graf 4 znázorňuje vývoj úspěšností žáků 1. ročníků středních škol v Moravskoslezském kraji v dovednostech v německém jazyce. Nejnižší úspěšnost vykazuje dovednost *stavba německé věty*, nejvyšší úspěšnost vykazuje dovednost *práce s předložkami*.



Graf 4 Úspěšnost žáků v jednotlivých dovednostech v německém jazyce v testování 1. ročníků středních škol v Moravskoslezském kraji v letech 2007 – 2011

Metodika tvorby úloh se zřetelem ke sledovaným dovednostem

Pro zpracování úloh je pro autory zpracována metodika, která obsahuje dvě stěžejní oblasti:

1. Způsoby tvorby a formulace úloh pro počítačem podporovanou diagnostiku.
 2. Způsoby určení dovedností, potřebných pro úspěšné řešení úlohy žákem.
1. Autoři jednotlivých úloh dostávají pro jejich přípravu podrobné pokyny v těchto oblastech:
 - možnosti zadání a formulace testových úloh;
 - možné druhy testových úloh;
 - uvedení možných odpovědí;
 - metodika určení dovedností v úlohách.
 2. Podíl dovedností potřebných k vyřešení úlohy se určuje expertní metodou. Základním pravidlem, kterým se metoda řídí, je fakt, že úloha se může „skládat“ maximálně ze tří dovedností. Stanovení míry dovednosti v úloze je určeno konsenzem autora a oponenta úlohy. Pokud se neshodnou, určí výslednou procentuální hodnotu dovednosti garant předmětu. Úlohy jsou do systému předávány v podobě dvourozměrné matice, kde jsou v řádcích úlohy a ve sloupcích procentuální zastoupení odpovídajících dovedností předmětu potřebných ke správnému vyřešení úlohy.

4.2 Predikce výsledků úspěšnosti žáků

Analýzy predikce výsledků úspěšnosti žáků byly realizovány na datech popsanych v části 2.3. Výpočty byly prováděny zvláště pro 5. a 9. ročník, protože testy i výsledky byly rozdílné. Záměrem analýz bylo stanovení vztahů mezi předměty, vztahy mezi předměty a výstupy z dotazníku a zjištění, zda lze v datech nalézt faktory predikce úspěšnosti žáků v jednotlivých předmětech.

5. ročník

Po odstranění extrémních hodnot byla provedena faktorová analýza položek z dotazníků postojů ke škole, která byla součástí testování znalostí. Tyto položky, kterých bylo celkem 13, měli žáci za úkol ohodnotit na 4-bodové stupnici od naprostého nesouhlasu po naprostý souhlas. Dále bylo do analýzy přidáno vzdělání rodičů, kde žáci u každého rodiče volili jednu z možností vzdělání: základní, střední bez maturity, střední s maturitou, vysokoškolské.

Faktorová analýza byla provedena metodou hlavních komponent (Principal components) s rotací Varimax. Výsledné hodnoty jednotlivých faktorů byly prostřednictvím Aderson-Rubinovy metody uloženy jako následující proměnné, které byly pojmenovány takto: vztah k učitelům, škola je zábavná, vzdělání rodičů, důležitost učiva a sebevědomí ve formě Z-skórů, kde průměr je nula a směrodatná odchylka se rovná jedné (viz tabulka 7) V tabulce 7 jsou uvedeny pojmenované faktory a míra, jak jsou tyto faktory syceny jednotlivými položkami dotazníku. Procento vysvětlené variability je 56,7%, jsou uvedeny pouze hodnoty větší než 0,300.

Tab. 7 Faktorová analýza dotazníku – 5. ročníky, rotované faktory

Výroky	Faktory				
	Vztah k učitelům	Škola je zábavná	Vzdělání rodičů	Důležitost učiva	Sebevědomí žáků
Třídní učitel mi pomáhá dosahovat co možná nejlepších výsledků	0,759				
Třídní učitel je férový a spravedlivý	0,702				
Třídní učitel se o mě zajímá a pomáhá mi se školními úkoly	0,693				
Většině učitelů na žácích opravdu záleží	0,533				
Dostávám známky, které si zasloužím	0,418			0,350	
Žáci vycházejí dobře s většinou učitelů	0,375				
Školní úkoly jsou zajímavé		0,734			
Vyučování je zábavné		0,728			
Rád/a dělám práci navíc		0,724			
Nejvyšší dosažené vzdělání matky			0,867		
Nejvyšší dosažené vzdělání otce			0,864		
Učím se to, co budu jednou v životě potřebovat				0,782	
To, co se učíme, považuji za důležité				0,671	
Vím, jak si poradit se školními úkoly					0,780
Jsem úspěšným žákem					0,766

V tabulce 8 jsou hodnoty Pearsonových korelačních koeficientů ze souboru 18 107 žáků 5. ročníků – respondentů studie. Vzhledem k velkému počtu žáků jsou i některé velice malé hodnoty již statisticky významné. Nejtěsnější vztahy jsou mezi mírou úspěšnosti žáků v jednotlivých předmětech. Pokud je tedy někdo úspěšný v jednom předmětu, lze předpokládat, že s velkou pravděpodobností bude úspěšný i v předmětu jiném. Vůbec nejtěsnější vztah u žáků 5. ročníků je mezi výsledky z českého jazyka a výsledky z matematiky.

Tab. 8 Pearsonovy korelační koeficienty výsledků a faktorů, 5. Ročník

	Český jazyk	Matematika	Anglický jazyk	Člověk a jeho svět	Vztah k učitelům	Škola je zábavná	Vzdělání rodičů	Důležitost učiva	Sebevědomí
Český jazyk	1	,583**	,536**	,563**	,082**	-,047**	,262**	,034**	,223**
Matematika	,583**	1	,499**	,532**	,068**	-,037**	,230**	,025**	,228**
Anglický jazyk	,536**	,499**	1	,483**	,057**	-,026**	,224**	,013	,200**
Člověk a jeho svět	,563**	,532**	,483**	1	,067**	-,052**	,219**	,020*	,203**
Vztah k učitelům	,082**	,068**	,057**	,067**	1	,000	,000	,000	,000
Škola je zábavná	-,047**	-,037**	-,026**	-,052**	,000	1	,000	,000	,000
Vzdělání rodičů	,262**	,230**	,224**	,219**	,000	,000	1	,000	,000
Důležitost učiva	,034**	,025**	,013	,020*	,000	,000	,000	1	,000
Sebevědomí žáků	,223**	,228**	,200**	,203**	,000	,000	,000	,000	1

Legenda: * – korelační koeficient je signifikantní na 5% hladině významnosti
 ** – korelační koeficient je signifikantní na 1% hladině významnosti

Z faktorů mají na výsledky v jednotlivých předmětech těsnější vztah pouze dva, a to vzdělání rodičů a sebevědomí žáků. Pokud se jedná o faktor škola je zábavná, ten má ke školním výsledkům dokonce záporný vztah, což znamená, že čím více žák souhlasí s tím, že vyučování je zábavné a školní úkoly jsou zajímavé, tím horší prospěch lze očekávat. Toto zjištění je velice překvapivé a zaslouží další analýzy.

Dalším krokem analýzy byl výpočet regresních rovnic prostřednictvím metody Stepwise (viz tabulka 9). V případě statistiky R^2 jde o takzvaný koeficient determinace, který vyjadřuje, jak velkou část z celkové variability hodnoty vysvětlované proměnné (odhad žákova výsledku) se podařilo vysvětlit naším modelem. Nejvyšší procento vysvětlené variability bylo v případě odhadu výsledků žáka v českém jazyce (47,5 %), nejnižší v případě odhadu výsledků žáka v anglickém jazyce (36,8 %). Celkově se tedy dá říci, že prostřednictvím zjištěných údajů dokážeme předpovědět výsledky testů v nejlepším případě s úspěšností maximálně 50 %.

Tab. 9 Vysvětlená variabilita regresních rovnic pro jednotlivé předměty

Předmět	Procento vysvětlené variability (R ²)
Český jazyk	47,5 %
Anglický jazyk	36,8 %
Matematika	43,2 %
Člověk a jeho svět	40,4 %

V tabulce 10 jsou uvedeny regresní rovnice. Počet prediktorů v rámci odhadu úspěšnosti žáků 5. ročníků v jednotlivých předmětech se pohybuje mezi pěti až osmi, přičemž faktory zjištěné prostřednictvím dotazníku ovlivňují výsledek jen velice málo. Nejdůležitější roli pro predikci výsledků úspěšnosti v jednom předmětu má tedy v našem případě úspěšnost v předmětech ostatních. Za zmínku dále stojí, že sice malou měrou, ale přesto ovlivňuje vztah k učitelům výslednou úspěšnost ve třech ze čtyř případů, a to v pozitivním směru.

Tab. 10 Regresní rovnice odhadu výsledků úspěšnosti žáků 5. ročníků

Předmět	Regresní rovnice
Český jazyk	0,3103 + 0,2192 (Matematika) + 0,2357 (Člověk a jeho svět) + 0,1917 (Anglický jazyk) + 0,01263 (Vzdělání rodičů) + 0,00873 (Sebevědomí) + 0,00494 (Vztah k učitelům) + 0,00279 (Důležitost učiva) - 0,00245 (Škola je zábavná)
Anglický jazyk	0,000742 + 0,3443 (Český jazyk) + 0,1870 (Matematika) + 0,2150 (Člověk a jeho svět) + 0,0113 (Vzdělání rodičů) + 0,0097 (Sebevědomí)
Matematika	0,0965 + 0,4202 (Český jazyk) + 0,2949 (Člověk a jeho svět) + 0,1991 (Anglický jazyk) + 0,0143 (Sebevědomí) + 0,0111 (Vzdělání rodičů) + 0,0032 (Vztah k učitelům)
Člověk a jeho svět	0,2174 + 0,3162 (Český jazyk) + 0,2066 (Matematika) + 0,1607 (Anglický jazyk) + 0,0075 (Sebevědomí) + 0,0069 (Vzdělání rodičů) - 0,0039 (Škola je zábavná) + 0,0023 (Vztah k učitelům)

9. ročník

V případě devátých ročníků byl postup obdobný s tím rozdílem, že počet předmětů se zvětšil ze čtyř na šest. Žáci byli testováni i v německém jazyce, ale ten není do výpočtů zařazen, protože není vhodné slučovat cizí jazyky do jediné kategorie, neboť test i výsledky byly odlišné. Velikost souboru, který byl použit pro analýzu, byl 13 435 žáků.

Po odstranění extrémních hodnot byla provedena faktorová analýza položek z dotazníku postojů ke škole, který byl součástí testování znalostí. Faktorová analýza dotazníku byla provedena podobně jako u souboru žáků 5. ročníků, tedy metodou hlavních komponent s rotací Varimax. Výsledkem je tentokrát 6 faktorů (viz tabulka 13). Do finálního výpočtu nakonec nebyl zařazen výrok: „Dostávám známky, které si zasloužím“, protože nebyl konzistentní se zvoleným modelem.

V tabulce 11 jsou uvedeny pojmenované faktory a míra, jak jsou tyto faktory syceny jednotlivými položkami dotazníku. Procento vysvětlené variability je 56,7%, jsou uvedeny pouze hodnoty větší než 0,300.

Tab. 11 Faktorová analýza dotazníku – 9. ročníky, rotované faktory

Výroky	Faktory					
	Vztah s třídním učitelem	Škola je zábavná	Vztah k učitelům	Vzdělání rodičů	Důležitost učiva	Sebevědomí žáků
Třídní učitel mi pomáhá dosahovat co možná nejlepších výsledků	0,846					
Třídní učitel je férový a spravedlivý	0,815					
Třídní učitel se o mě zajímá a pomáhá mi se školními úkoly	0,785					
Školní úkoly jsou zajímavé		0,787				
Vyučování je zábavné		0,726				
Rád/a dělám práci navíc		0,690				
Žáci vycházejí dobře s většinou učitelů			0,854			
Většině učitelů na žácích opravdu záleží			0,756			
Nejvyšší dosažené vzdělání otce				0,838		
Nejvyšší dosažené vzdělání matky				0,830		
Učím se to, co budu jednou v životě potřebovat					0,870	
To, co se učíme, považuji za důležité					0,726	
Vím, jak si poradit se školními úkoly						0,821
Jsem úspěšným žákem						0,802

Nejtěsnější vztah u žáků 9. ročníků je mezi českým a anglickým jazykem. Následuje vztah českého jazyka a matematiky. Vůbec nejméně mezi předměty spolu korelují fyzika a biologie. Opět se zde ukázalo, že z faktorů má k výsledkům nejtěsnější vztah vzdělání rodičů a žákovské sebevědomí. Toto sebevědomí souvisí zejména s výsledky v matematice, českém a anglickém jazyce. Vzdělání rodičů má nejvýraznější souvislost s úspěšností v jazycích.

V tabulce 12 jsou uvedeny Pearsonovy korelační koeficienty předmětů a nalezených faktorů.

Tab. 12 Pearsonovy korelační koeficienty výsledků a faktorů, 9. Ročník

	Český jazyk	Matematika	Anglický jazyk	Chemie	Fyzika	Biologie
Český jazyk	1	0,528**	0,633**	0,354**	0,271**	0,418**
Matematika	0,528**	1	0,517**	0,336**	0,374**	0,270**
Anglický jazyk	0,633**	0,517**	1	0,303**	0,335**	0,319**
Chemie	0,354**	0,336**	0,303**	1	0,231**	0,265**
Fyzika	0,271**	0,374**	0,335**	0,231**	1	0,182**
Biologie	0,418**	0,270**	0,319**	0,265**	0,182**	1
Vztah s třídním učitelem	0,055**	0,047**	0,071**	0,033**	0,025**	0,017*
Škola je zábavná	0,021*	-0,006	-0,003	0,018*	-0,024**	-0,010
Vztah k učitelům	0,054**	0,062**	0,051**	0,028**	0,036**	0,022*
Vzdělání rodičů	0,240**	0,208**	0,246**	0,090**	0,123**	0,119**
Důležitost učiva	-0,043**	-0,029**	-0,026**	0,001	-0,028**	-0,013
Sebevědomí	0,247**	0,258**	0,242**	0,147**	0,155**	0,140**

Legenda: * – korelační koeficient je signifikantní na 5% hladině významnosti
 ** – korelační koeficient je signifikantní na 1% hladině významnosti

V rámci výpočtu regresních rovnic prostřednictvím metody Stepwise (viz tabulka 13), statistika R^2 vyjadřuje, jak velkou část z celkové variability hodnoty vysvětlované proměnné (odhad žákova výsledku) se podařilo vysvětlit naším modelem. Nejvyšší variabilita se podařila u žáků 9. ročníků vysvětlit v případě predikce výsledků českého jazyka, a to rovných 50 % (viz tabulka 13). Naproti tomu u žádného z přírodovědných předmětů nedosahuje procento vysvětlené variability 20 %.

Tab. 13 Vysvětlená variabilita regresních rovnic pro jednotlivé předměty, 9. ročník

Předmět	Procento vysvětlené variability (R^2)
Český jazyk	50,0 %
Matematika	38,9 %
Anglický jazyk	46,7 %
Chemie	16,8 %
Fyzika	16,9 %
Biologie	18,7 %

V tabulce 14 jsou uvedeny regresní rovnice odhadu výsledků úspěšnosti žáků 9. ročníků v jednotlivých předmětech. Počet prediktorů se pohybuje od pěti (biologie) po osm (český

a anglický jazyk, fyzika). Z faktorů je v každé rovnici přítomno žákovské sebevědomí, ve čtyřech rovnicích vzdělání rodičů.

Tab. 14 Regresní rovnice odhadu výsledků úspěšnosti žáků 9. ročníků

Předmět	Rovnice
ČJ	$0,207 + 0,341$ (Anglický jazyk) + $0,179$ (Matematika) + $0,152$ (Biologie) + $0,096$ (Chemie) + $0,009$ (Vzdělání rodičů) + $0,007$ (Sebevědomí) - $0,004$ (Důležitost učiva) - $0,019$ (Fyzika)
Ma	$-0,038 + 0,296$ (Český jazyk) + $0,142$ (Fyzika) + $0,207$ (Anglický jazyk) + $0,135$ (Chemie) + $0,016$ (Sebevědomí) + $0,010$ (Vzdělání rodičů) + $0,005$ (Vztah k učitelům)
AJ	$0,148 + 0,535$ (Český jazyk) + $0,196$ (Matematika) + $0,091$ (Fyzika) + $0,014$ (Vzdělání rodičů) + $0,010$ (Sebevědomí) + $0,038$ (Biologie) + $0,006$ (Vztah s třídním učitelem) + $0,033$ (Chemie)
Chemie	$0,135 + 0,162$ (Český jazyk) + $0,139$ (Matematika) + $0,096$ (Biologie) + $0,047$ (Fyzika) + $0,035$ (Anglický jazyk) + $0,005$ (Sebevědomí)
Fyzika	$0,217 + 0,331$ (Matematika) + $0,223$ (Anglický jazyk) + $0,107$ (Chemie) + $0,008$ (Sebevědomí) + $0,044$ (Biologie) - $0,069$ (Český jazyk) - $0,005$ (Škola je zábavná) + $0,005$ (Vzdělání rodičů)
Biologie	$0,256 + 0,395$ (Český jazyk) + $0,145$ (Chemie) + $0,066$ (Anglický jazyk) + $0,031$ (Fyzika) + $0,004$ (Sebevědomí)

ZÁVĚR

V rámci výzkumné studie byly zevrubně analyzovány hlavní parametry diagnostického systému TEST a relativního přírůstku znalostí žáků. Nejprve byl popsán způsob určování kvality dovedností žáků v jednotlivých předmětech. Dále byly analyzovány a vyhodnoceny regresní modely relativního přírůstku znalostí žáků, zejména v souvislosti se vzděláním rodičů, motivací žáků k učení se apod. V další etapě studie byly korelovány výsledky z dotazníku s výsledky z didaktického testu a současně byla analyzována predikční validita všech výsledků testování žáků.

Počítačem podporovaný diagnostický systém TEST používali během testování žáci, učitelé využívali výsledky z testování zejména pro zvýšení kvality edukačního procesu. Někteří učitelé realizovali nad získanými daty a metodologiemi akční výzkum. Shrnutím výsledků a výstupů akčního výzkumu učitelů dostáváme následující výhody a nevýhody využití počítačem podporovaného diagnostického systému TEST:

Výhody:

Z hlediska organizace:

- Objektivita při hodnocení výsledků testů, stejná pravidla pro všechny respondenty, nestrannost.
- Variabilita testů, pokud se z databáze otázek na základě určených pravidel generují náhodné testy (rovnost podmínek je zajištěna dobře zpracovanou bankou úloh).

- Eliminace (alespoň v maximalizované míře) podvádění a opisování. Každý žák dostane vlastní variantu testu s jiným pořadím odpovědí.
- Ulehčení učitelů od zdlouhavého opravování někdy velkého počtu testů.
- Vždy aktualizované statistiky úspěšnosti.
- Úspora času.
- Finanční nenáročnost logistiky.
- Velká flexibilita vzhledem k času a místu testování.
- Možnosti dlouhodobého sledování výstupů a trendů z dat.

Z hlediska učení se žáků:

- Dobře zpracované a realizované testování může mít formativní charakter.
- Možné rozšíření o další typy úloh – multimedia, animace aj.
- Zvýšená motivace žáků při využití počítačů pro testování.
- Okamžitá dostupnost a přehlednost výsledků, okamžitá zpětná vazba.
- Okamžité srovnání žáků, tříd, škol.
- Rozšířené možnosti grafického zpracování výsledků.
- Možnost „sebetestování“ žáků jak přímo ve vyučovací hodině, tak i doma.
- Možnost nabídnout žákům doporučené úlohy pro jejich další rozvoj.
- Možnost využít rozšířené techniky testování – pojmové mapy, dovednosti, gramotnosti.

Nevýhody:

- Absence individuálního přístupu učitele k žákovi.
- Někdy je velmi těžké formulovat otázku tak, aby byla očekávaná odpověď jednoznačná.
- Obtížně se využívají a hodnotí otevřené úlohy s textovými odpověďmi.
- Neexistence ekvivalentu pro některé typy otázek v standardním papírovém testu – např. tvořivé úlohy.
- Nároky na výpočetní techniku, server s testovacím systémem, vybavenost žáků počítači a připojením k internetu.
- Někdy je technický problém zajistit autentizaci, tedy jestli skutečně test vykonává zkoušený žák.

Aplikace nových testovacích nástrojů ICT do standardního edukačního procesu, umožňující následné analýzy a na datech založené rozhodování nejen samotných učitelů a škol, ale ovlivňující také vztah k učení a učební procesy samotných žáků a studentů, představuje významný a aktuální úkol, který byl shledán natolik závažným, že se stal v podobě požadavku na realizaci „analýz učení“ (*learning analytics*) jedním ze střednědobých trendů v oblasti základního a středního vzdělávání (Johnson et al., 2014).

Cílem analýz učení je rozvíjet nové pedagogické postupy, zmocnit studenty, aby se aktivně podíleli ve svém učení, vyhledávat v populaci skupiny „rizikových studentů“ a posuzovat faktory ovlivňující studijní úspěšnost studenta a jeho pokroky v učení.

Studie podala přehled o systémovém uchopení počítačem realizované diagnostiky, o funkčním nástroji, který byl a je používám ke zjišťování výsledků vzdělávání, a následně i ke stanovení přidané hodnoty vzdělávání, a současně prezentovala vybrané výstupy analýz na velkých zkoumaných souborech dokumentující působení některých faktorů na školní úspěšnost žáků a studentů.

LITERATURA

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standardy pro pedagogické a psychologické testování*. Praha: Testcentrum, (2001).

Almond, R. G., & Mislevy, R. J. (2002). Enhancing the design and delivery of assessment systems: A four – process architecture (Electronic vision). *Journal of Technology, Learning and Assessment*. 1(5).

Boyle, J., & Fischer, S. (2007). *Educational Testing*. Malden: The British Psychological Society.

Dolata, R. (2006). Edukacyjna wartość dodana w komunikowaniu wyników egzaminów zewnętrznych. *Biuletyn badawczy CKE* [online], no. 8. Dostupné na WWW: http://www.cke.edu.pl/images/stories/badania/biul_8.pdf.

Fitz-Gibbon, C. T. (1999). *Monitoring Education*. London: Continuum.

Chráska, M. (1988). *Metody pedagogické diagnostiky*. Olomouc: UP Olomouc.

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). *NMC Horizon Report: 2015 K-12 Edition*. Austin, Texas: The New Media Consortium.

Kardanova, E. (2008). *Modelirovanie i parametrizacia testov: Osnovy teorii i prilozhenia*. Moskva: Federalnyj centr testirovania.

Krpec, R., Malach J., & Malčík, M. (2011). Seeking Relations in Value-Added Assessment Using the Method of Relative Gain of Knowledge. *Proceedings Information and Communication Technology in Education* (s. 435-444). Ostrava: University of Ostrava.

Liu, O. (2011). Value-added assessment in higher education: a comparison of two methods. *Higher Education*, s. 445-461.

Malach, J., & Malčík, M. (2014). Testování žáků a zjišťování relativního přírůstku znalostí jako systémový prvek řízení kvality škol a školství. In Niemierko, B., Szmigel, M.K. *Diagnozy edukacyjne. Dorobek i nowe zadania*. Kraków. gRUPA TOMAMI, (s. 181-200). ISBN 978-83-63873-13-4.

Malčík, M., & Kubincová, L. (2008). *Testing skills of the 1st year secondary schools pupils*. In Proceedings Information and Communication Technology in Education. Rožnov pod Radhoštěm. Ostrava: Ostravská univerzita. ISBN 978-80-7368-577-5.

Malčík, M., & Malach, J. (2010). Value-added assessment in postsecondary schools. Theoretical approaches and research results in the Czech republic. *Kultura i Edukacja*. Torun: Adam Marszalek, 5(79), 123-154.

Mechlová, E., & Malčík, M. (2011). Testing concept maps electronically. *ICETA 2011*. (s. 137-140). Košice: TU Košice.

Miklošiková, M. (2013). *Didaktika pro tvůrčí vysokoškolské učitele technických předmětů*. Žatec: Ohře Media.

Niemierko, B. (2009). *Diagnostyka edukacyjna*. Warsazawa: Wydawnictwo Naukowe WPN.

Norman, K.L., Friedman, Z., Norman, K., & Stevenson, R. (2001). Navigational Issues in the Design of On-line Self-Administered Questionnaires. *Behav. Info. Tech.* 20, 37-45.

PISA. Assessment Framework, Key Competencies in Reading, Mathematics and Science. Brusel: PISA, OECD, (2009).

Průcha, J. (1996). *Pedagogická evaluace*. Brno: MU.

Sanders, W.L., & Horn, S. (1994). The Tennessee Value-Added Assessment system (TVAAS): Mixed Methodology in Educational Assessment. *Journal of Personnel Evaluation in Education*. 8, 299-311.

Zjišťování přidané hodnoty. Celková zpráva projektu LS0602. Středisko vzdělávací politiky PdF UK Praha, 2008 [online]. [cit. 18.8.2011]. Dostupné z: <http://www.strediskovzdelavacipolitiky.info/download/VAM%20Celkova%20zprava%20projektu.pdf>

KONTAKT

Doc. PhDr. Josef Malach, CSc.

Katedra pedagogiky a andragogiky
Pedagogická fakulta Ostravské univerzity
Fráni Šrámka 3, 709 00 Ostrava
Tel.: +420731505280
E mail: josef.malach@osu.cz

Doc. RNDr. Martin Malčík, Ph.D.

Vysoká škola báňská – Technická univerzita
Katedra společenských věd
E-mail: martin.malcik@vsb.cz

Příloha 1 Číselník dovedností pro metodiku určování relativního přírůstku znalostí ve vzdělávání

Český jazyk – 3. a 5. ročník ZŠ

1. Ovládnání lexikálního pravopisu
4. Třídění slov
5. Porozumění obsahu textu

Český jazyk – 7., 9. ročník ZŠ a 1. a 3. ročník SŠ

1. Ovládnání lexikálního pravopisu
2. Ovládnání morfologického pravopisu
3. Ovládnání syntaktického pravopisu
4. Třídění slov
5. Porozumění obsahu textu

Matematika – všechny ročníky

1. Chápání čísla jako pojmu vyjadřujícího kvantitu; zápis celku různými způsoby
2. Numerické dovednosti
3. Práce se znaky (symboly)
4. Orientace v tabulce a práce s ní
5. Grafické vnímání, práce s grafem a diagramem
6. Poznání rovinných útvarů a prostorová představivost
7. Funkce jako vztah mezi veličinami
8. Správnost logické úvahy

Anglický jazyk – 5. ročník ZŠ

1. Poznat rýmy
2. Stavba anglické věty
5. Práce s předložkami
6. Slovní zásoba
7. Práce s tázacími výrazy
8. Reakce na jednoduchá sdělení

Anglický jazyk – 9. ročník ZŠ, 1. ročník SŠ

2. Stavba anglické věty
3. Stupňování přídavných jmen
4. Orientace ve slovesných strukturách
5. Práce s předložkami
6. Slovní zásoba
7. Práce s tázacími výrazy
8. Reakce na jednoduchá sdělení
9. Čtení kratšího souvislejšího textu s porozuměním

Německý jazyk – 7., 9. ročník ZŠ, 1., 3. ročník SŠ

1. Stavba německé věty
2. Orientace ve slovesných strukturách
3. Práce s předložkami
4. Práce s tázacími výrazy
5. Negace ve větách
6. Reakce na jednoduchá sdělení
7. Čtení kratšího souvislejšího textu s porozuměním

Přírodopis (přírodověda, biologie) – 5. ročník ZŠ

1. Vnímání širších souvislostí
2. Grafické dovednosti
3. Orientace v odborných pojmech a práce s nimi

Přírodopis (přírodověda, biologie) – 7., 9. ročník ZŠ a 1. a 3. ročník SŠ

1. Vnímání širších souvislostí
2. Grafické dovednosti
3. Orientace v odborných pojmech a práce s nimi
4. Objektivně popsat základní znaky biologických objektů a procesů.
5. Dokázat vysvětlit podstatu jevů, procesů a vztahů
6. Aplikovat poznatky a zkušenosti v praktických podmínkách

Fyzika – 7., 9. – ročník ZŠ, 1. ročník SŠ

1. Identifikace a správné používání pojmů
2. Kvalitativní a kvantitativní popis objektů, systémů, jevů a jejich klasifikace
3. Vysvětlování a předvídání jevů a kauzálních souvislostí
4. Grafické vnímání, tvorba a interpretace grafů, tvorba a interpretace schémat a náčrtů
5. Aplikace poznatků, užití fyzikálních zákonů k řešení problémů z praxe, řešení výpočtových a problémových úloh

Fyzika – 3. ročník SŠ

1. Identifikace a správné používání pojmů
2. Kvalitativní a kvantitativní popis objektů, systémů, jevů a jejich klasifikace
3. Vysvětlování a předvídání jevů a kauzálních souvislostí
4. Grafické vnímání, tvorba a interpretace grafů, tvorba a interpretace schémat a náčrtů
5. Aplikace poznatků, užití fyzikálních zákonů k řešení problémů z praxe, řešení výpočtových a problémových úloh
6. Pozorování, experimentování, měření

Chemie – 9. ročník ZŠ, 1., 3. ročník SŠ

1. Identifikace a správné používání pojmů
2. Kvalitativní a kvantitativní popis objektů, systémů, jevů a jejich klasifikace
3. Vysvětlování a předvídání jevů a kauzálních souvislostí
4. Aplikace poznatků
5. Pozorování, experimentování, měření

Geografie – 2. stupeň ZŠ a SŠ

1. Práce s pojmy
2. Využití školního atlasu světa
3. Orientace na mapě
4. Práce s grafy
5. Vnímání širších souvislostí

Pozn. Dovednosti jsou číslovány vzestupně podle ročníků, takže číslo některé dovednosti může v některém ročníku chybět.

E-OCENIANIE – OD PROJEKTU DO WDROŻENIA

Henryk Szaleniec, Wojciech Malecki

Abstrakt: Autoři kapitoly se zamýšlejí nad výsledky projektu, který v posledních osmi letech v Polsku pilotoval způsob elektronického hodnocení otevřených úloh. Popisují průběh projektu, který doplňují příklady konkrétních hodnocení vypracovaných úloh. Metodika využívající názor více hodnotitelů, kdy je zohledňována možnost efektu mírnosti nebo přísnosti, je v roce 2016 využívána pro gymnaziální zkoušky z matematiky pro celou populaci. Čtenář se může seznámit s průběhem vývoje podobného nástroje a utvořit si vlastní názor na možnosti uplatnění takového postupu v pro něj aktuálním edukačním prostředí.

Klíčová slova: e-hodnocení, eMarker, zkouška z matematiky, metodika hodnocení otevřených položek, příprava examinatorů

Abstract: Chapter authors reflect on the results of the project, which in Poland in the last eight years piloted a way of the electronic evaluation of open tasks. They describe the course of the project, which is complemented with the specific examples of elaborated tasks. The methodology, using multiple evaluators' opinion, when the possibility of the effect of moderation or strictness is taken into account, is used in 2016 for grammar exams in Mathematics for the entire population. One can get acquainted with the course of development of a similar tool and form his own opinion on the possibilities of such a procedure for his current educational environment.

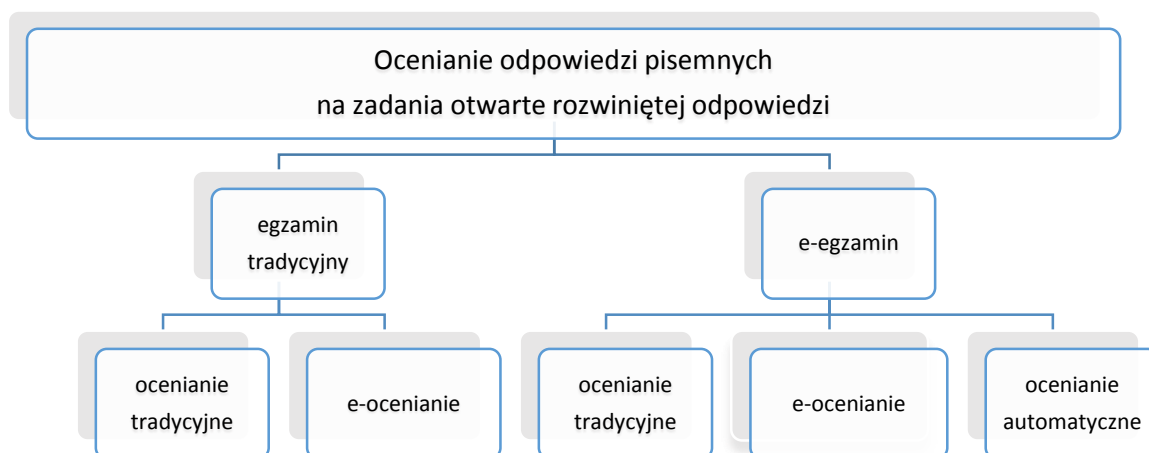
Key words: e-assessment, eMarker, exam in mathematics, methodology of evaluation of open items, preparation examiners

WPROWADZENIE

Projekt *Przygotowanie systemu informatycznego do e-ocenia* realizowany był przez Centralną Komisję Egzaminacyjną i współfinansowany z Europejskiego Funduszu Społecznego w ramach III Priorytetu Programu Operacyjnego Kapitał Ludzki 2007-2013, Działanie 3.2 *Rozwój systemu egzaminów zewnętrznych*. W latach 2015 i 2016 rozpoczęło się wdrażanie e-ocenia dla jednego egzaminu w całej populacji zdających egzamin gimnazjalny z matematyki.

E-ocenie (*electronic marking*) pisemnych prac egzaminacyjnych takich jak eseje z języka ojczystego, historii sztuki, filozofii, ekonomii, rozwinięte rozwiązania prac matematycznych, fizycznych, chemicznych, z ekonomii itp., to ocenianie z wykorzystaniem komputera i sieci. Jest ono rozwiązaniem różnym od tradycyjnego oceniania na papierze pod względem technologicznym, jak i organizacyjnym, ponieważ oceniający otrzymuje zamiast oryginalnej pracy egzaminowanego jej kopię na ekranie wraz z szeregiem narzędzi ułatwiających ocenianie. Warto tutaj jeszcze raz podkreślić, że mówimy o ocenianiu zadań otwartych rozwiniętej odpowiedzi, a nie zadań zamkniętych np. wielokrotnego wyboru, czy na dopasowanie, które mogą być automatycznie oceniane poprzez identyfikację zaznaczeń zdającego z zadaniem kluczem poprawnych odpowiedzi. E-ocenie może być zastosowane zarówno dla tradycyjnego egzaminu pisemnego, jak i w przypadku e-egzaminu. Może mieć

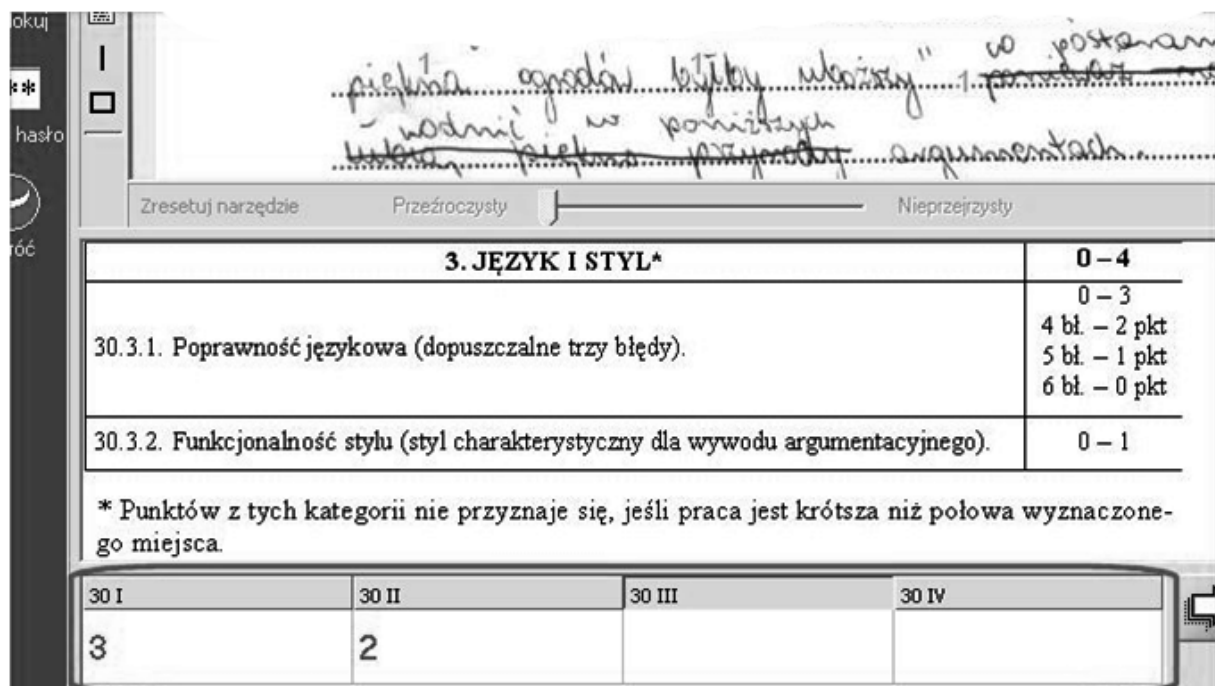
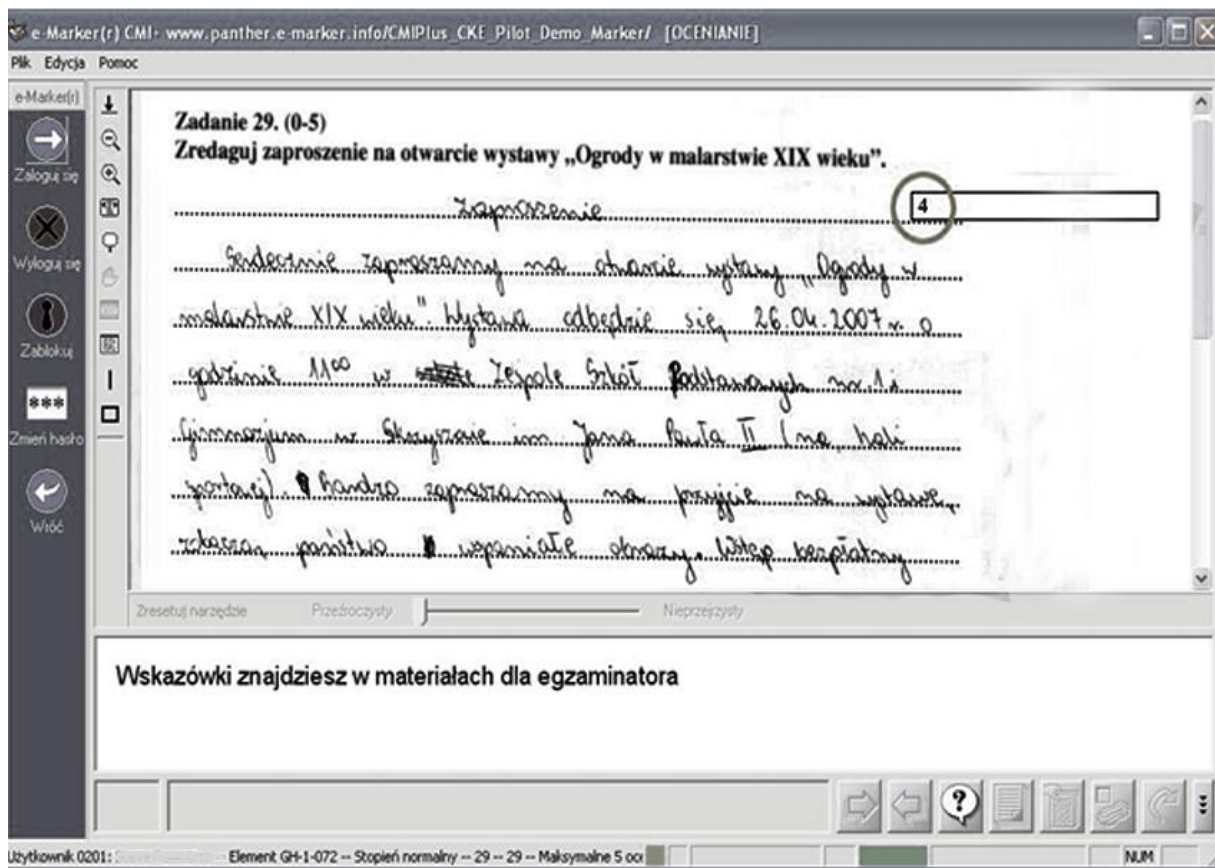
miejsce w masowych egzaminach zewnętrznych, jak i w egzaminach prowadzonych z poszczególnych przedmiotów w dużej uczelni, gdzie prace z tego samego egzaminu oceniane są przez kilku lub kilkunastu egzaminatorów.



Rysunek 1. Różne rozwiązania organizacyjne i technologiczne oceniania pisemnych prac egzaminacyjnych.

Rozwiązanie to pozwala na losowe przydzielanie prac egzaminatorom i bieżące monitorowanie w czasie rzeczywistym jakości oceniania i w razie konieczności kierowanie niektórych prac do ponownej oceny, czy też odsunięcie egzaminatora od pracy w przypadku nie zachowania wymaganych standardów. E-ocenie od kilku lat stosowane jest zarówno w egzaminach na uniwersytetach jak i w wielu systemach egzaminów zewnętrznych. Jak przedstawiono na rysunku 1. może być ono zastosowane w tradycyjnym egzaminie a w szczególności podczas e-egzaminu, kiedy to również zdający wykonuje pracę na komputerze, która jest archiwizowana i przesyłana z wykorzystaniem Internetu (egzamin online) lub za pomocą innego nośnika (egzamin offline) do egzaminatora. W szczególnym przypadku praca może być oceniana automatycznie z wykorzystaniem stosownego oprogramowania (Drasgow, Mattern, 2006; Larkey and Croft 2003), ale w tym artykule nie będziemy zajmować się takimi rozwiązaniami, ponieważ wciąż one są jeszcze dziedziną eksperymentowania i dalekie od zastosowania w językach słowiańskich.

Główna różnica między tradycyjnym ocenianiem prac pisemnych przez nauczycieli szkolnych, akademickich i oceniających zadania otwarte rozwiniętej odpowiedzi w egzaminach masowych a e-oceniem polega przede wszystkim na korzystaniu z innego niż papier nośnika, na którym dostarczane jest egzaminatorowi wypracowanie zdającego. Zamiast papierowego oryginału egzaminator otrzymuje obraz odpowiedzi na pytanie, obraz rozwiązania zadania, które może przeglądać i oceniać na ekranie komputera. Ma też do dyspozycji szereg narzędzi ułatwiających: analizę rozwiązania zadania, podjęcie decyzji, dokumentowanie komentarzy i rejestrowanie wyników.



Rysunek 2. Widok ekranu egzaminatora w programie e-Marker firmy DRS zastosowanym w 2008 roku w pilotażu e-oceny.

Kolejna różnica dotyczy kanału informacyjnego zapewniającego interakcje egzaminatora z ekspertem np. przewodniczącym zespołu oceniającego. Komunikacja odbywa się poprzez Internet. Istotnie zmienia się w porównaniu z dotychczasowymi rozwiązaniami stosowanymi szczególnie w systemie egzaminów zewnętrznych także sposób monitorowania pracy egzaminatora. Monitorowanie prowadzone jest w czasie rzeczywistym na bieżąco. Ponadto e-ocenie umożliwia w łatwy sposób przejście od punktowania przez egzaminatorów całych prac obejmujących od kilku do kilkudziesięciu zadań do specjalizacji w ocenianiu poszczególnych zadań w danej sesji.

Zmiana oceniania, o której tu mówimy, jest adresowana do znacznej części populacji nauczycieli oceniających prace w sesji egzaminacyjnej. Jak każda zmiana związana z implementacją technologii informatycznej do kolejnej dziedziny aktywności zawodowej wymaga czasu i starannego przygotowania tych, którzy będą pracować stosując nowe rozwiązanie. Takie zadanie związane z przygotowaniem kadry powinno mieć nawet wyższy priorytet niż przygotowanie technologiczne i organizacyjne. Wymaga ono nie tylko nabycia nowych umiejętności kluczowych, ale także przełamania barier mentalnych i psychologicznych.

Projekt przygotowania i wdrożenia e-oceniania obejmował kilka etapów.

1. Pilotaż e-oceniania dla 4 egzaminów (2008).
2. Studium wykonalności dla pozyskania rozwiązań informatycznych do projektu *Wdrożenie oceniania z wykorzystaniem technologii informatycznej* (2009).
3. Wybór rozwiązania informatycznego i dostosowanie go do specyfiki masowych egzaminów zewnętrznych w Polsce i szkolenie kluczowej kadry (2010-2011).
4. Studium wykonalności wdrożenia e-oceniania dla całej populacji z zastosowaniem aplikacji *scoris assessor* angielskiej firmy RM Education (2013).
5. Pilotaż na dużej próbie ocenianych prac i doskonalenie aplikacji *scoris* oraz towarzyszących jej rozwiązań informatycznych i organizacyjnych, szkolenie egzaminatorów (2012-2014).
6. Wdrożenie w jednym egzaminie dla połowy populacji i porównanie oceniania w dwóch technikach (2015).
7. Wdrożenie dla jednego egzaminu w całej populacji (2016).

1 PILOTAŻ E-OCENIANIA DLA CZTERECH EGZAMINÓW

We wprowadzaniu zmiany polegającej na przejściu od oceniania tradycyjnego na papierze do e-oceniania można wyróżnić pięć obszarów (Szaleniec, 2007).

1. Niewielkie zmiany w kompozycji arkuszy egzaminacyjnych w celu ułatwienia skanowania odpowiedzi na zadania otwarte.
2. Przygotowanie rozwiązań informatycznych (na poziomie sprzętu i oprogramowania) mających na celu ocenianie i administrowanie ocenianiem z wykorzystaniem nowej technologii.
3. Opracowanie rozwiązań logistycznych w zakresie zarządzania procesem e-oceniania w skali kraju w tym narzędzi do monitorowania jakości e-oceniania.
4. Przygotowanie egzaminatorów do oceniania prac, mając dostęp poprzez sieć do obrazów wypracowań egzaminowanych (*electronic marking*) i zapewnienie wsparcia informatycznego online (*help desk*).

5. Monitorowanie jakości oceniania z wykorzystaniem Sieci (*electronic marking*).

Pilotaż został przeprowadzony wraz z angielską firmą DRS *Data Services Limited* i wykorzystaniem oprogramowania e-Maker[®] dla czterech egzaminów: sprawdzianu na zakończenie szóstej klasy podstawowej, egzaminu gimnazjalnego w części humanistycznej, egzaminu gimnazjalnego w części matematyczno-przyrodniczej i egzaminu maturalnego z matematyki. Główna faza pilotażu (szkolenie egzaminatorów tuż przed ocenianiem i ocenianie) odbyła się w Krakowie w listopadzie 2008 roku. Egzaminatorzy oceniali w centrach oceniania zlokalizowanych w czterech ośrodkach, które zapewniły wystarczającą liczbę stanowisk komputerowych i łącze komputerowe zapewniające szybki transfer danych (obrazów rozwiązań uczniowskich i wyników oceniania) pomiędzy grupą 323 egzaminatorów i bazą danych i centrum przetwarzania wyników. Do oceniania wykorzystano prace archiwalne z wiosennej sesji egzaminacyjnej 2007. Z założenia zadania otwarte z każdej pracy były oceniane przez pięciu egzaminatorów zlokalizowanych w różnych zespołach. Sesje e-oceniania poprzedziło wstępne szkolenie egzaminatorów podczas jednodniowych spotkań, które było kontynuowane z wykorzystaniem e-learningu na platformie MOODLE. 8-9 listopada odbyło się szkolenie w zakresie dostępnych dla egzaminatorów funkcji programu e-Maker[®] dla 24 przewodniczących zespołów oceniających oraz trenerów szkolenia zespołów egzaminatorów. Jednocześnie został sprawdzony pod względem jakości bank zadań weryfikujących⁷, które losowo były umieszczane w puli zadań ocenianych przez każdego egzaminatora. Pilotaż wraz ze szkoleniem bezpośrednim przed ocenianiem trwał dwa dni (15-16 listopada) i uczestniczyły w nim oprócz przewodniczących, zespołu koordynującego i pracowników *helpdesku* 323 osoby reprezentujące wszystkie osiem okręgowych komisji egzaminacyjnych. Z czterech przedmiotów oceniono łącznie 90813 rozwiązań zadań oraz dodatkowo 27782 zadań weryfikujących. Przed przystąpieniem do oceniania egzaminatorzy oszacowali czas konieczny na ocenianie kierując się doświadczeniami z głównej sesji egzaminacyjnej. Porównanie szacowanego i rzeczywistego czasu oceniania przedstawia tabela 1.

Tabela 1. Porównanie oszacowanego czasu oceniania zadań w ocenianiu tradycyjnym i rzeczywistego czasu e-oceniania (źródło: Raport zespołu pilotażu i DRS).

Egzamin	Przewidywany czas oceniania	Rzeczywisty czas oceniania	Oszczędność czasu
Sprawdzian na zakończenie szkoły podstawowej	1011	449	55,6%
Egzamin gimnazjalny część humanistyczna	6945	4610	33,6%
Egzamin gimnazjalny część matematyczno-przyrodnicza	1634	477	70,8%
Matura z matematyki - poziom rozszerzony	825	438	46,9%

⁷ Zadania weryfikujące, to ocenione już zadania przez zespół sędziów kompetentnych, których ocena jest przez system porównywana z oceną danego egzaminatora. Wyniki tych porównań są przedmiotem analiz i ewentualnych konsultacji z przewodniczącym zespołu oceniającego, a w przypadku powtarzających się uchybień przekraczających przyjęty przedział tolerancji egzaminator nie jest dopuszczany do dalszego oceniania. Jego pula zadań może być ponownie oceniona przez innego egzaminatora.

Bezpośrednio z oprogramowania e-Maker[®] w aplikacji administratora egzaminu generowane były raporty eksportu wyników do bazy danych, raporty różnic w ocenianiu rozwiązań zadań przez pięciu egzaminatorów. System umożliwiał także analizy raportów dotyczące jakości pracy poszczególnych egzaminatorów, jakości oceniania poszczególnych zadań oraz funkcjonowania zadań weryfikujących.

Tabela 2. Fragment raportu z porównania sumy ocen za wszystkie zadania otwarte przez pięciu egzaminatorów dla egzaminu gimnazjalnego w części matematyczno-przyrodniczej (źródło: Raport zespołu pilotażu i DRS).

Szkoła	ID	Imię	Nazwisko	Egzaminator					Różnica
				A	B	C	D	E	
180605-	A21	Marek	...	17	17	17	17	18	1
180605-	A22	Barbara	...	11	12	12	12	12	1
180605-	B08	Szymon	...	22	22	22	22	22	0
180605-	B18	Damian	...	18	19	17	17	17	2
180605-	B21	Tomasz	...	23	23	23	23	23	0
180605-	C07	Bernadeta	...	18	18	18	19	19	1
180605-	C08	Ewa	...	25	25	25	25	25	0
180605-	C09	Marta	...	12	12	12	12	12	0
180605-	C10	Grzegorz	...	10	10	10	11	10	1
180605-	C11	Magdalena	...	8	8	8	8	8	0
180605-	C14	Michał	...	10	11	10	12	10	2

Jak można zauważyć analizując tabelę, egzaminatorzy oceniający prace z egzaminu gimnazjalnego w części matematyczno-przyrodniczej są wysoce zgodni w swoich ocenach, co oznacza mały wpływ efektu egzaminatora na końcowy wynik ucznia.

Inaczej wyglądała sytuacja w przypadku egzaminu w części humanistycznej co przedstawia tabela 3.

Tabela 3. Fragment raportu z porównania sumy ocen za wszystkie zadania otwarte przez pięciu egzaminatorów dla egzaminu gimnazjalnego w części humanistycznej (źródło: Raport zespołu pilotażu i DRS).

Szkoła	ID	Imię	Nazwisko	Egzaminator					Różnica
				A	B	C	D	E	
121608-	C14	Damian	...	11	11	13	11	16	5
121608-	C15	Bartłomiej	...	20	25	22	23	22	5
121608-	C16	Elżbieta	...	14	16	14	14	10	6
121608-	C17	Joanna	...	13	17	7	13	13	10
121608-	C18	Rafał	...	2	5	4	3	5	3
121608-	C19	Kamil	...	18	11	14	19	21	10
121608-	C21	Wioletta	...	5	8	13	16	12	11
121608-	C23	Wojciech	...	-	-	12	20	-	8
121608-	C25	Magdalena	...	-	-	10	9	-	1
121608-	C26	Sylwia	...	-	-	13	23	-	10
121608-	C27	Karolina	...	-	23	17	25	-	8

Na podstawie raportów generowanych w trakcie oceniania przewodniczący zespołu oceniającego pozyskiwał informację pozwalającą na kontrolę jakości oceniania i interwencję poczynając od wyjaśnień i dodatkowych zadań ćwiczeniowych aż do wykluczenia egzaminatora z sesji oceniającej.

1.1 Wnioski z pilotażu

Pomimo dużych wcześniejszych obaw egzaminatorów, to była bardzo efektywna sesja oceniania. Nowe umiejętności, zarówno w zakresie posługiwania się oprogramowaniem do e-oceniania, jak i komunikacyjne (online) z przewodniczącym zespołu egzaminatorów, zostały opanowane szybko. Odpowiedzi do krótszych, lepiej ustrukturyzowanych zadań były oceniane bardziej efektywnie niż zadania o dłuższych rozwiązaniach, w których obserwowany jest czynnik subiektywny - efekt oceniającego (Szaleniec i inni, 2015). Przy ocenianiu zadań rozwiniętej odpowiedzi w zadaniach z przedmiotów humanistycznych efekt oceniającego, wyrażający się zróżnicowaniem przydzielanej liczby punktów za to samo zadanie (przez różnych egzaminatorów), był obserwowany podobnie jak w ocenianiu tradycyjnym. Jednak dzięki zastosowaniu e-oceniania, efekt ten mógł być kontrolowany w czasie rzeczywistym poprzez zastosowanie zadań weryfikujących ukierunkowanych na zapobieganie rozbieżnościom w ocenianiu. Ponadto dzięki ocenianiu zadaniami, a nie całymi pracami, efekt surowości lub łagodności egzaminatora nie kumulował się w ocenie całej pracy danego ucznia. Zaobserwowano także znaczną oszczędność czasu oceniania (por. tabela 1). W kontroli jakości oceniania newralgicznym zadaniem okazało się przygotowanie przez zespół sędziów kompetentnych zróżnicowanych i wysokiej jakości ocenionych zadań weryfikacyjnych.

2 JAK PILOTAŻ POSTRZEGANY BYŁ PRZEZ UCZESTNICZĄCYCH W NIM EGZAMINATORÓW

Każda zmiana, aby mogła zaistnieć, wymaga akceptacji osób, których ona dotyczy. Dlatego też pilotaż pełnił także rolę wstępnego studium wykonalności. Pilotaż e-oceniania w czterech egzaminach pozwolił zaobserwować, że nie ma przeszkód merytorycznych i logistycznych oraz kadrowych, które uniemożliwiłyby w niektórych egzaminach przejście od oceniania tradycyjnego do e-oceniania. Uczestniczący w pilotażu egzaminatorzy, pełniący kluczowe funkcje w strukturze organizacji oceniania w ośmiu OKE i CKE, podkreślali szereg walorów oceniania prac z wykorzystaniem obrazów rozwiązań zdających, które są dostępne na monitorze komputera. Po pierwsze, egzaminator otrzymuje szereg dodatkowych narzędzi wspomagających proces oceniania. Po drugie, nowy proces monitorowania oceniania z wykorzystaniem zadań weryfikujących pozwala na poczucie komfortu wysokiej jakości oceniania. Po trzecie, możliwość specjalizowania się w ocenianiu poszczególnych zadań daje poczucie większego bezpieczeństwa w podejmowaniu decyzji o wynikach egzaminu. Ponadto w dalszej perspektywie istnieje szansa przejścia od oceniania w trybie „skoszarowanym” do oceniania przy własnym komputerze, co niewątpliwie daje szansę na lepsze gospodarowanie czasem, jaki jest pomiędzy egzaminem a terminem ogłoszenia wyników. Z perspektywy okręgowych komisji egzaminacyjnych, odpowiedzialnych za jakość oceniania, na szczególne podkreślenie zasługuje możliwość ciągłego monitorowania jakości oceniania i w razie konieczności odsunięcie od pracy egzaminatora, który nie radzi sobie z podejmowaniem decyzji w zakresie punktowania rozwiązań z godnie z ustalonymi przez CKE kryteriami.

3 STUDIUM WYKONALNOŚCI I REKOMENDACJA WDROŻENIA

Studium wykonalności dla wdrożenia oceniania z wykorzystaniem technologii informatycznej w polskim systemie egzaminów zewnętrznych to drugi i czwarty etap prowadzący od projektu do wdrożenia. Pierwsze studium zostało przygotowane w 2009 roku przez zewnętrzną firmę przy współpracy z zespołem projektowym CKE. Studium obejmowało, po pierwsze, analizę instytucjonalną z uwzględnieniem wykonalności instytucjonalnej projektu, gotowość do zmian kadry egzaminatorów w systemie egzaminów zewnętrznych oraz analizę prawną. Po drugie, analizę techniczną i technologiczną z uwzględnieniem wymagań funkcjonalnych wobec systemu, porównania dostępnych technologii w Europie, propozycji czterech alternatywnych wariantów infrastruktury technicznej i implementacji infrastruktury sprzętowej. Kolejne obszary analiz to ocena finansowa i ekonomiczna przedsięwzięcia, jakim jest wdrożenie e-oceniania.

Na bazie doświadczeń z pilotażu i analiz studium wykonalności w 2010 roku przygotowano rekomendację wdrożenia. Kluczową kwestią było pozyskanie aplikacji do e-oceniania, czyli systemu informatycznego (zakup i adaptacja lub zbudowanie od podstaw), który zapewniłby obsługę złożonego procesu oceniania egzaminu zewnętrznego w skali kraju począwszy od przygotowania bazy wypracowań egzaminowanych uczniów (kilkaset tysięcy zdających), standaryzację schematów i procesu oceniania, zarządzanie procesem oceniania na komputerach zlokalizowanych w domu egzaminatora, aż do wygenerowania wyniku egzaminacyjnego każdego ucznia, który przystąpił do egzaminu. Systemu informatycznego, który zapewniłby jednocześnie kontrolę jakości oceniania, koordynację na poziomie zespołów oceniających jak i całego krajowego egzaminu. W przyjętym rozwiązaniu e-oceniania niebagatelną sprawą jest też sprawne funkcjonowanie pomocy technicznej dla tysięcy egzaminatorów pracujących w rozproszeniu w własnych domach. To było zadanie *helpdesku*, do którego kadra musiała być stosownie przygotowana w ramach projektu.

4 TRZECI ETAP PRZYGOTOWANIA DO WDROŻENIA – TESTOWANIE OPROGRAMOWANIA I SZKOLENIA KADRY

W ramach trzeciego etapu przygotowania do wdrożenia e-oceniania została opracowana adekwatna do polskiego systemu egzaminacyjnego aplikacja *scoris assessor* angielskiej firmy RM Education. Testowanie aplikacji planowanej do wdrożenia zostało przeprowadzone w trakcie kolejnego pilotażu e-oceniania w tych samych egzaminach, które były włączone do pierwszego pilotażu. Podczas pilotażu e-oceniania w listopadzie 2011 r. w procesie przygotowania sesji i w jej przeprowadzeniu uczestniczyło łącznie 310 egzaminatorów i 11 informatyków z ośmiu okręgowych komisji egzaminacyjnych.

Dzięki uczestnictwu w pilotażu i testowaniu systemu *scoris assessor* kluczowa kadra do planowania, administrowania oraz zarządzania sesją e-oceniania pozyskała umiejętności praktyczne niezbędne do pełnienia różnych ról charakterystycznych dla e-oceniania. Kapitał wiedzy i umiejętności, pozyskany podczas szkolenia towarzyszącego pilotażowi, miał szczególne znaczenie dla kolejnego etapu przygotowania wdrożenia. Rysunek 3 przedstawia widok ekranu egzaminatora oceniającego pracę z egzaminu maturalnego z matematyki w aplikacji *scoris assessor* natomiast tabela 4 przybliży jeden z 44 raportów, które generowane z poziomu aplikacji umożliwiają monitorowanie oceniania w trakcie sesji i po jej zakończeniu.

scoris assessor v4 - [Tryb oceniania 6191854]

Plik Widok Obraz Znaczniki Rozwiązanie Narzędzia Okno Pomoc

Lista zadań Widok ogólny Tryb oceniania 6156838 Tryb oceniania 6191854 Zamknij

Poprzednie rozwiązanie Następane rozwiązanie

(1)

10 *Egzamin maturalny z matematyki
Poziom podstawowy*

ZADANIA OTWARTE

Rozwiązania zadań o numerach od 26. do 34. należy zapisać w wyznaczonych miejscach pod treścią zadania.

Zadanie 26. (2 pkt)
Rozwiąż nierówność $x^2 - x - 2 \leq 0$.

$x^2 - x - 2 \leq 0$
 $\Delta = b^2 - 4ac$
 $\Delta = 1^2 - 4 \cdot 1 \cdot (-2)$ ✓
 $\Delta = 1 + 8$
 $\Delta = 9$
 $x_1 = \frac{-b - \sqrt{\Delta}}{2a}$ ✓
 $x_1 = \frac{1 - 3}{2}$ ✓
 $x_1 = -1$
 $x_2 = \frac{-b + \sqrt{\Delta}}{2a}$ ✓
 $x_2 = \frac{1 + 3}{2} = \frac{4}{2} = 2$ ✓

$x \in (-\infty; -1] \cup [2; +\infty)$

Komponent MMA-P1A1P-102/01, zestaw zadań MMA1, Whole Paper : Tryb oceniania

Kliknij prawym przyciskiem, aby uzyskać opcje powiększania.

Egzaminator: FAMapze547

Postęp: 0%

Suma punktów: 0

Zadanie	Punkty
Whole Paper	
Z261	-
Z271	-
Z281	-
Z291	-
Z301	-
Z311	-
Z321	-
Z331	-
Z341	-

Zakończ ocenę

Element: Z261

2 ↩

MS - 1

/2

Brak Reset

Rysunek 3. Widok ekranu egzaminatora w aplikacji *scoris a assessor*.

Egzaminator siedząc przed swoim komputerem dysponuje całą gamą narzędzi ułatwiających analizę uczniowskiego rozwiązania widocznego na ekranie.

Tabela 4. Fragment raportu – Przegląd aktywności egzaminatora.

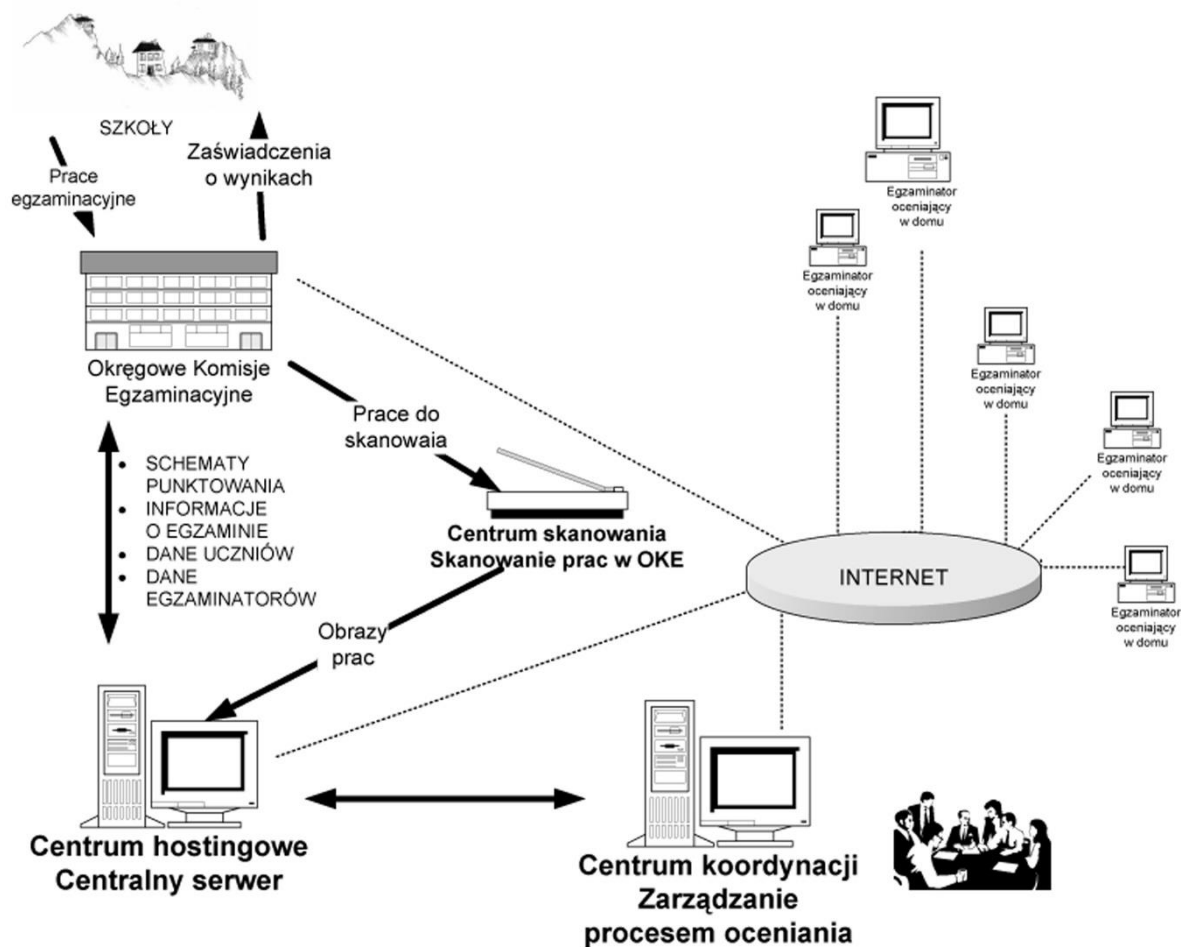
Nazwa egzaminu	Nazwa zadania	Egzaminator	ID egzaminatora	Docelowa liczba rozwiązań	Aktualnie oceniane	Rozwiązania trzymane (łącznie z procesem kontrolnym)	Rozwiązania zamknięte (łącznie z procesem kontrolnym)	Wstrzymane i zamknięte/ docelowa liczba rozwiązań	Docelowa data
Gimnazjum Mat.-Przyr. 2010	GM GZ 1 (arytmetyka)	<u>Maria</u>	xxx	90	1	0	89	98,9%	24/11
Gimnazjum Mat.-Przyr. 2010	GM GZ 2 (algebra i geometria)	<u>Jan</u> ...	zzz	90	0	0	70	77,8%	24/11

Innym bardzo ważnym zagadnieniem było przygotowanie egzaminatorów do oceniania uczniowskich wypracowań na ekranie własnego komputera. Egzaminatorzy musieli opanować nowe kluczowe kompetencje pracy w Sieci z wykorzystaniem narzędzi informatycznych. Jednocześnie konieczne było stworzenie warunków, aby wprowadzenie e-oceniania podczas rzeczywistej sesji odbyło się bez żadnych zakłóceń w ocenianiu innych egzaminów. Szkolenie egzaminatorów było głównym zadaniem projektu w 2013 roku.

5 WDROŻENIE E-OCENIANIA NA EGZAMINIE GIMNAZJALNYM Z MATEMATYKI W 2015 I 2016 R.

W wiosennej sesji egzaminacyjnej 2015 r. po raz pierwszy w Polsce przeprowadzono e-ocenianie zadań otwartych rozszerzonej odpowiedzi z prac z egzaminu gimnazjalnego w zakresie matematyki w połowie populacji gimnazjalistów. W kolejnym roku 2016 e-ocenianie zostało zastosowane dla tego samego egzaminu już dla całej populacji. Jak już wspomniano do e-oceniania wykorzystana została aplikacja *scoris* autorstwa angielskiej firmy RM Education dostosowana do realiów polskiego egzaminu w ramach projektu. Rezultaty e-oceniania były podstawą do ustalenia wyniku końcowego uczniów z tego egzaminu.

W odróżnieniu od oceniania tradycyjnego, gdzie egzaminatorzy oceniali w zespołach zlokalizowanych w centrach oceniania pod bezpośrednim kierownictwem przewodniczących zespołów (zwykle piątek sobota, niedziela), e-ocenianie odbywało się przy domowym komputerze. Podczas e-oceniania każdy egzaminator specjalizował się w ocenianiu tylko jednego zadania, a rozwiązania zdających przydzielane były losowo niezależnie od tego, w obrębie której okręgowej komisji egzaminacyjnej znajdowała się szkoła. Rysunek 4 przedstawia schematycznie relacje pomiędzy głównymi filarami systemu w procesie e-oceniania: komisje egzaminacyjne – centrum koordynacji e-oceniania – centralny serwer i rozproszeni w całej Polsce egzaminatorzy.



Rysunek 4. Schemat organizacji e-oceny.

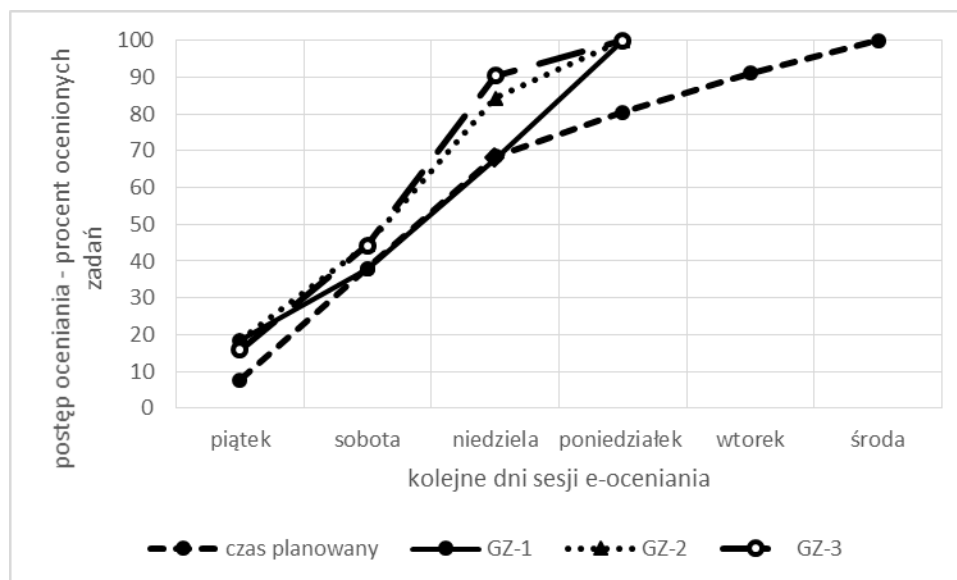
Jak już wspomniano, w 2015 roku w e-oceny uczestniczyła połowa (4) okręgowych komisji egzaminacyjnych (OKE). W drugiej połowie OKE prace z gimnazjalnego egzaminu z matematyki oceniane były w tradycyjny sposób. Takie rozwiązanie pozwoliło z jednej strony, wdrażać innowację skalując jej zasięg terytorialny, kadrowy i obciążenie pod względem informatycznym, a z drugiej, umożliwiło porównanie rezultatów na poziomie kraju przy zastosowaniu dwóch różnych rozwiązań. W trakcie sesji w e-oceny uczestniczyło 744 egzaminatorów zgrupowanych w 50 zespołach i oceniło łącznie prace przeszło 160 tysięcy uczniów. Pod względem organizacji oceniania istotną innowacją jest losowy przydział prac egzaminatorom niezależnie od ich przynależności do konkretnej OKE, jak również koordynacja zespołami bez podziału na komisje oraz ocenianie na własnym domowym komputerze a nie w centrach oceniania. Główna zmiana, oprócz pracy na komputerze z wykorzystaniem dedykowanej aplikacji *scoris*, dotyczyła jednak bardziej skutecznej niż w ocenianiu tradycyjnym kontroli jakości. Kontrolę jakości oceniania stanowiły trzy mechanizmy.

1. Sesja ćwiczeniowa online z wykorzystaniem aplikacji *scoris assessor*, która umożliwiała egzaminatorom najpierw opanowanie, a potem sprawdzenie swoich umiejętności posługiwania się oprogramowaniem i umiejętności stosowania modelu punktowania.

2. Sesja kwalifikacyjna, w wyniku której po poprawnym oceniu próbki zadań egzaminator dopuszczany był do pracy – uzyskiwał dostęp do puli zadań do oceny.
3. Ciągłe monitorowanie procesu poprzez ocenianie dostarczanych egzaminatorom losowo zadań kontrolnych stanowiących 6 % liczby sprawdzanych przez danego egzaminatora prac.

Nadzorowanie zachowania wysokich standardów oceniania prac egzaminacyjnych było zadaniem koordynatorów egzaminu gimnazjalnego z matematyki z okręgowych komisji egzaminacyjnych oraz Centralnej Komisji Egzaminacyjnej. Podczas sesji e-oceny mieli oni możliwość bieżącego monitorowania procesu logowania się egzaminatorów i przewodniczących zespołów oceniających do systemu, obserwowali proces kwalifikowania poszczególnych egzaminatorów, postęp oceniania oraz funkcjonowanie zadań kontrolnych. Pozwoliło to na ciągłe monitorowanie jakości oceniania. W razie potrzeby podejmowali też, wraz z przewodniczącymi zespołów oceny, stosowne decyzje merytoryczne i personalne dla zapewnienia jakości. Narzędziami, które umożliwiały monitorowanie, były cztery grupy raportów generowanych przez aplikację *scoris assessor*. Raporty te to: wgląd w postęp kwalifikacji członków zespołu, przegląd ogólny oceniania na poszczególnych etapach koordynacji, raport postępu koordynacji oceniania, podsumowanie aktywności egzaminatora. Ciągłość i bezawaryjność systemu zapewniał zespół informatyków i menedżerowie sesji.

Sesja e-oceny obejmująca szkolenie online przed egzaminem, ćwiczenia weryfikujące, na podstawie których oceniający dopuszczany jest do oceniania i proces oceniania w 2015 roku została zaplanowana na 6 dni. Jak wyglądało natężenie oceniania w zaplanowanym okresie podczas oceniania trzech zadań otwartych z matematyki GZ-1, GZ-2, GZ-2 I GZ przedstawia rysunek 5.



Rysunek 5. Postęp oceniania w poszczególnych dniach sesji.

Pomimo, że sesja planowana była od piątku do środy, to w zasadzie ocenianie zakończyło się w poniedziałek. Oszczędność czasu w stosunku do szacowanego zaobserwowana została już podczas pilotażu w 2008 roku (por tabela 1). Wtedy jednak e-ocenianie przeprowadzono w zespołach pracujących w centrach oceniania, a nie w domu. Niewątpliwie wpływ na zwiększenie sprawności oceniania miało rozwiązanie, w którym ocenijający specjalizował się w ocenianiu tylko jednego zadania.

Oprócz oszczędności czasowych, finansowych i większej sprawności organizacyjnej e-oceniania w porównaniu z ocenianiem tradycyjnym, najważniejszym aspektem wprowadzanej innowacji jest możliwość bieżącej kontroli jakości. Już samo odnotowanie skali ocen różniących się od ocen modelowych w zadaniach kontrolnych ma kolosalne znaczenie dla procesu doskonalenia egzaminów - nie tylko masowych egzaminów zewnętrznych. Wiadomo, że ocenianie zadań otwartych rozwiniętej odpowiedzi, szczególnie punktowanych na długiej skali, obarczone jest efektem egzaminatora (Szaleniec i inni, 2015; Kulon, Żółtak 2014). Wolne od tego efektu są zadania zamknięte. Ale to zadania otwarte rozwiniętej odpowiedzi mają wiele cennych zalet z punktu widzenia procesu kształcenia, jak również ze względu na dążenie do wyższej trafności oceniania. Efektu egzaminatora i błędów losowych egzaminatora nie da się całkowicie wyeliminować. Można je jedynie minimalizować. E-ocenianie daje wiele możliwości w tym zakresie, począwszy od identyfikacji natężenia tego zjawiska dla poszczególnych zadań egzaminacyjnych i wśród ocenających. Częstość błędnych ocen zadań kontrolnych w ocenianiu trzech zadań otwartych.

	GZ-1	GZ-2	GZ-3	Ogółem
Liczba egzaminatorów	269	260	215	744
Liczba egzaminatorów, którzy niepoprawnie ocenili co najmniej jedno zadanie	237	232	186	655
Odsetek egzaminatorów, którzy niepoprawnie ocenili co najmniej jedno zadanie	88,1%	89,2%	86,5%	88,0%
Liczba zastosowanych zadań kontrolnych w sesji e-oceniania	11176	10383	8411	29970
Liczba poprawnie ocenionych zadań kontrolnych	10112	9686	8079	27877
Odsetek poprawnie ocenionych zadań kontrolnych	90,5%	93,3%	96,1%	93,0%
Liczba niepoprawnie ocenionych zadań kontrolnych	1063	697	332	2092
Odsetek niepoprawnie ocenionych zadań kontrolnych	9,5%	6,7%	4,0%	7,0%

Tabela 5. Natężenie błędnych ocen zadań kontrolnych (źródło: CKE).

Jak można zauważyć analizując tabelę 5., najczęściej błędy oceny zadań kontrolnych występowały w zadaniu 1. (9,5% zastosowanych rozwiązań zadań kontrolnych), a najrzadziej w zadaniu 3. (4% rozwiązań zadań kontrolnych). Można wymienić trzy źródła, które przypuszczalnie leżą u podstaw takiego stanu. Po pierwsze, kompetencje merytoryczne i dydaktyczne z zakresu oceniania poszczególnych egzaminatorów. Po drugie, jakość zastosowanych na egzaminie zadań i po trzecie, jakość diagnostyczna zastosowanych zadań kontrolnych. Udzielenie odpowiedzi na pytanie, czy i w jakim stopniu wymienione źródła błędów miały wpływ na zaobserwowane zjawisko, wymaga dodatkowych badań jakościowych.

Tak jak już zaobserwowano na poszczególnych etapach przygotowania e-oceniania i jego wdrożenia, newralgicznym działaniem dla jakości oceniania okazało się dobre przygotowanie ekspertów z zespołu standaryzacyjnego wybierającego i oceniającego zadania weryfikujące (kontrolne). Ekspertów (sędziowie kompetentni) powołani do tego zadania powinni posiadać nie tylko duże doświadczenie ale i wiedzę pozwalającą oszacować, jakie rozwiązania uczniowskie posiadają potencjalnie dużą wartość diagnostyczną przydatną do kontroli jakości procesu oceniania. Jak podkreśla zespół ewaluacyjny sesji e-oceniania, ważne jest, aby ocenianie zadań wybranych jako weryfikujące odbyło się w dwóch etapach. I etap powinien polegać na samodzielnym punktowaniu odpowiedzi uczniowskich (bez wzajemnego porozumiewania się egzaminatorów), a II - na porównaniu ocen poszczególnych sędziów kompetentnych i omówieniu ewentualnych różnic w ocenie i następnie ustaleniu wspólnej wartości w odniesieniu do skali punktowej danego zadania.

6 PERSPEKTYWY E-OCENIANIA W POLSCE.

Przedstawiony proces przygotowania e-oceniania i wdrożenia go dla jednego tylko egzaminu ukazuje złożoność przedsięwzięcia. Kolejne etapy wymagały starannego opisu tradycyjnego oceniania zadań egzaminacyjnych. Okazało się to bardzo korzystne także dla istniejącej metodyki i organizacji dotychczasowego oceniania. Wymagało bowiem odpowiedzi na pytania, które nie były dotąd istotne. Można więc powiedzieć, że wdrożenie e-oceniania usprawniło i poprawiło jakość oceniania tradycyjnego. Doświadczenie zdobyte dzięki wprowadzeniu e-oceniania w egzaminie gimnazjalnym z matematyki stanowi punkt wyjścia do projektowania e-oceniania dla innych egzaminów.

Potencjalnie możliwe jest wdrożenie e-oceniania dla pozostałych egzaminów gimnazjalnych i egzaminów maturalnych. Wybór kolejności wdrażania uzależniony jest od kilku czynników. Jeden z ważniejszych czynników to możliwości techniczne, szczególnie w zakresie przetwarzania danych egzaminacyjnych. E-ocenianie wymaga określonych zasobów pamięci serwerów oraz Sieci o wymaganej szybkości transferu danych. Parametry serwerów i Sieci, dostateczne dla jednego egzaminu, nie wystarczą dla kilku egzaminów ocenianych jednocześnie w tej samej sesji. Konieczne więc będzie podjęcie decyzji o inwestowaniu w tej dziedzinie, bądź też o wykupowywaniu zasobów na serwerach zewnętrznych. Kolejny czynnik, to oczekiwany efekt poprawienia jakości oceniania. Konieczne jest przeprowadzenie analizy, dla których przedmiotów dotychczasowe efekty oceniania nie są zadowalające, a dzięki e-ocenianiu mogłyby zostać znacznie poprawione. Inny czynnik, to możliwości przygotowania arkuszy egzaminacyjnych z poszczególnych egzaminów w formie nadającej się do szybkiego przygotowania prac egzaminacyjnych do e-oceniania. To tylko kilka z szerszej gamy czynników. Obecnie prowadzone są prace studyjne i przygotowania pilotażu dla dwóch przedmiotów maturalnych, które mogłyby być oceniane przy wykorzystaniu e-oceniania.

PODSUMOWANIE

Wprowadzenie e-oceniania rewolucjonizuje ocenianie nie tylko z powodu zastosowania nowych narzędzi do pracy egzaminatora, ale także z powodu dostępności szerokiej gamy narzędzi do automatycznego raportowania, co umożliwi bieżące monitorowanie postępu i jakości oceniania oraz zwiększa widoczność (przezroczystość) wszystkich procesów sesji oceniania. Skonfigurowany w programie system kontroli jakości oceniania zadań otwartych rozwiniętej odpowiedzi zapewnia wysoki poziom porównywalności. Porównywalności poprzez utrzymywanie standardu oceniania dzięki zastosowaniu „wzorcowo” ocenionych rozwiązań – tzw. zadań kontrolnych do kwalifikowania egzaminatorów oraz do systematycznego monitorowania jakości ich pracy. Ponadto *scoris assessor* umożliwia prowadzenie szczegółowych analiz na poziomie zadań, grup zadań, arkuszy egzaminacyjnych i wyciąganie wniosków istotnych dla doskonalenia pracy twórców zadań egzaminacyjnych i jakości pracy egzaminatorów. Istotnym przesłaniem jest wykorzystanie tych wniosków do ciągłego i systematycznego doskonalenia oceniania zadań otwartych, które z natury swojej podatne są zarówno na błędy losowe, jak i systematyczne związane z efektem egzaminatora. Warto przy tym pamiętać, że to właśnie zadania otwarte posiadają szczególną wartość diagnostyczną i edukacyjną, która jest bardzo cenna dla procesu oceniania i procesu kształcenia.

Przeprowadzenie e-oceniania w 2015 roku z gimnazjalnego egzaminu z matematyki w czterech z ośmiu komisji okręgowych a następnie w całej populacji w 2016 roku potwierdziło obserwacje poczynione podczas pilotaży w ramach projektu. Wyeliminowane zostały niedogodności związane odbiorem, przekazywaniem egzaminatorom paczek z pracami uczniowskimi oraz bezpiecznym ich przechowywaniem i przesyłaniem do okręgowych komisji egzaminacyjnych. Egzaminatorzy zostali uwolnieni od wielu obowiązków administracyjnych w trakcie sesji oceniającej związanych z kodowaniem odpowiedzi na kartach, wpisywaniem identyfikatorów. Mogli też pracować w bardziej komfortowych warunkach i we własnym tempie podczas całej sesji oceniania. Mogli lepiej skupić się na analizie uczniowskich rozwiązań, przydzielając ocenę za rozwiązanie zadania w poczuciu odpowiedzialności za własną decyzję i w kontekście informacji zwrotnej uzyskiwanej na bieżąco w trakcie oceniania.

I na zakończenie jeszcze jedno stwierdzenie bardzo ważne z punktu widzenia trafności i rzetelności egzaminu. Wprowadzenie e-oceniania ułatwia ocenianie i czyni proces bardziej „przezroczystym”. Umożliwia łatwiejsze dostrzeganie jakości newralgicznych ogniw procesu. Ogniw, których jakość zależy od kompetencji ekspertów tworzących zadania, tworzących modele odpowiedzi i przygotowujących zadania kontrolne do diagnozowania, monitorowania i regulowania procesu oceniania w sesji.

BIBLIOGRAFIA

- Drasgow F., Mattern K., (2006). *New Test and New Items: Opportunities and Issues* [in:] Bartran D., Hambleton D. K., (Ed.), *Computer-Based Testing and the Internet. Issues and Advances*. John Wiley& Sons LTD, Chichester, West Sussex, England.
- Dudkiewicz-Świerzyńska M., Sławomir Sapanowski S., Zaborowski J., (2015), *Ewaluacja przebiegu E-oceniania w sesji 2015 r. Na podstawie raportów przewodniczących zespołów egzaminatorów*. Centralna Komisja Egzaminacyjna, Warszawa (niepublikowany raport).
- Jurdzińska, D., Modrzewska, E. (2015). *E-ocenianie jako innowacja technologiczna i jakościowa w systemie egzaminów zewnętrznych*. [w:] Niemierko, B., Szmigel M.K., (red.) *Zastosowania diagnozy edukacyjnej*. Kraków: gRUPA TOMAMI. 358-371.

Larkey L.S., Croft W.B., (2003). *A Text Categorization Approach to Automated Essay Grading* [in:] Shermis M.D., Burstein J., (Ed.) *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Lawrence Erlbaum Associates, Publishers, London.

Szaleniec H., (2007). *Refleksje na temat rozwoju systemu egzaminacyjnego*, [w:] Niemierko B., Szmigel M., K. (red.), *Uczenie się i egzamin w oczach uczniów*. Kraków: gRUPA TOMAMI.

Szaleniec H., Kondratek B., Kulon F., Pokropek A. Skórska p., Świst K., Wołodźko T., Żółtak M., (2015). *Efekt egzaminatora w ocenianiu prac maturalnych z języka polskiego i matematyki*. Warszawa Instytut Badań edukacyjnych.

Kulon, F., Żółtak, M. (2014). *Zróźnicowanie łagodności egzaminatorów między okręgowymi komisjami egzaminacyjnymi*. [w:] Niemierko, B., Szmigel M.K., (red.) *Diagnozy edukacyjne. Dorobek i nowe zadania*. Kraków: gRUPA TOMAMI.

KONTAKT

dr Henryk Szaleniec

Polskie Towarzystwo Diagnostyki Edukacyjnej

Wojciech Małecki

Okręgowa Komisja Egzaminacyjna we Wrocławiu,

Polskie Towarzystwo Diagnostyki Edukacyjnej

POSITIVE DIAGNOSIS IN A TEACHER'S/EDUKATOR'S WORK

POZITIVNÍ DIAGNÓZA V PRÁCI PEDAGOGA

Ewa Wysocka

Abstrakt: V článku je představen model pedagogické diagnózy v komplementárním pohledu, ve vztahu k předpokladům salutogeneze a patogeneze (pozitivní a negativní diagnóza). Tato koncepce byla základem pro sestavení modelu diagnózy a diagnostických nástrojů: *Měřítka Zvláštních Schopností (MZS)* a *Dotazník Osobnosti a Tvůrčího Myšlení (DOTM)* Edyty Charzyńskiej a Ewy Wysocké, a také *Dotazník Intrapersonálních a Interpersonálních Postojů vůči Světu a Životu (DIIPSŽ)* Ewy Wysocké. V článku jsou představeny teoretické a metodologické základy a podstatné psychometrické vlastnosti popsaných diagnostických nástrojů. Také jsou zdůvodněny možnosti jejich používání v pedagogické diagnóze.

Klíčová slova: pedagogická diagnóza, pozitivní diagnóza, negativní diagnóza, komplementární model diagnózy, diagnostické nástroje

Abstract: The present paper presents a model of diagnosis in pedagogy, as seen from the complementary perspective, i.e. with reference to the assumptions of the salutogenic and pathogenic approaches (positive and negative diagnosis). The concept provided a basis for developing a model of diagnosis and diagnostic instruments: *Special Abilities Scale (SAS)* as well as *the Identity and Creative Thinking Questionnaire (ICTQ)* devised by Edyta Charzyńska and Ewa Wysocka, *the Intrapersonal, Interpersonal and World – and Life-related Attitudes Questionnaire (IIWLQ)* devised by Ewa Wysocka. The author of this paper presented the theoretical and methodological basis for creating the aforementioned diagnostic instruments, and described their psychometric properties. It was also suggested that the instruments may prove useful in diagnosis in pedagogy.

Key words: diagnosis in pedagogy, positive diagnosis, negative diagnosis, complementary model of diagnosis, diagnostic instruments/tools

INTRODUCTION

One of the key factors impacting the effectiveness of teaching impact is a correct diagnosis. Both processes – diagnosis and post-diagnostic activity – should follow one theoretical and methodological paradigm, with diagnosis being an essential basis for designing teaching activities corresponding to its results. Both processes interpenetrate and are considered parallel.

The starting point in the conceptualization of areas, scope and subject of diagnosis made include the following – acceptable in pedagogy: a) the concept of personality as a system of interdependencies with the development of an individual treated as a variable process

impacted by a number of conflicting internal and external forces (Kaufmann 2004: 214); b) the concept of dynamic "system of attitudes towards the world" (Świda 1974, Marynowicz-Hetka 2006: 36) implied by the teaching and psychological perspective of an individual at the moment of transformation (creation, development process); c) the concept of *resilience* referring to processes and mechanisms which support correct (positive) functioning of an individual despite adversities, risk factors and threats to development, traumatic experiences occurring in their life – biography (Borucka 2011: 12; Borucka, Ostaszewski 2008; Christiansen, Evans 2005; Dezen 2007; Ferguson, Horwood 2003; Luthar, Cicchetti 2000; Luthar, Cicchetti, Becker 2000a,b; Luthar, Zelazo 2003; Ostaszewski, Zimmerman 2006; Rutter 2003; Taylor, Wang 2000; Zimmerman, Arunkumar 1994); d) the concept of salutogenesis by Aaron Antonovsky (1979, 1987, 1997, 2005), often treated as the breaking of the paradigm in studies on health and disorders of an individual and presented as a result of his research on people experiencing extreme stress levels (concentration camp, war-related experiences), who demonstrated a surprisingly good psychic health despite experiencing a long-lasting trauma. This defines the general subject of the diagnosis i.e. different conditions impacting the process of "an individual's transformation, and in particular the biopsychic development in a full life cycle in the natural, cultural and personal sense (Schulz 2004: 117, Marynowicz-Hetka 2006: 36).

What is important in the proposed approach – positive – are the various categories of an individual's diagnostic description: multidimensionality (different spheres of functioning), and the continuity (development process in the full cycle of life). Two directions of an individual's development are also taken into account: pro-social, pro-development vs. anti-social and destructive to development. It is assumed that for the process of change it is particularly important to identify potentials, both from the perspective of the efficiency of teaching activities, as well from the as humanist perspective and solving problems related to human development (positive approach and positive diagnosis).

Such an approach places the process of diagnosis in the positive psychology paradigm (Argyle 2004; Carr 2009; Czapiński 2004a,b; Seligman 1975, 1996, 2002, 2004) focused on creation and measures of life quality, primary understanding an individual, which makes it necessary to identify specific potentials of an individual (positive diagnosis).

1 POSITIVE DIAGNOSIS MODEL AND COMPLEMENTARITY OF THE APPROACH TO HUMAN DEVELOPMENT

While looking for theoretical references to justify the predominant role of positive diagnosis in teaching activity, one may refer to the concept of the development diagnosis by Irena Obuchowska (1983, 1997) who was the first to point out the importance of positive diagnosis. She integrated both approaches: "positive" (diagnosis of resources and potentials), and "negative" (diagnosis of deficits and disorders), but also pointed to the need to combine both approaches in one model.

While defining both concepts, I will focus only on the most important features. **Negative diagnosis**, which prevails in the current approaches to the diagnostic process, focuses on defining the essence of problems experienced by an individual. It exemplifies a theoretical **risk model** (*Risk-Need-Responsivity Model – RNR*) which assumes teaching activity in a dysfunctional area (Wysocka 2013, 2015, Andrews, Bonta 2003, Ward, Maruna 2007; Biel 2011: 150-153). It is related to the search for causes and mechanisms behind disorders and describes their consequences for an individual (their development) and environment (life environment). The diagnosis must be well preceded by a theoretical reflection on the nature

of disorders (symptomatology, etiology), the nature of the subject studied (development and social aspects), and adequate cognition methods (collection of information and its analysis). The obtained information is a basis for an initial diagnostic hypothesis concerning the experienced problem, its causes, mechanisms and development-related consequences (initial diagnosis). It must be noted that the evaluation conclusion is only a hypothesis which is sought to be confirmed in the activity and observation of a person studied "subjected" to various corrective activities. Diagnostic hypothesis is always made by referring the information obtained to the theoretical model adopted by the diagnostician (theoretical reflection). It is a basis for making the final evaluation conclusion (final diagnosis) which, however must be subject to constant verification in the course of further activities (post-diagnostic).

Positive diagnosis, often neglected, both in theoretical analyses and in teaching practice, focuses on identifying resources and potentials, various abilities and features of the subject studied which due to unfavorable development conditions (lack of support in development), could not emerge. However, they are an important factor for the process of change, fighting development crises and releasing autocreation mechanisms. Positive diagnosis reflects assumptions of the theoretical **Good Lives Model (GLM)**. The purpose of the positive diagnosis is the identification of positive features of an individual (predisposition), pro-development mechanisms and positive impact of the educational environment – the diagnosis focuses on the identification of an individual's strengths and their life environment (Wysocka 2013, 2015, Ward, Maruna 2007; Biel 2010: 153-156). The information collected is a basis for making a hypothetical positive diagnosis which must also be subjected to theoretical reflection in the context of the significance of identified states for an individual's development, and as regards its usefulness in designing post-diagnostic action.

Negative and positive diagnoses constitute a **complete diagnosis**, being a **complementary diagnosis model** in this approach. It is used for developing an intervention model based on two-way actions: eliminating deficits and strengthening potentials. According to the adopted principle, positive diagnosis is more important and significant for the process of designing intervention actions than a negative diagnosis. Positive diagnosis is based on factors that directly stimulate development whereas negative diagnosis focus on the elimination of factors hindering the development. This assumption is based on the statement that effective action supporting change focus on those elements that are efficient in an individual and develop in the right direction. However, what must not be neglected in the post-diagnostic plan, are actions improving the existing deficits still they must be of secondary or auxiliary nature (Obuchowska 1997: 13-14). The importance of positive diagnosis and actions stimulating and strengthening the identified potentials of an individual (individual's development), results from the fact that disorder symptoms cease to be functional which means that they cease to satisfy needs in a destructive way (if no other is available to an individual), and tend to decrease if obstacles to development are removed and an individual discovers creative ways to achieve their own goals and pursue personally meaningful values. Positive diagnosis falls within the humanist action model, being the result of theoretical assumptions of positive psychology with the concepts of *resilience* and *salutogenesis*.

It is worth mentioning that the problem of diagnosis complexity related to the principle of combining positive and negative diagnosis (the necessity to identify strengths and weaknesses of an individual, as well as conditions in which the development continues), is very important from the perspective of the phraseology of teaching activity since a failure to respect this principle limits the area of the planned modification activities and at the same

time reduces their efficiency. The principle of teaching activity is such an organization of this activity that is initially based on potentials (the principle of ergo tropic – activating actions based on a positive diagnosis) which should only be supplemented with actions directly eliminating symptoms of behavior dysfunctionality and living conditions (the principle of „succession” of semiotropic activities – eliminating disorder symptoms). It is also important to take into account those aspects of activity that are directed at the elimination of causative factors and thus on the elimination of the cause of disorders (significance of etiotropic activities and diagnosis of the etiology of disorders).

According to praxeological rules of corrective action, its effectiveness depends on the comprehensive diagnosis made both with respect to the level and type of disorders and dysfunctionality of the conditions of an individual’s development (symptomatological diagnosis – identification, causative and consequential, directed negatively at disorders, related to semiotropic and etiotropic activity), as well as the correct functioning of an individual in their life environment (diagnosis of potentials – identification, directed positively, related to ergotropic activity).

Thus post-diagnostic planning has to take into account the positive diagnosis as it is a basis for activities which directly eliminate disorders (Tokarczyk 1997: 54). In pedagogical practice an initial selective diagnosis (a basis for classification for specific intervention) includes mainly conclusions evaluating the scope of deficits of a subject of the research, usually without any information on their development potential. It restricts post-diagnostic planning with respect to ergotropic activities and additionally causes problems concerning the diagnostic and therapeutic relation (establishing and developing the relation), which is no less important from the point of view of intervention activities. Activity correcting deficits identifies difficulties in building a diagnostic and therapeutic relation, something that is absent from activities supporting the development of identified potentials. The relation is then based on trust and direct convincing of an individual that a diagnostician or a therapist (teacher) follows the principle of protecting their welfare and not protecting others (usually negatively viewed by a disturbed individual because of their own experience and a conviction that the world and people are bad; Wysocka, Ostafińska 2014).

Complementary model of diagnostic and intervention activities should thus be an example of both models: risk and good life, however the focus should be shifted towards positive diagnosis and intervention which underlines the importance of identification and reliance in the process of change on the resources (potentials) of a developing individual (and subject to the educational process), which reflects both the principles of efficient activity (praxeology of action), and humanist activity (subjectivity of an individual in the process of change). In other words, the good lives model (humanist, subjective), should be the starting point for a diagnosis and post- diagnostic planning, whereas the risk model may only play an auxiliary role in pedagogy, its purpose being mainly the identification and elimination of existing obstacles to development (chart 1).

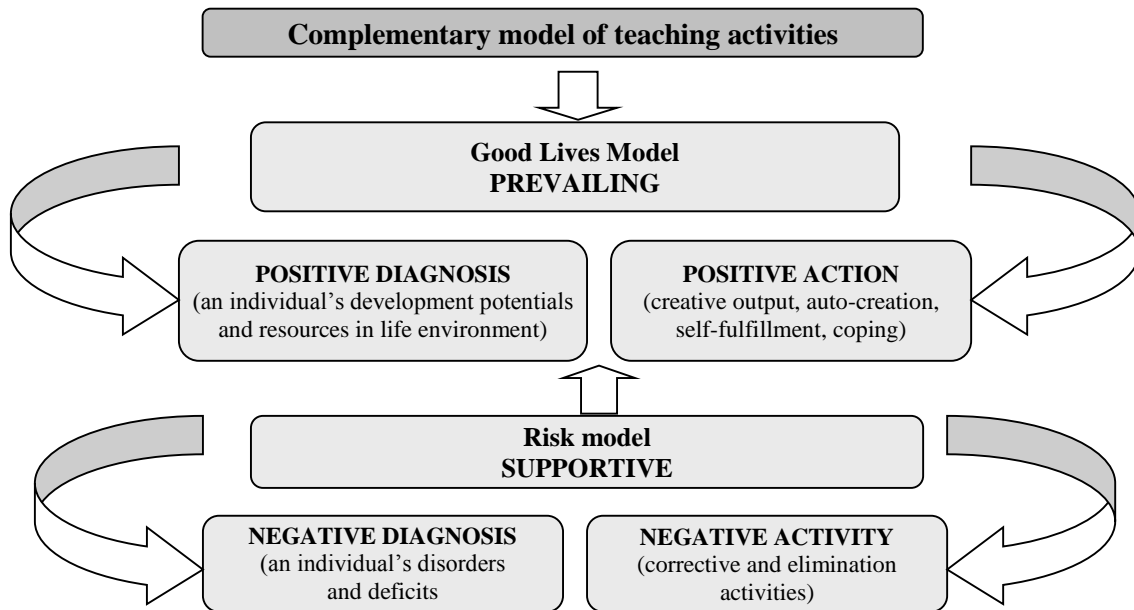


Chart 1 Complementary model of activities in pedagogy (developed by the author)

2 POSITIVE DIAGNOSIS – PROPOSED DIAGNOSTIC TOOLS

Referring to the theories mentioned in the introductory section (personality as a system of interdependencies "I – the world", a system of attitudes towards the world, salutogenesis, *resilience*, risk model, good lives model), the most important areas of the complementary diagnosis (positive and negative), may, in the cognitive approach, be presented in the following way (chart 2).

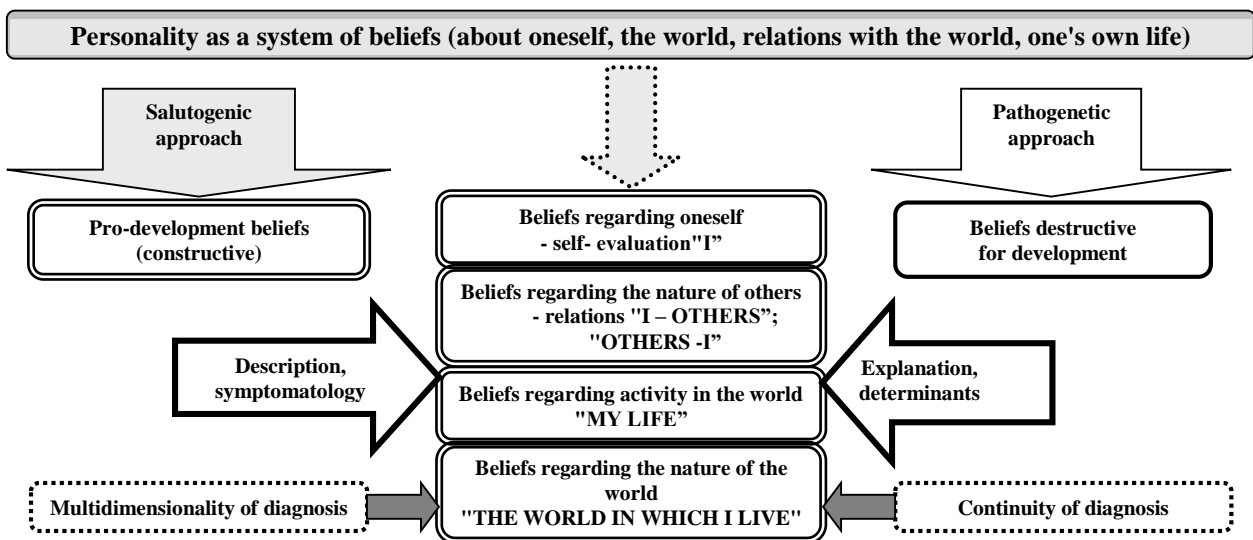


Chart 2 Schematic overview of diagnosis from the cognitive perspective (author's own study, Wysocka 2015, p. 163)

The adoption of specific theoretical assumptions regarding the phenomenon studied (e.g. cognitive concepts), determines the methodology of its cognition i.e. the use or creation of specific methods and diagnostic tools (e.g. positive, quality or quantity approach). Based on the assumption that the positive approach prevails in the pedagogical diagnosis, several versions of 3 diagnostic tools were created to measure (quantitative approach) beliefs about one's own personality and potentials: **1. Intrapersonal, Interpersonal and World – and Life-related Attitudes Questionnaire (IIWLQ)** devised by Ewa Wysocka (2011), Joanna Aksman (Aksman, Wysocka 2011), Bartłomiej Gołek (Gołek, Wysocka 2011) and Joanna Góźdź (Wysocka, Góźdź 2011); **2. Special Abilities Scale (SAS)**, devised by Edyta Charzyńska and Ewa Wysocka (2015b); **3. Identity and Creative Thinking Questionnaire (ICTQ)** devised by Edyta Charzyńska and Ewa Wysocka (2015a). The procedure for creating the tools was in all cases analogous but not identical. **Rational strategy** (deductive, theoretical) was applied to generate test items, with the starting point being a specific theoretical concept and hypotheses concerning the structure of the measured construct (variable). Two-stage expert evaluation (1 stage - qualification of test items to individual categories of variables; 2 stage – evaluation of instructions, response format, developmental adjustment, validity of item content). The number of items was reduced at both stages. The research was conducted in stages: **(1)** pre-pilot phase – a study among students, one class selected at random from a given type of school (primary, junior high school, high school); students played the role of expert judges evaluating the difficulty of items, correspondence of items to their life situation, comprehensibility of statements; **(2)** pilot phase study –conducted on a sample of 100-150 students attending each type of school, in order to receive further feedback from students on the scale structure; primary statistics were prepared: descriptive, discriminant power and reliability of scales; further reducing test items; **(3)** proper research facilitating the final verification of accuracy, reliability and discriminant power of a tool was conducted all over Poland on samples of approx. 1000 individuals. The final structure of tools (number of variables) was defined based on the explorative factor analysis, whereas the structure was confirmed by means of confirmative factor analysis (it does not apply to the first tool in the case of which the confirmative analysis was not performed). Tools were normalized (standard ten, separately for both genders and with age taken into account). The tools are of a self-descriptive character (are based on self-consciousness, *self-report* type tools), and are a pencil and paper method.

1. *Intrapersonal, Interpersonal and World – and Life-related Attitudes Questionnaire* (IIWLQ) is used to diagnose personality understood as a system of beliefs – cognitive approach (Epstein 1985, 1990a, b, 1991a,b, 2003, 2006) – makes it possible to make measurements in the area of: (a) **intrapersonal attitudes** understood as the image of oneself and global self-assessment (a system of beliefs concerning one's own "I") comprised of a general self-evaluation (non-specific), and partial self-evaluations (specific - cognitive and intellectual, physical, social and moral and characterological spheres); (b) **interpersonal attitudes** understood as an image of other people and relations with them (beliefs about functioning in interpersonal relations): "*others towards me*" (support from others and appreciation vs. indifference of others and lack of appreciation; threat from others vs. sense of security "*me towards others*" (public spirit, altruism, sociability vs. egocentrism, alienation, aggressiveness vs. lack of aggressiveness); (c) **attitudes towards the world** understood as the image of the world (beliefs concerning the sense and kindness of the world towards people); (d) **attitudes towards one's own life** understood as the image of one's own life (sense of effectiveness and control vs. sense of helplessness). Surveys are used for the self-study of children and youth at 4 stages of development – 4 diagnostic

tools: primary school – grades: 1-3¹; grades 4-6, junior high school, high school. Psychometric properties of all versions of tools are satisfactory but differentiated for individual stages of development. Reliability measured with Cronbach's alpha internal consistency coefficient is presented in table 1.

Table 1 Cronbach's alpha internal consistency coefficient

Scales		Cronbach's alpha			
		High school	Junior high school	Primary school grades IV-VI	Primary school grades I-III
General self-evaluation, non-specific		0.836	-	-	-
Detailed self-evaluation, specific		0.852	0.719	0.757	0.68
Interpersonal functioning		0.832	0.731	0.774	0.77
Image of the world		0.675	0.636	0.668	0.41
Image of life	General	0.834	0.688	0.675	0.71
	Situations at school	-	-	-	0.63
	Situations outside school	-	-	-	0.55
Scale of lie		0.700	-	-	-
General		0.940	0.874	0.895	0.85

Author's own study, et alli: Wysocka (2011), Wysocka, Gózdź (2011), Aksman, Wysocka (2011), Golek, Wysocka (2011)

2. *Special Abilities Scale* (SAS), devised by Edyta Charzyńska and Ewa Wysocka (2015b), comprises 10 sub-scales for evaluation of special skills, performed based on self-evaluation (*self-report*-type tool). Description of special skills was adopted from Howard Gardner (1983, 1995, 1999a,b, 2009a,b), Kenneth W. Tupper (2002), Danah Zohar and Ian Marschall (2001) and Robert Emmons (1999, 2000), and distinguishes 10 categories i.e. logical-mathematical, intrapersonal, musical, visual-spatial, naturalist, spiritual, interpersonal, existential and linguistic. It was assumed that these abilities are manifested by willingness, preferences and passions and thus the subjects studied have insight in them. Results are calculated separately for all 10 scales (abilities). Psychometric properties are acceptable, and the reliability measured with *Cronbach's alpha* internal consistency coefficient is presented in table 2.

Table 2 Cronbach's alpha internal consistency co-efficient

Abilities	Cronbach's alpha	
	Junior high school	High school
Logical-mathematical	0.883	0.885
Musical	0.847	0.882
Existential	0.817	0.859
Naturalist	0.820	0.854
Interpersonal	0.785	0.804
Intrapersonal	0.762	0.788
Linguistic	0.691	0.774
Kinesthetic	0.623	0.672
Visual-spatial	0.634	0.615
Spiritual	0.708	0.845

Author's own study, et alli: Charzyńska, Wysocka (2015b)

3. *Identity and Creative Thinking Questionnaire* (ICTQ) devised by Edyta Charzyńska and Ewa Wysocka (2015a) is used for the analysis of creative abilities (creative thinking), and creative personality, whereas from the methodological point of view it is based on the idea of potential creativity⁸. Holistic concept of creativity developed by Klaus K. Urban (1990, 2003, 2004) and Hans G. Jellen (Urban, Jellen 1986), understood as a set of intellectual skills and motivational and personality properties, was applied. Only the first two elements analyzed as personal beliefs regarding one's potential were taken into account from the three-element model of creativity and its conditions: (1) cognitive factors, (2) personality factors, (3) environment-related factors). The developed tool has two parts used for evaluation: a) cognitive-intellectual (creative thinking), b) personality (creative personality traits and its correlates), performed based on self-evaluation. The tool comes in two versions: a) for junior high school students, b) for high school students and potentially for other types of education above high school level. Final scales for creative personality include: a) **version for junior high school** - involvement and durability, readiness for risk-taking, entrepreneurship, transgression, strong ego, empathy, high self-evaluation, b) **version for high school** - readiness for risk taking, involvement and durability, transgression, strong ego, openness, non-conformity, entrepreneurship. Creative thinking includes the following scales: divergent thinking, spatial-visual thinking, explorative thinking (analogous in both cases). Scores are calculated for general dimensions: creative personality (conditions determining the emergence of creative potential), and creative thinking (creative potential available and perceived by the subject studied), and for 7 traits of creative personality and 3 types of creative thinking. Psychometric properties of the tools are acceptable and reliability measured with *Cronbach's alpha* internal consistency coefficient is presented in table 3.

Table 3 Cronbach's alpha internal consistency coefficient

Creative personality	Cronbach's alpha		Creative thinking	Cronbach's alpha	
	Junior high school	High-school		Junior high school	High school
Involvement and durability	0.844	0.741	Divergent	0.765	0.814
Readiness for risk taking	0.843	0.904	Visual transformational	0.746	0.753
Entrepreneurship	0.820	0.663	Explorative	0.656	0.685
Transgression	0.777	0.763	CREATIVE THINKING	0.832	0.849
Strong ego	0.764	0.794			
Empathy	0.798	-			
High self-esteem	0.799	-			
Openness,	-	0.783			
Non-conformity	-	0.690			
CREATIVE PERSONALITY	0.901	0.899			

Author's own study, et alli: Charzyńska, Wysocka (2015a)

⁸ Two main trends are distinguished in studies on creativity: a) effective creativity (actual, real), connected with analyzing works with respect to creative elements manifested therein; b) potential creativity connected with studying creative skills - creative thinking and creative personality.

The tools developed fill the gap existing in the area of positive pedagogical diagnosis concerning the potentials of the individuals studied (and indirectly deficits as well), facilitate insight into one's own potentials which condition the selection of the educational and vocational path and make it possible to adjust the educational methodology to a student's personality profile and their skills. They may be used in individual diagnosis (case diagnosis), and in scientific research as well.

FINAL CONCLUSION

A teacher makes a diagnosis to provide effective help i.e. to support development, but to make it feasible, the diagnosis must be accurate and reliable. Teacher's diagnostic technique despite its apparent diversity does in fact lack diagnostic tools verified psychometrically which may only constitute its part but still a very important one. Teaching tools, which meet the criteria of theoretical and methodological "correctness", refer in principle to a diagnosis that is strictly educational, school-related (mainly knowledge and skills tests), whereas the process of pedagogy development, which determines elaboration and emergence of new areas of interest (new disciplines and sub-disciplines), creates new diagnosis-related needs. It must be noted that at least in Poland the diagnostic technique of a teacher is still not developed well enough, yet the accuracy of the diagnosis made depends on the quality of information obtained which requires the use of specific methods to obtain it. There is an obvious shortage of such methods, and thus the attempt to develop tools verified psychometrically and based on standards is an unquestionable and urgent need in the field of pedagogy. The tools developed meet methodological requirements of objectivity, accuracy and reliability, as well as a theoretical requirement – a reference to an individual's potential and thus create a basis for the positive diagnosis which is a priority for the pedagogical activity and determines its effectiveness (as claimed in the introductory section). Furthermore, the theoretical accuracy of the presented model of diagnosis and diagnostic tools also depends on their placement in cognitive concepts. While studying of attitudes and self-perception (potentials and difficulties experienced, perception of oneself, the world, one's own life and place in the world), we refer to concepts (cognitive) which have so far credibly explained the functioning and development of an individual. The proposed theoretical and methodological approach they are based on seems to be justified, and the tools developed may be used as non-specific in various disciplines of pedagogy: in educational – teaching, educational and care, social, rehabilitation or even special diagnostics.

REFERENCES

- Aksman, J., & Wysocka, E. (2011). *Kwestionariusz Nastawień Intrapersonalnych, Interpersonalnych i Nastawień wobec Świata (KNIIS)*. Podręcznik testu – wersja dla uczniów szkoły podstawowej klas VI-VI. Kraków: MEN.
- Andrews, D.A., & Bonta, J. (2003). *The psychology of criminal conduct*. Cincinnati, Ohio: Anderson Publishing Blackburn.
- Antonovsky, A. (1979). *Health, stress and coping: new perspectives on mental and physical well-being*. San Francisco: Jossey Bass.
- Antonovsky, A. (1987). *Unraveling the mystery of health. How people manage stress and stay well*. San Francisco: Jossey Bass.

- Antonowsky, A. (1997). *Poczucie koherencji jako determinanta zdrowia*. In: I. Heszen-Niejodek, H. Sęk (eds.), *Psychologia zdrowia* (pp. 206-231). Warszawa: Państwowe Wydawnictwo Naukowe. Translated by D. Gawlikowska, J. Łuczyński, M. Olejnik.
- Antonowsky, A. (2005). *Rozwikłanie tajemnicy zdrowia. Jak radzić sobie ze stresem i nie zachorować*. Warszawa: Wydawnictwo Instytutu Psychiatrii i Neurologii. Translated by H. Grzegołowska-Klarkowska.
- Argyle, M. (2004). *Psychologia szczęścia*. Wrocław: Wydawnictwo Astrum. Translated by N. Oparska.
- Biel, K. (2011). *Zła resocjalizacja czy resocjalizacja zła*. In: K. Biel, M. Sztuka (eds.), *Resocjalizacja wobec tajemnicy zła* (pp. 133-160). Kraków: Wydawnictwo WAM, „Ignatianum”.
- Borucka, A. (2011). *Koncepcja resilience. Podstawowe założenia i nurty badań*. In: W. Junik (ed.), *Resilience. Teoria – badania – praktyka* (pp. 11-28). Warszawa: Wydawnictwo Edukacyjne PARPAMEDIA.
- Borucka, A., & Ostaszewski, K. (2008). Koncepcja resilience. Kluczowe pojęcia i wybrane zagadnienia. *Medycyna Wieku Rozwojowego*, 2(12), 587-597.
- Carr, A. (2009). *Psychologia pozytywna. Nauka o szczęściu i ludzkich siłach*. Poznań: Zysk i S-ka. Translated by Z.A. Królicki.
- Charzyńska, E., & Wysocka, E. (2015a). *Kwestionariusz Osobowości i Myślenia Twórczego (KOMT). Podręcznik testu – książka użytkownika (wersje dla uczniów gimnazjum i liceum)*. Katowice: Fundacja Pomocy Osobom Niepełnosprawnym.
- Charzyńska, E., & Wysocka, E. (2015b). *Skala Zdolności Specjalnych. Podręcznik testu – książka użytkownika (wersje dla uczniów gimnazjum i liceum)*. Katowice: Fundacja Pomocy Osobom Niepełnosprawnym.
- Christiansen, E.J., & Evans, W.P. (2005). Adolescent Victimization. Testing Models of Resiliency by Gender. *Journal of Early Adolescence*, 3(25), August, 298-316.
- Czapiński, J. (2004a). Spotkanie dwóch tradycji: hedonizmu i eudajmonizmu. In: J. Czapiński (ed.), *Psychologia pozytywna. Nauka o szczęściu, zdrowiu, sile i cnotach człowieka* (pp. 13-17). Warszawa: Państwowe Wydawnictwo Naukowe.
- Czapiński, J. (2004b). Wstęp. In: J. Czapiński (ed.), *Psychologia pozytywna. Nauka o szczęściu, zdrowiu, sile i cnotach człowieka* (pp. 7-10). Warszawa: Państwowe Wydawnictwo Naukowe.
- Dezen, K.B.S. (2007). *Promoting resilience among maltreated youth: Policy and procedural implications*. Washington: Georgetown Public Policy Institute (A Thesis submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University in partial fulfillment of the requirements for the degree of Master of Public Policy in the Georgetown Public Policy Institute; April 16, 2007).
- Emmons, R.A. (1999). *The psychology of ultimate concerns: motivation and spirituality in personality*. New York: Guilford Press.
- Emmons, R. (2000). Is spirituality an intelligence? Motivation, cognition, and the psychology of ultimate concern. *International Journal for the Psychology of Religion*, 10(1), 3-26.
- Epstein, S. (1985). The implications of cognitive-experimental self-theory for research in social psychology and personality. *Journal for the Theory of Social Behaviour*, 15(3), October, 283-310.

- Epstein, S. (1990a). Cognitive-experiential self-theory. In: L.A. Pervin, (ed.), *Handbook of personality and research: Theory and research* (pp. 165-192). New York: Guilford Publications, Inc.
- Epstein, S. (1990b). Wartości z perspektywy poznawczo-przeżyciowej teorii „ja”. In: J. Reykowski, N. Eisenberg, E. Staub (eds.), *Indywidualne i społeczne wyznaczniki wartościowania* (pp. 11-32). Wrocław–Warszawa–Kraków–Gdańsk–Łódź: Zakład Narodowy im. Ossolińskich, Wydawnictwo PAN.
- Epstein, S. (1991a). Cognitive-experiential self-theory. Implications for developmental psychology. In: M.R. Gunnar, L.A. Sroufe (eds.), *Self-processes and development* (vol. 23, pp. 79-123). Minnesota Symposia on Child Psychology Series. Hillsdale, NJ: Erlbaum Associates.
- Epstein, S. (1991b). Cognitive-experiential self-theory. An integrative theory of personality. In: R. Coleman Curtis (ed.), *The relational self: Convergences in psychoanalysis and social psychology* (pp. 111-137). New York: Guilford Press.
- Epstein, S. (2003). Cognitive-experiential self-theory of personality. In: T. Millon, M.J. Lerner (eds.), *Handbook of psychology: Personality and social psychology* (vol. 5, pp. 159-184). Hoboken, NJ: John Wiley and Sons Inc.
- Epstein, S. (2006). Conscious and unconscious self-esteem from the perspective of cognitive-experiential self-theory. In: M.H. Kernis (ed.): *Self-esteem: Issues and answers. A sourcebook of current perspectives* (pp. 69-76). New York and Hove: Psychology Press, Taylor and Francis Group.
- Fergusson D.M., & Horwood L.J. (2003). Resilience to childhood adversity: Results of a 21 year study. In: S.S. Luthar (ed.), *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 130-155). New York: Cambridge University Press.
- Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences*. New York: Basic Books.
- Gardner, H. (1995). *Leading minds: An anatomy of leadership*. New York: Basic Books.
- Gardner, H. (1999a). Are there additional intelligences? In: J. Kane (ed). *Education, information, transformation: Essays on learning and thinking* (pp. 111-131). Upper Saddle River, NJ: Prentice-Hall.
- Gardner, H. (1999b). *Intelligence reframed: multiple intelligences for the 21st century*, New York: Basic Books.
- Gardner, H. (2009a). *Inteligencje wielorakie. Nowe horyzonty w teorii i praktyce*. Warszawa: MT Biznes, Laurum. Translated by A. Jankowski.
- Gardner, H. (2009b). *Pięć umysłów przyszłości*, Warszawa: MT Biznes. Translated by D. Bakalarz.
- Gołek, B., & Wysocka, E. (2011). *Kwestionariusz Nastawień Intrapersonalnych, Interpersonalnych i Nastawień wobec Świata (KNIIS). Podręcznik testu – wersja dla uczniów szkoły gimnazjalnej*. Kraków: Ministerstwo Edukacji Narodowej.
- Kaufmann, J-C (2004). *Ego: socjologia jednostki. Inna wizja człowieka i konstrukcji podmiotu*. Warszawa: Oficyna Naukowa. Translated by K. Wakar.
- Luthar, S.S., & Cicchetti, D. (2000). The construct of resilience: implication for intervention and social policies. *Development and Psychopathology*, 12(4), 857-885.

- Luthar, S.S., & Cicchetti, D., Becker, B. (2000a). Research on resilience: reply to commentaries. *Child Development*, 71(3), 573-575.
- Luthar, S.S., & Cicchetti, D., Becker, B. (2000b). The construct of resilience: a critical evaluation and guidelines for future work. *Child Development*, 71(3), 543-562.
- Luthar, S.S., & Zelazo, L.B. (2003). Research on resilience: An integrative review. In: S.S. Luthar (ed.) *Resilience and vulnerability: Adaptation in the context of childhood adversities* (pp. 510-550). New York: Cambridge University Press.
- Marynowicz-Hetka, E. (2006). *Pedagogika społeczna* (vol. 1). Warszawa: Wydawnictwo Naukowe PWN.
- Obuchowska, I. (1983). *Dynamika nerwic. Psychologiczne problemy zaburzeń nerwicowych u dzieci i młodzieży*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Obuchowska, I. (1997). Diagnoza psychologiczna w poradniach psychologiczno-pedagogicznych. *Problemy Poradnictwa Psychologiczno-Pedagogicznego*, 2(7), 5-15.
- Ostaszewski, K., & Zimmerman, M.A. (2006). The effects of cumulative risks and promotive factors on urban adolescent alcohol and other drug use: A longitudinal study of resiliency. *American Journal of Community Psychology*, 3(38), 237-249.
- Rutter, M. (2000). Resilience reconsidered: conceptual considerations, empirical findings and policy implications. In: J.P. Shonkoff, & S.J. Meisels (eds.), *Handbook of early childhood intervention* (second edition, pp. 651-682). New York: Cambridge University Press.
- Schulz, R. (2004). Pedagogika jako nauka o człowieku: próba dookreślenia idei. In: T. Lewowicki (ed.), *Pedagogika we współczesnym dyskursie humanistycznym* (pp. 113-124). Warszawa-Kraków: Oficyna Wydawnicza „Impuls”.
- Seligman, M.E.P. (1975). *Helplessness*. San Francisco: Freeman.
- Seligman, M.E.P. (1996). *Optymizmu można się nauczyć*. Poznań: Media Rodzina. Translated by A. Jankowski.
- Seligman, M.E.P. (2002). *Authentic happiness: using the new positive psychology to realize your potential for lasting fulfillment*. New York: Free Press.
- Seligman, M.E.P. (2004). *Psychologia pozytywna*. In: J. Czapiński (ed.), *Psychologia pozytywna. Nauka o szczęściu, zdrowiu, sile i cnotach człowieka* (pp. 18-32). Warszawa: Państwowe Wydawnictwo Naukowe. Translated by J. Radzicki.
- Szmidt, K.J. (2013). *Pedagogika twórczości*. Sopot: GWP.
- Świda, H. (1974). Osobowość jako system ustosunkowań wobec świata. *Studia Pedagogiczne* (vol. 30, pp. 151-201). Wrocław: Ossolineum.
- Taylor, R.D., & Wang, M.C. (eds.) (2000). *Resilience across contexts: family, work, culture, and community*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Tokarczyk, E. (1997). Diagnoza dla celów resocjalizacji – specyfika i wynikające z niej trudności, *Problemy Poradnictwa Psychologiczno-Pedagogicznego*, 2(7), 52-54.
- Tupper, K.W. (2002). Entheogens and existential intelligence: the use of plant teachers as cognitive tools. *Canadian Journal of Education*, 27(4), 499-516.
- Urban, K.K. (1990). Recent trends in creativity research and theory in Western Europe. *European Journal of High Ability*, 1, 99-113.

Urban, K.K. (2003). Toward a componential model of creativity. In: D. Ambros, & L.M. Cohen, A.J. Tennenbaum (eds.), *Creative intelligence. Toward theoretic integration* (pp. 81-112). Cresskill, NJ: Hampton Press Inc.

Urban, K.K. (2004). Assessing Creativity: The Test for Creative Thinking – Drawing Production (TCT-DP). The concept, application, evaluation, and international studies. *Psychology Science*, 46(3), 387-397.

Urban, K.K. (2005). Assessing creativity: The Test for Creative Thinking – Drawing Production (TCT-DP). *International Education Journal*, 6(2), 272-280.

Urban, K.K., Jellen H.G. (1986). Assessing creative potential via drawing production: The Test for Creative Thinking-Drawing Production (TCT-DP). In: A.J. Cropley, & K.K. Urban, H. Wagner, W.H. Wiczerkowski (eds.), *Giftedness: A continuing worldwide challenge* (pp. 182-169). New York, NY: Trillium Press.

Ward, T., & Maruna, S. (2007). *Rehabilitation: beyond the risk paradigm*. London-New York: New York University Press.

Wysocka, E. (2011). *Kwestionariusz Nastawień Intrapersonalnych, Interpersonalnych i Nastawień wobec Świata (KNIIS)*. Podręcznik testu – wersja dla uczniów szkoły ponadgimnazjalnej. Kraków: MEN.

Wysocka, E. (2013). *Diagnostyka pedagogiczna. Nowe obszary i rozwiązania*. Kraków: Oficyna Wydawnicza „Impuls”.

Wysocka, E. (2015). *Diagnoza pozytywna w resocjalizacji. Model teoretyczny i metodologiczny*. Katowice: Wydawnictwo UŚ.

Wysocka, E., & Gózdź, J. (2011). *Kwestionariusz Nastawień Intrapersonalnych, Interpersonalnych i Nastawień wobec Świata (KNIIS)*. Podręcznik testu – wersja dla uczniów szkoły podstawowej klas I-III. Kraków: MEN.

Wysocka, E., & Ostafińska-Molik, B. (2014). Nastawienia życiowe młodzieży niedostosowanej społecznie i prawidłowo przystosowanej – analiza porównawcza. *Przegląd Naukowo-Metodyczny, Edukacja dla Bezpieczeństwa*, 22(1), 233-254.

Zimmerman, M.A., & Arunkumar, R. (1994). Resiliency research: Implications for schools and policy. *Social Policy Report. Society for Research in Child Development*, 4(8), 1-18.

Zohar, D., & Marshall, I. (2001). *Inteligencja duchowa: najwyższa z inteligencji*. Poznań: Rebis. Translated by P. Turski.

KONTAKT

Ewa Wysocka, dr hab. Prof. UŚ

University of Silesia, Poland, Faculty of Education and Psychology, Department of Education Theory, Grażyńskiego 53, 40-126 Katowice

48 (32) 3599 790; 48 519 694 975

ewa-wysocka@hotmail.pl; ewa.wysocka@us.edu.pl

DIAGNOZA I EWALUACJA W POLSKICH SZKOŁACH. CO JEST DOBRE, A NAD CZYM NALEŻY PRACOWAĆ?

DIAGNOSIS AND EVALUATION IN POLISH SCHOOLS. WHAT IS GOOD AND WHAT WE NEED TO WORK ON?

Maria Kocór

Abstrakt: W podjętym temacie chodzi o przypomnienie głębokiego sensu diagnozy i ewaluacji w szkołach różnego typu i szczebla w każdym miejscu i czasie, jakim jest lepsza i bardziej przyjemna oraz skuteczna edukacja. Diagnoza służy projektowaniu działań edukacyjnych, zaś ewaluacja ocenie efektów, ich trafności i skuteczności. Efekty edukacji są widoczne w kompetencjach uczniów czy studentów. Kompetencje te to z kolei gotowość do skutecznego działania czy zachowania się w różnych sytuacjach. Zatem jeden cel diagnozy i ewaluacji w szkołach towarzyszy – wyższa jakość i skuteczność działań. Czy jednak w polskich szkołach tak jest? Po nakreśleniu celów i zamierzeń diagnozy i ewaluacji autorka zastanawia się nad stanem rzeczywistym w szkołach polskich – nad kryteriami, standardami, kompetencjami nauczycieli, itp. Poszukuje mocnych stron diagnozy i ewaluacji w polskich szkołach oraz tych, nad którymi należy pracować. Przykładów trudności i ograniczeń diagnozy i ewaluacji dla wyższej jakości edukacji w Polsce jest wiele, ale są też sukcesy i przykłady dobrych praktyk, które należy promować. Podjęte rozważania są tylko próbą wzniesienia krytycznej dyskusji i niepokoju dla poszukiwania lepszych jakościowo stanów.

Klíčová slova: szkoła, uczeń, nauczyciel, diagnoza, ewaluacja, kompetencje, jakość

Abstract: In the undertaken subject author wants to remind about deep meaning of diagnosis and evaluation in schools which is better, more comfortable and effective. Diagnosis is used to design educational actions, and evaluation to grade it's efficiency. The effects of the education can be seen in student's competences. In turn, these competences are a readiness for effective actions or behaving in different situations. Therefore diagnosis and evaluation are accompanied by one goal - higher quality and effectiveness of actions. However, is this like in polish schools? After sketching the diagnosis and evaluation goals and assumptions the author wonders about the actual state. She looks for strong points of the diagnosis and evaluation in polish schools and these, which needs some work. There are lots of the examples of difficulties and restrictions of diagnosis and evaluation, but there are also successes and instances of good practices, which should be promoted. Undertaken considerations are only an attempt to raise a critical discussion and anxiety in order to search for higher quality states.

Key words: school, student, teacher, diagnosis, evaluation, competence, quality

WPROWADZENIE

Każde działanie wymaga namysłu, rozeznania ograniczeń, trudności i możliwości, jakie mogą mu towarzyszyć. Rozpoznanie to jest potrzebne zarówno na starcie, jak i w trakcie działania konieczne jest obserwowanie innych i siebie, uświadamianie własnych odczuć i emocji, itp. Jest jeszcze jeden aspekt postrzegania, gromadzenia informacji o czyjejs lub własnej działalności czy zachowaniu, które ma miejsce po nim, kiedy zastanawiamy się, czy dobrze zrobiliśmy, czy nam lub innym udało się, powiodło, a jeśli nie, to jaki był tego powód. Takie poszukiwanie, namysł nad obecną, przeszłą lub mającą nastąpić sytuacją czy zdarzeniem jest zbliżone do pojęcia diagnozy, choć do końca nim nie jest. Chciałam jednak tymi słowami zaakcentować, jak ważne jest rozpoznawanie warunków, w jakich działamy, własnych reakcji i emocji oraz potrzeb i możliwości osób, z którymi lub, dla których przyszło nam pracować.

Przypomnieć należy, że *diagnoza* pierwotnie była pojęciem medycyny jako rozpoznanie choroby. Ogólnie jednak pochodzi od greckiego słowa *diagnosis* i oznacza *rozpoznanie, rozróżnienie* w oparciu o zebrane dane i ich krytyczną analizę (za: Mazurkiewicz, 2003, s. 52). Mówi się najprościej, że znając symptomy zjawiska, sytuacji bądź zachowania, diagnosta poszukuje czynników sprawczych uświadamiając sobie, które z nich są sterowalne, a które należy tylko wkalkulować w działanie pedagogiczne, czyli uwzględnić stan zdrowia ucznia, jego specjalne potrzeby stosując odmienne metody. Dlatego wyróżnia się diagnozę poznawczą, której wyniki służą wyjaśnianiu i upowszechnianiu zebranych informacji i diagnozę decyzyjną służącą podejmowaniu decyzji (Mazurkiewicz, 2003). Bliżej różne modelowe ujęcia, klasyfikacje i ich opisy oraz obszary problemowe i zakres diagnozy realizowanej w obrębie różnych subdyscyplin pedagogicznych, w tym głównie w odniesieniu do szkoły i innych wychowawczych środowisk znajdziemy w pracach specjalistycznych wielu autorów (Radlińska, 1961; Ziemiński, 1973; Konarzewski, 1999; Obuchowska, 2002; Nikitorowicz, 2006; Jarosz, 2001; Jarosz i Wysocka 2006; Wysocka, 2007, 2013; Niemierko, 2009; Mazurkiewicz, 2003; Mizerek, 2010-2011; Ciężka, 2010-2011; Skałbania, 2011; Barłóg, 2008, 2011, 2015). Na podstawie ich krytycznej analizy i własnych przemyśleń oparłam niniejsze rozważania. Odnoszą się one do poszukiwania odpowiedzi na pytania: Czym jest diagnoza i ewaluacja w edukacji i, na czym ich sens się opiera? Co je łączy a co różni i, jaki jest ich stan w polskiej szkole? Jakie można przypisać im mocne i słabe strony? Co mogą wnieść dobrego dla ucznia i, przed jakimi praktykami go chronić by nauka nie była oparta na współzawodnictwie, ale na współdziałaniu i współpracy?

1 ISTOTA I SENS DIAGNOZY ORAZ EWALUACJI W EDUKACJI

1.1 Diagnoza w pedagogice i jej założenia teoretyczne

W pedagogice należącej do nauk humanistycznych i społecznych trudno precyzyjnie określić przedmiot diagnostyki pedagogicznej. Niewątpliwie, jest subdyscypliną, która zajmuje się diagnozowaniem, diagnozą wychowawczą i edukacyjną w zależności od tego, czy dotyczy pedagogiki przedszkolnej czy wczesnoszkolnej, pedagogiki opiekuńczej czy resocjalizacyjnej. Możemy mówić o diagnozie szkolnej, środowiskowej i diagnozie rodziny jako ważnego działu pedagogiki społecznej, szkolnej lub analogicznie pedagogiki rodziny.

W pedagogice społecznej i opiekuńczej wiele miejsca poświęca się diagnozie środowiskowych uwarunkowań rozwoju człowieka. Mówi się o potrzebie diagnozy środowiska: rodzinnego, szkolnego, rówieśniczego czy szerzej lokalnego, odkrywaniu i aktywizowaniu jego sił ku rozwojowi jednostki, rodziny czy tzw. *Małej Ojczyzny*. Wskazuje się głównie prakseologiczny i rozwojowy sens diagnozy, która jest warunkiem projektowania działań dydaktycznych, wychowawczych, opiekuńczych, resocjalizacyjnych, itp. w każdym

miejscu i czasie. Akcentuje się sens diagnozy pełnej, całościowej, na którą składa się kilka aspektów: identyfikacyjny czy opisowy, genetyczny, celowościowy, fazy czy rozwojowy dotyczący zaawansowania badanego stanu ku prognozowaniu i projektowaniu (Ziemski, 1973). Diagnoza może być prowadzona jednorazowo na jednej osobie lub grupie albo wielokrotnie. Wówczas mamy diagnozę dynamiczną realizowaną w odstępach czasu chcąc badać dynamikę potrzeb, aspiracji, postaw. Mówi się też o postdiagnozie, która wnika w czas przeszły, a diagnosta zastanawia się jak było w danym go obszarze (Mazurkiewicz, 2003).

Przede wszystkim, niezależnie od tego, czy jest to diagnoza społeczna, środowiskowa, resocjalizacyjna, psychologiczna, pedagogiczna czy edukacyjna najważniejszy jej podział dotyczy diagnozy pozytywnej i negatywnej. Diagnoza pozytywna dotyczy pozytywnych cech, które diagnosta chce ujawnić u jednostki, grupy i ich wychowawczych środowisk. Częściej jednak w praktyce pedagogicznej wciąż eksponuje się diagnozę negatywną dotyczącą badania stanu, który niepokoi pedagogów nauczycieli wychowawców, rodziców. Dzieje się tak, gdy uczeń ma trudności w kontaktach rówieśniczych lub w uczeniu się, gdy źle się zachowuje, stosuje przemoc, zamyka się w sobie, spożywa alkohol czy wagaruje. Wówczas szukamy przyczyn i stosujemy diagnozę negatywnych cech, zachowań i stanów, ale ten rodzaj diagnozy nie wystarcza w uruchamianiu pomocy i rozwiązywaniu problemów, aby uczeń stawał się silny i sam potrafił radzić sobie. Dlatego równoległe do diagnozy negatywnej potrzeba rozpoznawania mocnych stron jednostki i środowiska czyli tzw. czynników chroniących, sił ludzkich by zastępować nimi zauważane braki i popełniane błędy. Tak samo jest w diagnozie edukacyjnej czy w rozpoznawaniu niekorzystnej sytuacji szkolnej lub rodzinnej ucznia. Ona też powinna być ukierunkowana na kompensację, wzmacnianie, rozwój, autonomię i odpowiedzialność w myśl słynnej *spirali wsparcia* S. Kawuli (1996): „Pomagam tobie tak, abyś sam sobie umiał pomóc”. Istotne jest by środowisko było korzystne dla rozwoju jednostki. Należy diagnozować jego siły i ukierunkowywać na rozwój. Jak pisze K. Barłóg podstawą wspomagania jest obserwacja dziecka, poznanie jego indywidualnych możliwości, zdolności, zainteresowań, a dopiero wówczas uruchamianie wsparcia, pomocy i dostosowania metod pracy z dzieckiem i z jego rodziną (2015, s. 40).

Trafne i proste rozumienie diagnozy podaje H. Radlińska, J. Korczak czy S. Ziemski, którzy akcentują jej prakseologiczny sens stawiając na jej całościowy wymiar. W nawiązaniu do eksponowanych kilka aspektów czy też elementów S. Ziemski wyjaśnia jej sens w taki oto sposób: „Diagnoza jest to rozpoznanie na podstawie zebranych objawów i znanych ogólnych prawidłowości badanego złożonego stanu rzeczy przez przyporządkowanie go do typu albo gatunku, dalej przez wyjaśnienie genetyczne i celowościowe, określenie jego fazy obecnej oraz przewidywanego rozwoju” (Ziemski, 1973, s.68). Także H. Radlińska jako prekursorka problematyki stawia na diagnozę będącą podstawą decyzji wzmacniających i doskonalących zdiagnozowany stan lub wyrównania, korekty, naprawy to co szkodliwe. Określa diagnozę jako ciągły proces poznawania przyczyn badanego stanu, oceny jego objawów (dodałabym widocznych efektów) i możliwości zmian w kierunku pedagogicznie pożądanym (Radlińska, 1961, s. 370). A Mazurkiewicz analizując teorie wskazanych autorów podsumowuje, że diagnoza to opis wyników badania rzeczywistości wychowawczej, edukacyjnej dokonany na podstawie wiedzy z różnych źródeł, która jest podstawą oceny, a szerzej ewaluacji. Obejmuje opisanie jego symptomów, zakwalifikowanie do danego typu, wskazanie etiologii i przyczyn, znaczenia rozwojowego, etapu rozwoju i możliwości zmian (Mazurkiewicz, 2003, s. 50-55).

W tak krótkim tekście nie da się szeroko opisać poruszanych zagadnień i prób krytyki. Warto jednak wskazać najważniejszy wniosek odnoszący się do sensu diagnozy podkreślając, że powinna być zacznem każdej działalności pedagogicznej - profilaktycznej, pomocowej, naprawczej czy doskonalącej, twórczej, rozwojowej. Jest instrumentem osiągnięcia głębszych

celów edukacji. Stąd musi być **ukierunkowana na rozwój** (Obuchowska, 2002; Wysocka, 2007), a więc diagnozie negatywnej zawsze powinno towarzyszyć rozpoznawanie mocnych stron, sił i potencjału (Por. Wysocka i Jarosz, 2006), którymi można rekompensować zdiagnozowane niedociągnięcia i braki. Tylko wtedy **diagnoza będzie skuteczna**. Diagnoza osiągnie swój cel, gdy będzie wyrażać perspektywę wychowawcy, wychowanka i środowiska, gdy będzie mieć subiektywny i obiektywny aspekt, potraktowanie ilościowo-jakościowe, a przede wszystkim **holistyczne ujęcie** (Nikitorowicz, 2006). Mówimy wówczas o całościowej, **pełnej diagnozie** mającej sens poznawczy, a głównie prakseologiczny.

1.2 Diagnoza a ewaluacja w szkole

Musimy mieć świadomość, że zarówno diagnoza poznawcza, jak i decyzyjna są podstawą, warunkiem oceny pedagogicznej i jej „ciągiem dalszym” wskazującym na kierunek działań doskonalących jakość pracy szkoły i pracy pedagogicznej w każdej placówce edukacyjnej. Diagnoza poznawcza służy ustalaniu standardów, prawidłowości, norm, wzorców, do których nauczyciel może porównać zdiagnozowany stan czyli dane na temat osiągnięcia danego standardu, poziomu opanowania wiedzy w danym zakresie i ukształtowania umiejętności. Świadczą one bowiem o wynikach pracy szkoły czy innej placówki, a więc o rezultatach widocznych w kompetencjach uczniów, które podlegają ocenie i są ważnym wskaźnikiem pomiaru pedagogicznego czy elementem ewaluacji wewnętrznej i zewnętrznej w szkole.

W świetle przepisów prawa oświatowego przez **ewaluację** „należy rozumieć proces gromadzenia, analizowania i komunikowania informacji na temat wartości działań podejmowanych przez szkołę lub placówkę; wyniki ewaluacji są wykorzystywane w procesie podejmowania decyzji skierowanych na zapewnienie wysokiej jakości organizacji procesów kształcenia, wychowania i opieki oraz ich efektów w szkole lub placówce” (Rozporządzenie Ministra Edukacji Narodowej z dnia 27 sierpnia 2015...). Zgodnie z nim istnieje obowiązek **ewaluacji zewnętrznej** prowadzonej przez organy nadzoru pedagogicznego i **ewaluacji wewnętrznej** organizowanej przez dyrektora. W ramach ewaluacji zewnętrznej jest **ewaluacja całościowa** realizowana odnośnie wszystkich wymagań stawianych szkołom i **ewaluacja problemowa** odnosząca się do wybranych standardów.

W szkole **ewaluacja zewnętrzna** jest realizowana w zakresie wymagań określonych w *rozporządzeniu Ministra Edukacji Narodowej z dnia 6 sierpnia 2015 r. w sprawie wymagań wobec szkół i placówek* (Dz. U. poz. 1214) i obejmuje: zbieranie i analizowanie informacji o działaniach szkoły w zakresie wymagań państwowych; ustalenie, czy szkoła spełnia te wymagania oraz opracowanie raportu z ewaluacji. Ewaluacja zewnętrzna ma być demokratycznym narzędziem wsparcia szkół w rozwoju. Ma inspirować do wprowadzania zmian i nowatorskich rozwiązań. Ma służyć rodzicom i uczniom w zdobywaniu informacji o szkole, ale i motywować do aktywnego uczestnictwa w kreowaniu warunków pracy szkoły, umożliwiać realny wpływ podmiotom szkoły na jej funkcjonowanie. Przy ewaluacji istotne są więc opinie dyrektora i nauczycieli, uczniów, rodziców, podmiotów lokalnych współpracujących ze szkołą (<http://www.npseo.pl/action/externalevaluation>; 26.07.2016).

Dyrektor szkoły we współpracy z innymi nauczycielami zajmującymi stanowiska kierownicze, w ramach nadzoru przeprowadza natomiast **ewaluację wewnętrzną** zwaną **autoewaluacją** i wykorzystuje jej wyniki do doskonalenia jakości pracy szkoły. Kontroluje przestrzeganie przez nauczycieli przepisów prawa i wspomaga nauczycieli w realizacji ich zadań, w szczególności przez: diagnozę pracy szkoły, planowanie działań rozwojowych, w tym motywowanie nauczycieli do doskonalenia zawodowego, prowadzenie działań rozwojowych. Ewaluację wewnętrzną przeprowadza się w odniesieniu do zagadnień uznanych w szkole za istotne z punktu widzenia jej specyfiki. W tym celu dyrektor szkoły

obserwuje prowadzone zajęcia dydaktyczne, wychowawcze i opiekuńcze, inne sytuacje i czynności wynikające z działalności statutowej szkoły. Opracowuje na każdy rok szkolny plan nadzoru, który przedstawia na zebraniu rady pedagogicznej. Zawiera on przedmiot i termin ewaluacji, zakres i czas kontroli przestrzegania przez nauczycieli przepisów prawa i zakres wspomaganie w realizacji zadań. Do końca roku dyrektor przedstawia wyniki i wnioski z dokonanej ewaluacji. (<http://www.npseo.pl/action/internalevaluation>; 26.07.2016).

Szkoła może spełniać wymagania na pięciu poziomach: niskim, podstawowym, średnim, wysokim i bardzo wysokim, gdzie A oznacza najwyższy poziom, a E - najniższy. Raporty zawierają wyniki badań i rozmów, jakie wizytatorzy prowadzą z nauczycielami, uczniami, rodzicami czy z przedstawicielami instytucji współpracujących ze szkołą, ale też zapisy z obserwacji podczas lekcji i przerw. **Zatem z ewaluacją wiąże się diagnoza, która jest jej nieodłącznym elementem.** Każdy dowie się z raportu, jakie szkoła proponuje indywidualne programy, zajęcia pozalekcyjne, humanistyczne, artystyczne czy sportowe. Sprawdzą, jakim sprzętem dysponuje, o której zaczynają się lekcje, czy jest monitoring, jaka część nauczycieli prowadzi ciekawe lekcje (Suckiel, <http://www.zs9.bydgoszcz.pl/ewaluacja-zewn%C4%99trzna-ocena-szko%C5%82y>; 26.06.2016).

Próbując ustosunkować się do niniejszego punktu zatytułowanego „Diagnoza a ewaluacja w szkole” chciałabym podkreślić ich wzajemne relacje. Bowiem **diagnoza edukacyjna służy ocenie pracy szkoły**, jej jakości, skuteczności, trafności podejmowanych działań czyli **ewaluacji** właśnie. **Diagnoza służy zbieraniu informacji z różnych źródeł** od uczniów, nauczycieli, rodziców, innych pracowników, specjalistów w szkole i poza nią, itp. **O warunkach, przebiegu i rezultatach edukacji** szkolnej jako procesu zamierzonego, zaplanowanego, mającego wymiar instytucjonalny i społeczno-kulturowy. Informacje te są potrzebne do oceny wartości działania edukacyjnego, zasadności podejmowanych zmian i decyzji edukacyjnych. W tym miejscu warto odwołać się więc do stanowiska autorytetu diagnostyki edukacyjnej na gruncie polskim Bolesława Niemierko. W swojej pracy o związkach diagnozy z ewaluacją pisze, że ewaluacja jest szersza od diagnozy edukacyjnej, bo jest nastawiona na więcej źródeł informacji, na szerszy kontekst i różne obszary pracy szkoły, z którymi wiąże się podejmowanie decyzji w zakresie metodycznym, organizacyjnym, kadrowym i materialno-finansowym. Ewaluacja jest związana z badaniami stosowanymi (2009, s. 317-319), które wiążą się z wdrażaniem zmian, działań doskonalących pracę szkoły skierowanych na wyższą jakość edukacji. Posiłkuje się też diagnozą potrzeb i możliwości rozwojowych uczniów, jak też nauczycieli, stanem, współpracy szkoły z rodziną, klimatem i atmosferą panująca w szkole, poziomem zaangażowania uczniów, nauczycieli i rodziców w życie szkoły, sytuacją materialną szkoły, niepowodzeniami, samorządnością itp. Zatem nie można ich od siebie odłączyć. Badania diagnostyczne są warunkiem lepszego dostosowania działalności szkoły i pracy nauczycieli do potrzeb uczniów i specyfiki środowiska lokalnego.

Także znawca problematyki H. Mizerek podkreśla różnicę między ewaluacją a diagnozą wyjaśniając, że ewaluacja jest specyficznym rodzajem badań społecznych czy, dodałabym badań nad edukacją i w edukacji. Ewaluację od diagnostyki odróżnia sposób wykorzystania zgromadzonych informacji. Ewaluacja koncentruje się na gromadzeniu wiedzy, która jest podstawą formułowania sądów na temat wartości podejmowanych działań. Istnieje jednak ogromna różnica między formułowaniem sądów wartościujących a potocznie rozumianym ocenianiem. Często tendencja utożsamiania ewaluacji z ocenianiem ogranicza możliwość wykorzystania wyników ewaluacji w praktyce (Mizerek, 2010-2011, s. 21).

2 MOCNE I SŁABE STRONY DIAGNOZY I EWALUACJI W POLSKIEJ SZKOLE

Poszukując czynników chroniących i zagrażających wyższej jakości i skuteczności pracy polskiej szkoły w aspekcie omawianych zagadnień, warto uświadomić sobie determinanty skutecznej diagnozy i ewaluacji w edukacji. Składają się nań uwarunkowania podmiotowe i przedmiotowe, czy też prawne, polityczne, społeczne i kulturowe, w tym w szczególności pedagogiczne związane ze stylami myślenia i kompetencjami nauczycieli praktyków jako diagnostów i autoewaluatorów. Skuteczna diagnoza i ewaluacja to taka, która osiąga swe cele związane z wysoką jakością działań edukacyjnych dostosowanych do potrzeb i możliwości każdego ucznia, ale też możliwości nauczyciela, szkoły i środowiska. Dlatego ewaluacja jest związana ze świadomością tzw. *ewaluatorów zewnętrznych* pracujących w organach nadzoru pedagogicznego. Wszystkim im bowiem musi towarzyszyć jeden cel jakim jest dobro ucznia i jego wszechstronny rozwój. Musi on być odległy od potocznie rozumianej sztucznej oceny czy diagnozy jako orzekania o zastanym, kreowanym lub uzyskanym stanie dla samej znajomości i wiedzy o nim. Mając na uwadze sposoby definiowania, rozumienia diagnozy i ewaluacji w szkołach przez znawców zagadnienia (do takich należy inicjator corocznych konferencji i wielotomowych publikacji na temat diagnostyki i ewaluacji pedagogicznej Josef Malach (2013, 2015) kierownik Katedry Pedagogiki i Andragogiki i jego współpracownicy z Wydziału Pedagogicznego Uniwersytetu Ostrawskiego) oraz uwzględniając procedurę diagnostyczną i ewaluacyjną w polskich szkołach zauważa się przykłady dobrych praktyk, które należy promować, ale też niedociągnięcia, nad którymi warto pracować. Przykłady te ujęłam w tabeli.

Tab. 1 Mocne i słabe strony diagnozie i ewaluacji ku wyższej jakości w polskiej szkole

Polska szkoła w zakresie DIAGNOZY i EWALUACJI	MOCNE STRONY (czynniki chroniące)	SŁABE STRONY (czynniki zagrażające)
	<ul style="list-style-type: none"> - możliwość współpracy i wysokie kompetencje specjalistów w zakresie diagnozy interdyscyplinarnej ucznia; - wielość naukowych opracowań i bogatych propozycji diagnostycznych i ewaluacyjnych dla szkół z zakresu pedagogiki i nauk współpracujących: psychologii, socjologii i pomocniczych jak: prawo czy medycyna; - zasoby szkół wyższych: kadrowe, metodyczne, biblioteczne w zakresie przygotowywania przyszłej kadry do skutecznej diagnozy oraz ewaluacji wewnętrznej i zewnętrznej w szkołach; - zagwarantowana prawnie autonomia szkół i nauczycieli (liczne akty prawne i rozporządzenia,) w rozpoznawaniu potrzeb, możliwości uczniów i ich środowisk, co pozwala na dobór strategii do prowadzonej diagnozy, autonomia szkoły w zakresie 	<ul style="list-style-type: none"> - mała liczba specjalistów w szkołach lub współpracujących ze szkołami w zakresie diagnozy wielospecjalistycznej i ewaluacji w odniesieniu do ucznia o specjalnych potrzebach; - niski wciąż poziom kompetencji głównie metodologicznych nauczycieli w zakresie badań diagnostycznych i ewaluacyjnych, mała znajomość technik i narzędzi, procedur; - niski poziom gotowości nauczycieli do diagnozowania i autoewaluacji dla podejmowania trafnych decyzji i działań z uwagi na brak czasu i przeciążenie pracą wymaganą od nauczycieli w szkołach, ciągłe zmiany przepisów, nie zawsze są zasadne, itp. - ograniczone środki finansowe na prowadzenie badań diagnostycznych i ewaluację oraz na wszechstronne opracowanie ich wyników; duże przeciążenie nauczycieli dodatkową pracą z racji niskiego uposażenia i małej motywacji; - mała współpraca nauczycieli w zakresie diagnozy i ewaluacji edukacyjnej z obawy

<p>ewaluacji wewnętrznej nastawiona na współpracę;</p> <ul style="list-style-type: none"> - bogate zasoby i różne możliwości uczelni kształcących w obszarze nauk humanistycznych i społecznych, w tym głównie z pedagogiki w zakresie prowadzenia badań diagnostycznych i ewaluacyjnych na potrzeby praktyki szkolnej z wykorzystaniem potencjału badawczego studentów i kadry akademickiej; - duże możliwości szkół i nauczycieli w pozyskiwaniu środków finansowych przeznaczonych na mało zbadane obszary edukacji – problemowe czy też związane z potencjałem uczniów, zasobami szkoły i środowiska, w zakresie współpracy badawczej szkół, regionów czy różnych krajów nawet; - liczne przykłady dobrych praktyk w zakresie skutecznej diagnozy i ewaluacji w szkołach nakierowanych potencjał na ucznia i mocne strony, ale i ograniczenia środowiska szkolnego oraz rodzinnego; - platformy internetowe, organizacje działające na rzecz promowania jakości edukacji i wykorzystywania w tym celu diagnozy i ewaluacji ku rozwojowi; Jako przykłady dobrych praktyk można wskazać działalność: Polskiego Towarzystwa Diagnostyki Edukacyjnej, które od lat pod przewodnictwem B. Niemierko cyklicznie organizuje różne warsztaty, konferencje, a 23-25 września 2016 r. jest XXII krajowa konferencja diagnostyki edukacyjnej w Krakowie nt. <i>Diagnozowania twórczości uczniów i nauczycieli</i>; Inne wartościowe przykłady to Stowarzyszenie Szkół Aktywnych (ssa.edu.pl); Towarzystwo Szkół Twórczych; Stowarzyszenie Nauczycieli Olimpijskich; Instytut Badań Edukacyjnych (ibe.edu.pl) prowadzi badania naukowe nad funkcjonowaniem i efektywnością systemu edukacji w Polsce; Warto też wskazać na Ośrodek Rozwoju Edukacji i System Ewaluacji Oświaty 	<p>przed okazaniem niewystarczających kompetencji i małe zaufanie do organizatorów oświaty;</p> <ul style="list-style-type: none"> - mała motywacja nauczycieli do podejmowania diagnozy i ewaluacji, traktowanie przez wielu pedagogów jako pracy nadobowiązkowej, której nie muszą wykonywać z racji małej świadomości korzyści płynących ze stosowania jej w pracy; - nie zawsze wiarygodny materiał pochodzący z jednego zamiast z kilku źródeł i stereotypy myślowe uczących, błędna interpretacja wyniku; - ograniczony dostęp do fachowej literatury dla małych środowisk oraz brak wsparcia w tym zakresie dyrektorów, którzy stawiają na wysokie noty z egzaminów, konkursów, olimpiad w pogoni za renomą, dużą liczbą uczniów w szkole; - mało dynamicznego, systemowego i interdyscyplinarnego ujęcia diagnozy niewystarczające zasoby kadrowe do prowadzenia diagnozy ucznia i jego specjalnych potrzeb – z niepełnosprawnością, z deficytami środowiskowymi, a także wybitnie zdolnego; - mało jest w diagnozie i ewaluacji w szkołach wykorzystywania metod i analiz jakościowych - interpretowania, rozumienia, wartościowania, a często dominuje pomiar dydaktyczny oparty na suchych danych liczbowych i procentowych, na średnich, standardach, tendencjach; - największym mankamentem wydaje się małe ukierunkowanie w diagnozie, która często ma charakter wymuszony, interwencyjny, bo jest ukierunkowana na negatywne stany, a mało jest dopełniania diagnozą potencjału czyli mocnych stron jednostki i środowiska, choć tylko wówczas diagnoza osiągnie swój najgłębszy sens i wymiar. - mała współpraca a duże współzawodnictwo podmiotów szkoły, małe zaangażowanie i zaufanie do siebie, wiele udawania w prowadzeniu ewaluacji, która ma służyć podnoszeniu jakości, a nie ocenie nauczyciela; - nie zawsze wyniki diagnozy i ewaluacji są obiektywne, tzn. bezstronne, bo wciąż wśród podmiotów szkoły, a głównie u nauczycieli, dyrektorów szkół dominuje adaptacyjne myślenie; Chcą oni wykazać, że ich szkoły są
--	--

	<p>(ore.edu.pl; npseo.pl), itp.</p> <p>- jest wielu wartościowych nauczycieli i kreatywnych pedagogów w polskich szkołach, którzy kochają swoją pracę z dziećmi i młodzieżą. Mają więc mocną wewnętrzną motywację do zbierania informacji na temat ich potrzeb, możliwości i sposobów pomocy by stawali się samodzielni i kompetentni;</p>	<p>dobrze nie zawsze myśląc o tym, jak zrobić, aby były bardzo dobre, a wręcz znakomite!;</p> <p>- Podobnie jest ze szczerością i bezstronnością ze strony rodzica, który nie zawsze chce przyznać, że jest oporny na współpracę, zaś często nauczycielowi wytyka błędy i braki, a przecież w diagnozie i ewaluacji chodzi o to by uchwycić prawdziwy stan i szukać rozwiązań;</p>
--	--	--

Źródło: opracowanie własne

Specjalnego potraktowania wymaga problem przygotowania i umotywowania szkoły i nauczycieli do diagnozy i ewaluacji skierowanej na doskonalenie swojej pracy i rozwijanie współpracy z różnymi podmiotami ku bardziej wartościowej i satysfakcjonującej edukacji. Zapewne wartościowe są niektóre prawne rozwiązania, ale są one bardziej oparte na kontroli i zewnętrznej ocenie, niż na szczerości, zaufaniu i odpowiedzialności wszystkich podmiotów tworzących system szkolny: uczniów, nauczycieli, rodziców, reprezentantów nadzoru pedagogicznego i kuratorium oświaty, pracowników naukowych szkół wyższych i studentów. Szczególnego akcentu wymaga też przygotowanie metodologiczne i cechy osobowościowe osób, które prowadzą lub będą prowadziły diagnozę dla wyjaśniania tego co niewidoczne gołym okiem (za K. Konarzewski, 1999), rozpowszechniania wiedzy o nim i projektowania swojej pracy związanej z ujawnionym stanem. Od lat przedmiotem moich zainteresowań są nauczyciel i szkoła, od lat też prowadzę zajęcia wśród czynnych i przyszłych nauczycieli pedagogów z zakresu metodologii badań pedagogicznych czy szerzej badań w pedagogice. Od lat więc zauważam trudność w opanowywaniu wiedzy i rozwijaniu umiejętności czy szerzej kompetencji z zakresu metodologii badań ilościowych i jakościowych, teoretycznych i praktycznych. Przede wszystkim dostrzegam trudność w głębokim rozumieniu potrzeby triangulacji badań oraz współpracy badaczy teoretyków i praktyków. O tych możliwościach, ograniczeniach i trudnościach potraktuję w artykule na temat kompetencji metodologicznych kandydatów na nauczycieli do diagnozy i ewaluacji ku wyższej jakości szkolnej edukacji.

ZAKOŃCZENIE

Podjęte analizy prowadzą do wniosku, że zapewnianie, ocena i doskonalenie jakości pracy szkoły spoczywa na dyrektorze i z jego perspektywy są nierzadko oceniane i opisywane efekty i jakość działań edukacyjnych, ale też diagnostycznych i ewaluacyjnych. Ważnym zadaniem, a zarazem wyzwaniem staje się uspołecznianie procesów ewaluacji i zarządzania jakością (włączanie w ten proces uczniów, nauczycieli, rodziców). Budowana dzięki takim działaniom wspólnota szkolna lepiej może rozpoznać potrzeby szkoły i jej uczestników, w tym głównie uczniów i ich rodziców. Taka wspólnota trafniej będzie dokonywać ewaluacji poszukując mocnych i słabych stron szkoły jako systemu. Zarządzanie jakością służące rozwojowi szkoły zakreśla szeroki horyzont działań i obszarów współpracy oraz korygowania błędów, uzupełniania braków i doskonalenia tego, co okazało się dobre i pozytywne. Konieczne jest więc rozwijanie potencjału kadry pedagogicznej oraz lepsze merytoryczne, metodologiczne i osobowościowe przygotowanie kandydatów na nauczycieli do diagnozy rozwojowej i ewaluacji ku wyższej jakości w szkołach. Istotne jest położenie nacisku na zaangażowanie nauczycieli i delegowanie uprawnień, aby w szkole mogli rozwijać rolę i kompetencje liderów edukacyjnych- wsparcia w rozwoju i w trudnych sytuacjach.

Rozwijanie krytycznej autorefleksji i wynikającej z tego odpowiedzialności za jakość działań pozwoli uniknąć przerzucania swojej odpowiedzialności na inne osoby i czynniki zewnętrzne (http://www.npseo.pl/action/requirements/wymaganie12_zarzadzanie_szkola_lub_placowka_sluzuj_jej_rozwojowi; 26.07.2016).

Jak wskazuje znawca problematyki H. Mizerek: „Konieczne jest budowanie systemu ewaluacji efektywnej. W takim systemie ewaluacje zewnętrzne różnią się co do szczegółowych celów, łączą je natomiast wspólne kryteria i wskaźniki oraz podobne narzędzia” (Mizerek, 2010-2011, s. 20). Jest to ważny wniosek, jaki nasuwa mi się w wyniku podjętych rozważań postulując w szkole jako organizacji uczącej się w szczególności zmianę świadomościową i mentalną. Istotną myśl końcową chcę przekazać za B. Ciężką: „Ewaluacja jest grupą wszystkich narzędzi do pozyskiwania informacji mających wspomagać proces decyzyjny i budować współodpowiedzialność za osiągnięcie rezultatów pracy, nie oferuje natomiast gotowych rozwiązań czy sposobów postępowania” (2010-2011, s. 83). I o tym należy pamiętać w przygotowywaniu nauczycieli, w budowaniu lepszej szkoły i optymalnych warunków edukacji oraz rozwoju dzieci, młodzieży, dorosłych w każdym miejscu i czasie. Podsumowując rozważania dotyczące diagnozy jako fundamentu warto powołać się na szeroki kontekst rozwojowy i kompleksowy charakter diagnozy, jaki przedstawia K. Barłóg. Autorka akcentuje zmianę paradygmatu diagnozy z modelu tradycyjnego, który skupiał się na rozpoznawaniu deficytów w kierunku poszukiwania szans i przestrzeni rozwoju (2008, s. 85). Akcentuje potrzebę „formułowania diagnoz w sposób pozwalający wybrać właściwe działania pomocowe i ustalić strategię ich realizacji” (2011, s. 25). Tak też jest z ewaluacją.

W końcowej części rozważań pragnę podkreślić raz jeszcze konieczność badania, oceniania, kontrolowania, ewaluacji wewnętrznej i zewnętrznej zawsze dla doskonalenia i rozwoju. Zatem nie możemy w badaniu jakości pracy szkoły i działań nauczyciela skupiać się jedynie czy w głównej mierze na niedociągnięciach, błędach, brakach, potknięciach. Jeśli więc ten negatywny aspekt występuje jesteśmy zobligowani do poszukiwania czynników chroniących i mocnych stron szkoły jako systemu w wymiarze instytucjonalnym i społeczno-kulturowym. Diagnoza i ewaluacja dla rozwoju obliguje do wzajemnej szczerości, zaufania, zaangażowania i odpowiedzialności podmiotów szkoły i jej partnerów. Aby jednak tak się stało muszą oni zmieniać swe myślenie z adaptacyjno-instrumentalnego na krytyczno-kreatywne. Lider edukacji, pedagog krytycznie myślący i z poczuciem pasji będzie tych stereotypowych myśli i instrumentalnych działań zaprzeczeniem. Zatem kolejny problem, jaki jawi się z podjętych rozważań to poszukiwanie odpowiedzi na pytanie: Jaki nauczyciel, dyrektor, pedagog, lider edukacji będzie promotorem diagnozy i edukacji dla rozwoju? Jakie są mu potrzebne kompetencje, jak je diagnozować, doskonalić i wspierać by realizował po partnersku i z odpowiedzialnością mądrą i wartościową edukację dla wartościowego życia?

LITERATURA

Barłóg, K. (2008). Wspomaganie rozwoju dzieci z niepełnosprawnością intelektualną w stopniu lekkim w różnych formach edukacji wczesnoszkolnej, Wyd UR Rzeszów.

Barłóg, K., Kensy, E., Mach A., Rorat, M. i Zaborniak-Sobczak, M. (2011). Wczesne wspomaganie rozwoju i edukacja dzieci zagrożonych niepełnosprawności i niepełnosprawnych w wieku przedszkolnym, Wyd. UR: Rzeszów.

Barłóg, K. (2015). Wsparcie wczesnorozwojowe dzieci zagrożonych niepełnosprawnością i niepełnosprawnych. Wyd. UR: Rzeszów, s. 33-47.

Ciężka, B. (2010-2011). Planowanie ewaluacji wewnętrznej w szkole (placówce) wraz z przykładami projektów ewaluacji. W: Mazurkiewicz, G. (red.) *Ewaluacja w nadzorze pedagogicznym. Autonomia*. Wyd. UJ: Kraków, s. 83-92.

Jarosz, E. (2001). *Wybrane obszary diagnozowania pedagogicznego*. UŚ: Katowice.

Jarosz, E. i Wysocka, E. (2006). *Diagnoza psychopedagogiczna*. Wyd. „Żak”: Warszawa.

Kawula, S. (1996). Spirala życzliwości: od wsparcia do samodzielności. W: „Wychowanie na co Dzień”, nr 10/11), s. 14-17.

Konarzewski, K. (1999). Komu jest potrzebna diagnostyka oświatowa? W: Niemierko, B. I Machowska, B. (red.) *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. ODESiP: Gdańsk.

Malach, J. (2015). Rozvoj a hodnoceni smyslu pro iniciativu a podnikavost žáků základních škol. Pedagogická fakulta Ostravská univerzita v Ostravě: Ostrava.

Malach, J. (2015). Přínos programovaného učení k rozvoji pedagogické teorie a praxe. Pedagogická fakulta Ostravská univerzita v Ostravě: Ostrava.

Mazurkiewicz, E. A. (2003). Diagnostyka w pedagogice społecznej. W: Pilch, T. i Lepalczyk, I. *Pedagogika społeczna*. Wydawnictwo „Żak”: Warszawa, s. 51-70.

Mizerek, H. (2010-2011). Efektywna autoewaluacja w szkole. Jak ją sensownie zaprojektować i przeprowadzić. W: Mazurkiewicz, G. (red.) *Ewaluacja w nadzorze pedagogicznym. Autonomia*. Wyd. UJ: Kraków, s. 20-63.

Niemierko, B. (2009). *Diagnostyka edukacyjna*. PWN: Warszawa.

Nikitorowicz, J. (2006). Próba holistycznego ujęcia procesu diagnostycznego w naukach społecznych. W: Bolejko, A. i Zińczuk, M. (red.) *Terapia pedagogiczna w teorii i praktyce*. Wyd. „Logopeda radzi”: Białystok.

Obuchowska, I. (2002). Osoby niepełnosprawne: diagnoza dla rozwoju. W: Lotz D., Wenta K., Zeidler W. (red.) *Diagnoza dla osób niepełnosprawnych*. Wyd. UŚ: Szczecin, s. 40-45.

Radlińska, H. (1961). *Pedagogika społeczna*. Wyd. Ossolineum: Wrocław.

Skalbania, B. (2011). *Diagnostyka pedagogiczna. Wybrane obszary badawcze i rozwiązania praktyczne*. „Impuls”: Kraków.

Wysocka, E. (2007). Człowiek a środowisko życia – podstawy teoretyczno-metodologiczne diagnozy. Wyd. „Żak”: Warszawa.

Wysocka, E. (2013). *Diagnostyka pedagogiczna. Nowe obszary i rozwiązania*. „Impuls”: Kraków.

Ziemski, S. (1973). *Problemy dobrej diagnozy*. „Wiedza Powszechna”: Warszawa.

KONTAKT

dr Maria Kocór

Wydział Pedagogiczny, Uniwersytet Rzeszowski

ul. Ks. J. Jałowego 24, 35-010 Rzeszów

Telefon: +48 696 031 355

e-mail: mariakoc@wp.pl

OCENIANIE W KONSTRUKTYWISTYCZNYM MODELU KSZTAŁCENIA

EVALUATING IN THE CONSTRUCTIVIST MODEL OF EDUCATION

Tomasz M. Zimny

Abstrakt: Konstruktivismus je jedním z důležitých teorií vzdělávání v Polsku. Vytváří nástroje, které by mohly být použity k nahrazení "staré" vyučovací metody založené na přenosovém modelu vzdělávání. Srovnání chování a konstruktivistické modely vzdělávání v posledním desetiletí ukazují, že první podporuje různé formy podřízenosti vedení ideologie adaptace. Tato situace stimuluje jednotnost myšlení studentů, které ve školním kontextu znamená, že studenti dosahují ne své vlastní cíle, ale jejich učitele. To vede k lineárnímu myšlení, jehož cílem je přesně definovaný výstup, spíše než divergentní myšlení.

V oblasti vzdělávání je velmi důležité stanovit cíle vzdělávání a metody ověřování jejich dosažení. Závěrečné zkoušky ve škole kontrolují pracovní výsledky jak studentů, tak učitelů. Ty by se měly zaměřit především na studenta: přizpůsobit jeho / její vývoj a schopnost učit se (včetně možnosti vyhledávání, hodnocení, výběru, zpracování a používání informací za úkol úspěch), tak i praktickým dovednostem. Zadáání cíle vzdělávání v těchto kategoriích poruší tradici školy a vyžaduje významné změny v hodnotících metodách. Ve skutečnosti, navrhovaný přístup zavádí správný vztah mezi různými účastníky vzdělávání: učitelem a žákem.

Klíčová slova: hodnocení školy, cíle vzdělávání, zkoušky, vyšetření, konstruktivismus

Abstract: Constructivism is one of the important theories of education in Poland. It creates tools that could be used to replace the "old" teaching methods, based on the transmission model of education. Comparison of behavioral and constructivist education models in the last decade show that the first promotes various forms of subordination guided by the ideology of adaptation. This situation stimulates an uniformity of thinking in students, which in the school context means that students are achieving not their own goals, but theirs teachers. It leads to the linear thinking, aimed at a strictly defined purpose, rather than divergent thinking.

In the education, it is crucial to define learning objectives and methods of verifying their achievement. Final school exams checks the work results of both student and teacher. They should primarily focus on the student: to conform his/her development and learning skills (including the ability to searching, evaluating, selecting, processing, and use of information for task accomplishment) and also practical skills. Specifying the objectives of education in these categories breaks the tradition of the school, and requires significant changes in evaluation methods. In effect, proposed approach establishes the proper relationship between different participants of education: the teacher and the student.

Key words: evaluation of school, aims of education, exams, constructivism

WPROWADZENIE

W XXI wieku edukacja transmisyjna oparta na behawioryzmie nie spełnia oczekiwań społecznych i jednostkowych uczniów, rodziców, nauczycieli. Spowodowane jest to zmianami społecznymi wynikającymi z postępu technologicznego, organizacyjnego. W czasach XIX i XX wieku zarysowany został wyraźnie podział na trzy fazy życia człowieka faza zabawy do okresu obowiązku szkolnego, faza uczenia się do podjęcia pracy zawodowej, oraz faza pracy zawodowej. Obecnie ten podział jest już nieaktualny ze względu przynajmniej na dwa postępujące zjawiska, a mianowicie tempo zmian jest na tyle duże, że po ukończeniu szkoły by pracować konieczna jest kontynuacja uczenia się, lecz dzięki postępowi pracownik dysponuje znacznie większym budżetem czasu wolnego, który na uczenie się i na zabawę. Te zmiany cywilizacyjne powodują także, to że szkoła ma nie tylko przygotować ucznia czy studenta do podjęcia pracy zawodowej, lecz także a może przede wszystkim przygotować do kierowania własnym rozwojem, do kierowania procesem własnego uczenia się. Te zmiany wymagają wniknięcia w istotę procesu uczenia się i zmiany celów i kontrolowania i oceniania realizacji tych celów edukacji w trakcie i po każdym etapie. Być może zmiany wymagają także metody, środki i warunki realizacji celów edukacji. Skoro tempo zmian jest na tyle duże, że szkoła nie może przygotować ucznia do pracy w całym jego dorosłym życiu, to powinna przygotować go do kierowania własnym rozwojem i do uczenia się, co należy przyjąć za cel edukacji szkolnej.

Podstawą behawioryzmu jest model bodziec – reakcja. Jeśli do tego określimy, że nauczyciel ma dobierać bodźce generujące pożądane reakcje ucznia. Takie podejście prowadzi do myślenia o uczniu jako o jednostce biernej mającej się podporządkować systemowi szkolnemu. Wystarczy jednak przyjąć procesualny model Tomaszewskiego (1966) zadanie – działanie – wynik, by całkowicie zmienić ogląd sytuacji ucznia. Po pierwsze zadanie zawsze jest czymś zadaniem, w tym przypadku jest zadaniem ucznia, a jego działanie jest ciągiem bodźców i reakcji ucznia i jego środowiska w którym działa. Ważne jest by nie traktować zadania jako nadanego zewnętrznemu w stosunku do ucznia. Chodzi tu o zadanie, które uczeń sam określi albo przyjmie jako własne do realizacji przez siebie. W procesie uczenia się wynikiem realizacji zadania są zmiany w uczniu, w jego wiedzy, umiejętnościach i ustosunkowaniach do siebie do swojego środowiska i do relacji swojej do środowiska. Jednak taka zmiana oglądu sytuacji szkolnej ucznia jest niewystarczająca z uwagi na brak zmiany celów edukacji szkolnej.

Proces edukacji jest co najmniej dwupodmiotowy i może być zmieniany w zakresie celów, metod, środków, warunków oraz sposobów oceniania wyników. Dwa elementy są kluczowe, cele oraz ocenianie wyników. Ocenianie wyników jest ważne z uwagi na to, że decyduje o faktycznie realizowanych celach. Jeśli ocenianie wyników nie jest zgodne z celami to realizowane cele są różne od deklarowanych. Realizuje się takie cele których realizacja jest kontrolowana i oceniana.

Zmiany cywilizacyjne, postęp technologiczny oraz organizacyjny wymuszają zmiany w określaniu celu kształcenia na przygotowanie ucznia do kierowania własnym rozwojem własnym procesem uczenia się. Skoro przygotowany program edukacji szkolnej określający cele wymagania programowe i treści i behawioralny sposób jego realizowania jest już niewystarczający, ponieważ nie przygotowuje ucznia do kierowania własnym rozwojem do planowania celów własnego uczenia się i poszukiwania w swoim środowisku możliwości ich realizacji z powodu braku podmiotów planujących edukację i bodźcujących do realizowania planu. Skoro uczeń ma być przygotowany do kreowania własnej osoby przez uczenie się i zdobywanie doświadczeń, to należy całkowicie odwrócić strategię edukacji z nastawionej na realizowanie jednego wspólnego programu kształcenia i jednego sposobu i organizacji procesu uczenia się uczniów na zindywidualizowane cele szczegółowe, metody,

środki, dostosowane do możliwości i specyfiki każdego ucznia. Na szczęście postęp technologiczny i organizacyjny nie tylko stawia przed edukacją szkolną nowe wymagania, ale daje jej nowe narzędzia i nowe możliwości organizowania edukacji. Perspektywa konstruktywistyczna edukacji, która ostatnio zyskuje wielu zwolenników jest przydatna do wyjaśniania procesu uczenia się i poszukiwania nowych form działania edukacyjnego.

1 KONCEPCJE KONSTRUKTYWISTYCZNE

Jest wiele koncepcje konstruktywistycznych lecz wszystkie mają kilka wspólnych cech. Najważniejsza dla edukacji to ta, że wiedza jest zawsze czyjaś wiedzą (tak samo jak umiejętności czy ustosunkowania są zawcze czyjeś) i nawet jeśli przyjmiemy, że rzeczywistość obiektywnie istnieje, to wiedza o niej jest kreowana przez każdą jednostkę na podstawie własnych doświadczeń. Konstruowana wiedza o rzeczywistości zależy nie tylko od rzeczywistości lecz także od własności podmiotu poznającego, w konsekwencji należy przyjąć, że jej poznanie jest zawsze subiektywne. Prowadzi to do oczywistego jak się wydaje wniosku, że rzeczywistość nie jest poznawalna obiektywnie, niezależnie od podmiotu poznającego, a to co poznajemy jest konstruktem poznającego. Jest to zgodne z przekonaniem, że ten sam element rzeczywistości (obiekt czy zdarzenie) jest różnie spostrzegany przez różne osoby a nawet przez tą samą osobę w różnym czasie w zależności od wcześniejszych doświadczeń, ustosunkowań, potrzeb, emocji. Mózg człowieka przetwarza informację, organizuje, interpretuje ją. Koncepcje konstruktywistyczne dają możliwość rozwijania autonomii ucznia jego samodzielności w przeciwieństwie do wcześniejszego podejścia zmierzającego do unifikacji i narzucenia wszystkim uczniom standardów edukacyjnych.

Proces uczenia się każdego człowieka zależy od dwóch elementów: od samego człowieka jako podmiotu świadomego i działającego oraz jego środowiska społecznego w tym edukacyjnego jako, że człowiek jest istotą społeczną. Dlatego biorąc pod uwagę te dwie perspektywy psychologiczną i społeczną przywołam tu teorię rozwoju człowieka Piageta i Wygotskiego.

Zdaniem J. Piageta (1966) wiedza jest aktywnie konstruowana przez uczącego się. Istotą rozwoju intelektualnego człowieka są jego interakcje ze swoim środowiskiem. Jest to proces w którym człowiek jest aktywny, dąży do zrozumienia swojego środowiska przez analizowanie i manipulowanie jego elementami. Stara się poznać jakie daje możliwości, jakie stwarza zagrożenia. Piaget marginalizuje znaczenie interakcji społecznych jako dostarczających wiadomości a posługiwanie się językiem uważa za wynik rozwoju nie zaś za narzędzie służące rozwojowi. Rozwój psychiczny człowieka uznaje za przejaw dojrzewania struktur poznawczych (Piaget, 1967). Według Piageta struktury poznawcze człowieka ulegają ciągłym przekształceniom w wyniku konfrontacji swojej wiedzy i oferowanej przez środowisko w tym w szczególności przez szkołę. Odbywa się to przez asymilację czyli włączanie nowych wiadomości do już posiadanych oraz akomodacji czyli zmienianiu (uspójnianiu) posiadanych wiadomości.

Zdaniem L. Wygotskiego rozwój umysłowy to zasadniczo proces społeczno-kulturowy (Schaffer, 2006). Podobnie jak Piaget Wygotski uznaje, że człowiek jest współtworcą swojego rozwoju, a nie odbiorcą doświadczeń innych ludzi. Zdaniem Wygotskiego rozwój człowieka odbywa się na trzech płaszczyznach kulturowej, interpersonalnej i indywidualnej. Na płaszczyźnie kulturowej człowiek przyswaja narzędzia kulturowe i to zarówno materialne takie jak smartfony, mapy itp jak i niematerialne takie jak systemy symboliczne np. język. Dzięki temu dziecko nie odkrywa swojego środowiska na nowo lecz negocjuje znaczenie jego elementów w trakcie doświadczania. Na płaszczyźnie interpersonalnej Wygotski szczególnie znaczenie przypisuje osobom ze środowiska bardziej doświadczonym rodzicom,

nauczycielom, czy prowadzącym szkolenie (Wygotski, 1989). Osoba ta jako posiadająca większą wiedzę, umiejętności czy ustosunkowania i zatem większy zasób narzędzi kulturowych i przekazuje je osobom mniej doświadczonym by wzmocnić ich aktywność intelektualną. Wygotski określa jeszcze sferę aktualnego i najbliższego rozwoju. Sferę aktualnego rozwoju stanowi to co człowiek już przyswoił i potrafi, natomiast sfera najbliższego rozwoju stanowi to co może przyswoić w najbliższym czasie przez własną aktywność bądź także ze wsparciem drugiej osoby. Sfera najbliższego rozwoju zmienia się wraz z procesem uczenia się, zatem wymaga nieustającego monitorowania (Gruszczyk-Kolczyńska, 2008, s. 193).

W perspektywie koncepcji rozwoju człowieka L. Wygotskiego teoria J. Piageta wydaje się być niekompletna, pomijająca znaczenie kultury w której żyjemy i którą tworzymy. Chciałbym podejść do tych dwóch różnych stanowisk nie jako wzajemnie wykluczających się lecz wzajemnie uzupełniających się. Mianowicie koncepcję Wygotskiego jako bardziej ogólną wzbogacić o teorię płaszczyzny indywidualnej traktującej o procesach rozwoju struktur poznawczych człowieka, procesach które zachodzą w mózgu o dokonania neurobiologii i neurodydaktyki (Spitzer, 2015).

2 PROCES KONTROLOWANIA I OCENIANIA WYNIKÓW UCZENIA SIĘ UCZNIĄ

Wynikiem edukacji są zmiany w uczniu, które dokonały się zgodnie z programem edukacji określone w czterech wymiarach, wiedzy, umiejętności, sprawności, ustosunkowań. Rozpatrując proces kontrolowania i oceniania konieczne jest udzielenie odpowiedzi na pytania: W jakim celu kontrolujemy? Kto ma kontrolować? Co ma być kontrolowane? Kiedy ma być kontrolowane? W jakich warunkach ma być kontrolowane? W jaki sposób ma być kontrolowane?

Jeśli celem edukacji jest przygotowanie ucznia do kierowania własnym rozwojem, własnym procesem uczenia się, to celem kontrolowania oprócz informacji o postępach ucznia jest jego wdrożenie do samokontrolowania i samooceny. Kontrolowanie i ocenianie dokonuje uczeń przy nadzorze nauczyciela i ewentualnie innych uczniów. Odpowiedź na pytanie co ma być kontrolowane i ocenianie wydaje się być najbardziej kłopotliwa, ponieważ uczniowie mają realizować zadania zgodnie ze swoimi możliwościami, ze swoją strefą najbliższego rozwoju. W konsekwencji z uwagi na zróżnicowanie uczniów ta strefa najbliższego rozwoju dla różnych uczniów może być różna. Kontrolowane i oceniane powinna być zatem poprawność czy dopuszczalność wykonanego czy zrealizowanego zadania przez ucznia i to pozwalająca wnioskować o uczniu we wszystkich 4 wymiarach, wiedzy, umiejętności, sprawności, ustosunkowań oraz postępy które uczeń poczynił w każdym wymiarze w określonym okresie np. miesiąca, półrocza itp. Proces kontrolowania powinien następować w zależności od okresu realizowania zadania w jego trakcie, by w razie potrzeby można korygować realizację oraz po zrealizowaniu.

Kontrolowanie wyników kształcenia dokonuje się przez kontrolowanie spełniania przez uczniów wymagań programowych przez sprawdzanie wykonania postawionych mu zadań i obejmuje w szerokim rozumieniu funkcje kontrolowania, a w nim:

- kontrolowanie w wąskim rozumieniu, czyli sprawdzanie zaistnienia rezultatu: czy działanie dało jakikolwiek rezultat, czy nie dało żadnego rezultatu,
- mierzenie zaistniałego rezultatu działania przez porównywanie go z celem jako wzorcem ideowym lub z wzorcem realizacji celu, dające w wyniku pomiar rezultatu,

- ocenianie zgodności uzyskanego pomiaru rezultatów kształcenia z wymaganiem dające ocenę pomiaru rezultatu jako wynik oceniania.

Mierzenie jest poznawaniem własności rzeczy ze względu na daną właściwość. Polega ono na odwzorowywaniu empirycznie poznawanych relacji między własnościami rzeczy w formalny system relacyjny oraz nazywaniu własności. Odwzorowanie to, jest homomorficzne czyli jednoznaczne i stanowi funkcję pomiarową. Funkcja pomiarowa posiada: dziedzinę (zbiór argumentów) jako zbiór możliwych relacji między własnościami oraz przeciwdziedzinę (zbiór wartości) czyli zbiór relacji między symbolami (określony przez formalną zmienną relacyjną). Dokładność mierzenia jest określona przez rozpoznane relacje między własnościami rzeczy: równości bądź różności, porządku liniowego, różnic, ilorazów. Zmienne relacyjne są wymienione od najmniej do najbardziej dokładnej informacji. Dla nazywania członów rozpoznanych relacji stosowane są następujące nazwy:

- różny, równy – numery, albo nazwy własne,
- porządku liniowego – liczby porządkowe (rangi),
- różnic – liczby,
- ilorazów – liczby.

Pamiętać jednak należy o uwadze K. Walenty, że występują tu relacje pozorne. Zauważmy, że między liczbami zachodzą wszystkie wymienione relacje. K. Walenta (1971 s. 43) relacje nierozpoznane między własnościami rzeczy, a zachodzące między liczbami nazywa relacjami pozornymi.

Mierzenie staje się teoretycznie możliwe, gdy potrafimy wskazać właściwość ze względu na którą chcemy poznawać rzeczy, oraz empirycznie możliwe, gdy potrafimy zmysłowo poznać relacje między własnościami rzeczy, oraz gdy są porównywane ze względu na tą samą właściwość.

Poznanie ludzkie najpierw jest synkretyczne i polega na rozróżnianiu rzeczy przez odróżnienie ich od tła, następnie analityczne przez wyróżnienie ich własności. Stąd najpierw rozróżniamy rzeczy i nadajemy im identyfikatory, później mierzymy ich własności.

Operacja rozróżniania nie mieści się w przedstawionej definicji mierzenia, natomiast je poprzedza. Obiekt musi być odróżniony od innych, aby można mierzyć jego własności. Rozróżniamy własności rzeczy: jakościowe oraz wielkościowe, oraz własności zbioru rzeczy ilościowe. Własności jakościowe możemy rozróżniać jakościowo albo stwierdzić, że nie potrafimy ich rozróżnić i wtedy orzec, że są równe pod danym względem, (innych relacji oprócz równy i różny nie można sensownie określić). Własności wielkościowe możemy rozróżniać wielkościowo i można wskazać która z rzeczy jest mniejsza, a która większa albo orzec, że są równe w przypadku gdy różnica między nimi jest mniejsza od naszego progu wrażliwości i nie potrafimy wskazać ani mniejszej ani większej. Jeżeli ponadto wyróżnimy dwa stany i odległość między nimi przyjmiemy za jednostkę miary bądź za wielokrotność jednostki miary własności danego rodzaju, to możemy wskazać o ile takich jednostek jedna własność jest mniejsza albo większa od drugiej. A jeśli jeden z tych stanów określimy jako wartość zerową tzn. taką której osiągnąć nie można, ale teoretycznie można znaleźć własności dowolnie bliskie, to możemy określić relację ilorazu między własnościami czyli powiedzieć, ile razy jedna własność jest mniejsza albo większa od drugiej. Własności ilościowe możemy rozróżniać ilościowo z dokładnością do elementu zbioru. Dla każdego zbioru możemy wskazać liczbę jego elementów. Istnieje odwzorowanie homomorficzne między licznosciami zbiorów, a zbiorem liczb naturalnych.

Wyróżnia się następujące skale pomiarowe (Finkelstein 1988, s. 41-43):

- dla własności jakościowych rzeczy:
 - skala różnościowa (matching scale),
- dla własności wielkościowych rzeczy:
 - skala rangowa (ranking scale),
 - skala porządkowa (ordinal scale),
 - skala przedziałowa (interval scale),
 - skala stosunkowa (ratio scale),
- dla własności ilościowych zbioru rzeczy:
 - skala naturalna (natural scale),
 - skala ilorazowa (quotient scale),
 - skala różnicowa (difference scale).

Rozróżniamy właściwości i wyróżnione ze względu na nie własności rzeczy bezpośrednio obserwowalne oraz bezpośrednio nieobserwowalne. Pomiar własności bezpośrednio obserwowalnych jest pomiarem bezpośrednim, natomiast bezpośrednio nieobserwowalnych jest pomiarem pośrednim. Bezpośrednio obserwowalne to właściwości aktualne, a bezpośrednio nieobserwowalne to właściwości potencjalne, np.: funkcje psychiczne czyli zdolności oraz produkty ich aktualizacji czyli wiadomości, umiejętności i ustosunkowania, o których wnioskować możemy na podstawie ich aktualizacji, oraz właściwości abstrakcyjne i wyznaczone przez nie lub definiujące je własności określone za pomocą pojęć konkretnych, obserwowalnych, potencjalnych jest możliwe jedynie za pośrednictwem własności obserwowalnych aktualnych (po zaktualizowaniu, zoperacjonalizowaniu) albo abstrakcyjnych za pośrednictwem własności konkretnych (po skonkretyzowaniu) jako ich wskaźników wyrażających je.

Rozróżniamy wskaźnikowanie aktualizacyjne (operacjonalizacyjne) właściwości potencjalnych oparte na relacji skutkowo-przyczynowej właściwości wskaźnikowej obserwowalnej względem właściwości wskaźnikowanej nieobserwowalnej - zachodzi tu rozumowanie redukcyjne ze skutku na przyczynę, oraz wskaźnikowanie konkretyzacyjne właściwości aktualnych abstrakcyjnych trudno obserwowalnych przez właściwości-właściwości podporządkowane konkretne następuje za pośrednictwem mierzenia ich właściwości wskaźnikowych, obserwowalnych.

Należy określić skalę pomiarową dla właściwości wskaźnikowych, a wtórnie dla właściwości wskaźnikowanej nieobserwowalnej.

Skala pomiarowa zmiennej wskaźnikowanej wynika z przyjętych założeń w procesie wskaźnikowania, co do natury relacji między zmiennymi: wskaźnikowaną tą nieobserwowalną i wskaźnikową tą bezpośrednio mierzoną z możliwymi interpretacjami wyników mierzenia zmiennych wskaźnikowych. Dotyczy to mierzenia takich właściwości człowieka jak: wiadomości, umiejętności, ustosunkowania w zakresie określonym przez program kształcenia.

Specyfika pomiaru dydaktycznego wynika z jego przedmiotu (potencjalny i określony treściowo) jest rodzajem pomiaru pośredniego. Wymagania programowe dotyczą wiadomości, umiejętności i ustosunkowań, dalej klasyfikować można je: jakościowo

po treściach których dotyczą w ramach tej samej treści według złożoności. Złożoność wiadomości dotyczy ich dokładności oraz ogólności i szczegółowości zarazem. Złożoność umiejętności dotyczy złożoności wykonywanych operacji, a w szczególności: złożoności i różnorodności elementów na których wykonywane są operacje, złożoność samych operacji, liczby i różnorodności wykonywanych operacji. Złożoność ustosunkowań dotyczy szczegółowości określenia przedmiotu ustosunkowania.

Wymagania programowe na danym etapie kształcenia należy wielowymiarowo uporządkować biorąc pod uwagę: ich rodzaj, ich treść, ich złożoności. Na uporządkowanej strukturze wymagań należy wskazać: wymagania minimalne, które powinien spełnić każdy uczeń. Zróżnicowanie uczniów jest tylko czasem potrzebnym na ich spełnienie. Wymagania które należy spełnić by podjąć naukę szkolną w określonym profilu klasie, lecz absolutnie nie obowiązujące wszystkich uczniów. To uczniowie powinni decydować które wymagania poza minimalnymi chcą spełniać i w jakim czasie.

Pomiar dydaktyczny wymaga wskazania wymagań programowych określonych w strukturze wymagań spełnianych przez uczenia. Skale pomiarowe stosowane w pomiarze dydaktycznym to (Zimny 2010):

- skala różnościowa (matching scale) dla określenia:
 - wiadomości o określonej treści,
 - umiejętności o określonej treści,
- skala rangowa (ranking scale) dla wskazania ocen przyporządkowanych do określonych wymagań w strukturze w proponowanym modelu konstruktywistycznym proponuje całkowitą rezygnację ze stopni szkolnych i wtedy ta skala przestaje mieć zastosowanie.
- skala porządkowa (ordinal scale) dla ustalenia
 - sprawności aktualizowania się umiejętności o określonej treści,
 - ustosunkowania wobec określonego przedmiotu,
- skala przedziałowa (interval scale) dla ustalenia
 - sprawności aktualizowania się umiejętności o określonej treści,
- skala naturalna (natural scale) dla ustalenia
 - liczby spełnionych przez ucznia wymagań na danym etapie kształcenia.

Celem stosowania pomiaru dydaktycznego jest monitorowanie postępów ucznia. Porównywać można uczniów między sobą, lecz szczególnie interesujące są porównania własności ucznia w kolejnych momentach czasu i sledzenie jego rozwoju. Teraz stosuje się stopnie szkolne które nie niosą informacji często nawet dla nauczyciela, nie spełniają podstawowej funkcji wdrażania ucznia do samokontrolowania i samooceny, oraz sledzenia postępów w czasie. Można powiedzieć, że stopnie szkolne odzwierciedlają zadowolenie nauczyciela z pracy ucznia, ponieważ są subiektywna, a kryteria są często mgliste, są agregatem wielu nie zawsze wiadomo jakich składowych.

Pomiar dydaktyczny stosuje się dla diagnozy spełniania przez uczniów wymagań programowych, a to może być dokonywane w dwu różnych powodów, a mianowicie: dla określenia strefy najbliższego rozwoju i planowania dalszego przebiegu procesu kształcenia ucznia (pomiar prospektywny), dla kwalifikowania ucznia do dalszego etapu kształcenia albo do wykonywania określonych prac np. (egzamin na prawo jazdy) (pomiar kwalifikatywny).

Pomiar prospektywny dotyczy diagnozowania stanu wiedzy, umiejętności, sprawności i ustosunkowań ucznia po to, by móc dostosować proces edukacji do możliwości i preferencji ucznia i by szybciej uzyskać oczekiwane rezultaty. Jest (regulatorem) instrumentem sterującym dalszym procesem kształcenia ucznia, umożliwia kształcenie zróżnicowane a nawet zindywidualizowane w zakresie programu, czasu uczenia się, środków dydaktycznych, kontekstu zadaniowego. Może i powinien być dokonywany przez nauczyciela i ucznia w sposób ciągły jako istotny element kontrolowania procesu kształcenia modyfikując i dookreślając dalszy jego przebieg. Wymaga jak najlepszej znajomości ucznia, zatem może być dokonywane na podstawie wszystkich możliwych metod poznawczych.

Pomiar kwalifikacyjny ma charakter ogólny i finalny. Jest syntetyczny i wykorzystywany do kwalifikowania uczniów do dalszego etapu kształcenia bądź do wykonywania określonych czynności. Powinien być dokonywany przez osobę zewnętrzną względem procesu kształcenia po to, by nie był zniekształcony wiedzą o uczniu, pozbawiony subiektywizmu, by zachowana została zasada sprawiedliwości zewnętrznej by wszyscy uczniowie byli badani w takich samych warunkach. Dokonywany przez osobę zewnętrzną względem procesu kształcenia powoduje także zmianę relacji nauczyciel – uczeń. Nauczyciel z egzekutora staje się mistrzem pomagającym uczniowi w przyswojeniu wiadomości, umiejętności, sprawności i ustosunkowań. Pomiar kwalifikacyjny powinien przebiegać tak, by wyniki uzyskane przez różnych uczniów w różnych miejscach i różnym czasie były porównywalne. Porównywalność może zapewnić tylko standaryzacja warunków przeprowadzania kontrolowania, mierzenia i oceniania. Warunki kontrolowania dotyczą: sposobu kontrolowania, warunków czasoprzestrzennych, osoby kontrolującej i oceniającej, zadań do rozwiązania czy wykonania, sposobu rejestrowania wyników kontrolowania i ich interpretacji oraz relacji uczniów do wymienionych warunków.

W pomiarze dydaktycznym prospektywnym można stosować: obserwację, wypytywanie, test; natomiast w pomiarze dydaktycznym kwalifikacyjnym tylko test dydaktyczny albo testy zdolności specjalnych (Zimny 2005). Należy tu jednak zwrócić uwagę na dwie istotne kwestie związane z konstruowaniem testów dydaktycznych, a mianowicie nie powinno się stosować zadań zamkniętych z podanymi możliwymi odpowiedziami do wyboru, zadań polegających na uzupełnieniu luk itp. Najlepsze zadania testowe to zadania dywergencyjne i to zarówno w odniesieniu do wyniku jak i sposobu wykonania czy rozwiązania zadania. Wtedy jednak należy podać warunki graniczne akceptowalności realizacji zadania oraz ewentualnie stosowane kryteria poprawności prezentowania realizacji. Ponadto ważne jest, że test dydaktyczny jest w swej konstrukcji inny i inne stosuje się normy do rozwiązywania zadań testowych niż w testach psychologicznych. Testy psychologiczne dotyczą sprawności aktualizowania się funkcji psychicznych i mają normy statystyczne, natomiast testy dydaktyczne dotyczą produktów tych aktualizacji i mają normy zawsze treściowe.

PODSUMOWANIE

Konstruktywizm stwarza możliwości myślenia o edukacji w nowy sposób zrywający z behawioralną tradycją. Jest to ważne z uwagi na nowe wyzwania stawiane edukacji szkolnej przez postęp cywilizacyjny, technologiczny i organizacyjny. Myślenie to przede wszystkim inaczej lokuje w procesie edukacji ucznia jako autonomiczną jednostkę i tą jego autonomię rozwija. W procesie edukacji to kontrolowanie i ocenianie postępów ucznia jest tym elementem który kreuje relacje między uczniem i nauczycielem. Konstruktywizm odbiera nauczycielowi dominującą rolę w edukacji. Staje się on osobą wspomagającą i doradzającą i dbającą o usamodzielnienie ucznia. Kontrolowanie i ocenianie wyników edukacji uczniów mam nadzieję odejście od jednowymiarowych stopni szkolnych często dezinformujących,

zniewalających uczniów i pozbawiających ich motywacji wewnętrznej. Takie konstruktywistyczne myślenie o edukacji wymaga jednak stosownego oprzyrządowania nauczyciela by mógł dalece zróżnicować proces edukacji. To oprzyrządowanie daje jednak dzisiaj postęp technologiczny.

Proces kontrolowania i oceniania powinien zachodzić na dwóch płaszczyznach perspektywnej w sposób ciągły realizowany przez ucznia i nauczyciela by wdrożyć ucznia do samokontrolowania i samooceny a także monitorowania postępów ucznia w czasie by móc na bieżąco wyznaczać sferę najbliższego rozwoju, oraz kwalitatywnej by określić czy uczeń posiadał pożądane kompetencje do wykonywania określonych czynności bądź do podjęcia dalszej edukacji. Przeprowadzane kontrolowanie i ocenianie wyników edukacji powinno być realizowane przy zastosowanych odpowiednich metod diagnostycznych oraz interpretowane zgodnie z uzyskaną informacją. Obecnie pierwszy warunek nie jest zawsze spełniany, natomiast drugi jest zawsze niespełniany. Stosowanie stopni szkolnych a tym bardziej wyliczanie z nich średniej arytmetycznej jest operacją nieuprawnioną, a ich wyniki są nieinterpretowalne.

LITERATURA:

Gruszczyk-Kolczyńska, E. (2008). *Dzieci ze specyficznymi trudnościami w uczeniu się matematyki*. Warszawa: WSiP.

Finkelstein L. (1988). *Teoria i filozofia pomiaru*, In P. H. Sydenham (Eds.), *Podręcznik metrologii*. Warszawa: WKiŁ.

Piaget J. (1966). *Studia z psychologii dziecka*. Warszawa: PWN.

Piaget J. (1967). *Rozwój ocen moralnych dziecka*. Warszawa: PWN.

Schaffer H. R. (2006). *Psychologia dziecka*. Warszawa: PWN.

Spitzer M. (2015). *Jak uczy się mózg*. Warszawa: Wydawnictwo Naukowe PWN

Tomaszewski T. (1966). *Psychologia jako nauka o człowieku*. Warszawa: KiW

Wygotski L.S. (1989). *Myślenie i mowa*. Warszawa: PWN.

Walenta K. (1971). *Podstawowe pojęcia teorii pomiaru*. In J. Koziński (Eds.), *Problemy psychologii matematycznej*. Warszawa: PWN.

Zimny T. (2005). *Stosowanie testów dydaktycznych do diagnozowania wyników kształcenia na progach edukacyjnych*, In K. Wenta E. Perzycka (Eds.), *Diagnoza pedagogiczno-psychologiczna wobec zagrożeń transformacyjnych*, Szczecin: Oficyna Wydawnicza CDiDN.

Zimny T. (2010). *Mierzenie w badaniach społecznych*, In E. Syrek (Eds.) *Czas społeczny akademickiego uczestnictwa w rozwoju i doskonaleniu civil socjety*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

KONTAKT

dr hab. Tomasz Zimny

Uniwersytet Szczeciński Instytut Pedagogiki

70-781 Szczecin ul. Czerwona 12/3

telefon +48531001211

e-mail: tomzimny@o2.pl

RESUMÉ

Monotematicky zaměřená monografie s názvem **Aktuální problémy plošného testování** je výsledkem dlouhodobě kultivovaných odborných vazeb vysokoškolských pracovišť v Česku, na Slovensku a v Polsku, která se zaměřují na otázky pedagogické, resp. edukační diagnostiky a evaluace, a státních pracovišť garantujících provádění rozmanitých testovacích šetření širokého rozsahu a zpracovávajících příslušné zprávy.

Práce prezentuje teoretické základy i zkušenosti pracovníků polské i slovenské národní agentury pro zjišťování výsledků vzdělávání, které mohou být jen stěží nahraditelným zdrojem poučení pro budování současné a verifikované teorie plošného testování u nás. Ta má potenciál k přesvědčivému argumentování pro realizaci plošného testování, z něhož mohou být generovány výstupy v podobě vypočítané přidané hodnoty vzdělávání či stanovení úrovně pokroku žáka v učení nebo jeho silných a slabých stránek či rozdílů mezi jeho učebním potenciálem a jeho reálně dosahovanými učebními výsledky. V několika kapitolách jsou zkoumány otázky diagnostikování, resp. testování v prosazujícím se paradigmatickém rámci pedagogického konstruktivismu.

Pokrok oblasti edukačních aplikací informačních a komunikačních technologií vytváří nové podmínky pro tvorbu, administraci i vyhodnocování všech variant didaktických testů v podobě e-testování, které se v okolních státech pro své přednosti úspěšně rozvíjí. Nepochybně mohou vedle hlavních cílů plošného testování vzniknout jeho další záměrné nebo vedlejší efekty v podobě e-autotestování, adaptivního testování nebo formativního testování.

Monografie je určena odborné pedagogické, případně andragogické komunitě, především z řad akademických pracovníků. Dobře poslouží všem, na nichž leží tíha zavádění plošného testování na různých stupních škol, včetně škol vysokých a institucí celoživotního vzdělávání.

SUMMARY

The single topic-focused monograph titled **Current Issues of National Assessment** is a result of long-term cooperation between Czech, Slovak and Polish university departments. The main research focus of those departments is pedagogy, i.e. educational diagnostics and assessment; national institutions supervising observing the execution of various broad surveys dealing with particular reports.

The monograph presents theoretical bases and experience of the employees of both Polish and Slovak National Associations for the Evaluation of Educational Achievement. Those will serve as the basis for creating an up-to-date and verified theory of national assessment in the Czech Republic which provides reasons for the implementation of national assessment. National assessment could lead to calculating the added value of education and determining the student's level of progress with regard to their learning, strong and weak points and the differences between their potential and achieved results. Moreover, a number of chapters are dedicated to testing in the up-and-coming paradigmatic framework of pedagogical constructivism.

The progress in the area of application of information and communication technologies paves the way for the creation, administration and assessment of all variants of didactic tests through the form of e-testing, which has been developing in the neighboring countries. Aside from the abovementioned main objectives, national assessment may also give rise to further actions such as e-self-testing, adaptive testing and/or formative testing.

The monograph is intended for the members of pedagogical and andragogical communities, particularly academic scholars. However, it will also be of use to those involved in the implementation of national assessment at all school levels, including universities and lifelong learning institutions.

REJSTRÍK

A

absolutní ukazatel, 54
administrácia testovania, 29
Aksman, 96
Antonovsky, 92
attitudes towards one's own life, 96
attitudes towards the world, 96

B

banka úloh, 35, 47, 50
behawioryzmus, 115

C

celoplošné testovania, 26
complementary diagnosis model, 93
complementary model of diagnostic and
intervention activities, 94
concept of creativity, 98

D

databáze, 47
diagnosis, 92, 93
diagnosis complexity, 93
diagnostic hypothesis, 93
diagnostic tool, 95
diagnoza, 105, 107
dyrektor szkoły, 107

E

e-egzamin, 76
egzamin, 76
EGZAMIN MATURALNY, 8, 9
egzaminator, 84
elektronické testovania, 41
e-Maker[®], 79, 80
e-ocenie, 75, 78, 81, 84
e-Test, 34, 35, 41
ewaluacja, 107
Ewaluacja, 108
externá maturita, 24
externé testovanie, 30

G

Gólek, 96
Good Lives Model (GLM), 93

Gózdź, 96

Ch

Charzyńska, 97, 98

I

Identity and Creative Thinking
Questionnaire (ICTQ), 98
intrapersonal attitudes, 96

J

Jellen, 98

K

koncepcje konstruktywistyczne, 116
konstruktywizm, 121

M

Malach, 109
maturitná skúška, 23, 25, 28
maturitný test, 25
měření přidané hodnoty vzdělávání, 53
metoda relativního přírůstku znalostí, 55
model přidané hodnoty, 54
MONITOR, 23, 25, 26

N

národné testovanie, 22
Národný ústav certifikovaných meraní
vzdelávania (NÚCEM), 22, 24, 27, 29, 33,
34, 43
negative diagnosis, 92

O

objektivita hodnotenia, 30
ocenie, 81

P

Piaget, 116
pilotáž, 79, 81
počítačem podporovaný diagnostický
systém, 47
Pokyny pro řešení úloh, 47
pomiar dydaktyczny, 120
pomiar kwalifikatywny, 121

pomiar perspektywny, 121
Portál IS Diagnostic, 49
positive psychology paradigm, 92
post-diagnostic planning, 94
Prezentační proces, 46
pridaná hodnota vo vzdelávaní, 32
Proces hodnocení úlohy, 46
proces uczenia, 116
Proces výběru testu, 46
Proces vyhodnocení a zpětné vazby, 47
procesualny model Tomaszewskiego, 115
program edukacji szkolnej, 115
projekt Kvalita 2007 až 2020, 51

R

rational strategy, 96
relativní přírůstek znalostí, 55, 56
relativní přírůstek znalostí žáka žáka, 53
relativní ukazatel, 54
risk model, 92

Ř

Řešení úlohy, 47

S

scoris assessor, 82, 85
skala pomiarowa, 119, 120
skladač, 48
Souhrnný záznam žáka, 47
Související parametry řešení úloh, 47
Special Abilities Scale (SAS), 97
spirala wsparcia, 106
systém TEST, 46, 48, 58, 68
szkoła, 108, 111

T

tester, 49
testovací nástroj, 29
testovací nástroje, 28
testování, 50, 52
testovanie, 22, 23, 27, 33
tradycyjne ocenianie, 76

U

uczeń, 115
uczien, 120
úloha, 36, 62
Urban, 98

V

vkladač, 48

W

Wiedza o społeczeństwie, 19
własności ilościowe, 118
własności jakościowe, 118
własności wielkościowe, 118
Wygotski, 116
wykonalność, 82
wynik edukacji, 117
wynik kształcenia, 117
Wysocka, 96, 97, 98

Z

zadanie kontrolne, 88
Zewnętrzny egzamin maturalny, 7
zmiana oceniania, 78
zmiany cywilizacyjne, 115

Aktuální problémy plošného testování

Vydala Ostravská univerzita, Pedagogická fakulta

Editoři: doc. PhDr. Josef Malach, CSc., doc. RNDr. Martin Malčík, Ph.D.,
Mgr. Bc. et Bc. Martina Rozsypalová
Recenzovala: doc. RNDr. Mária Lucká, PhD.
Vydání: první, 2016
Počet stran: 128
Tisk: REPRONIS s. r. o., Ostrava
Cena: neprodejné

© Ostravská univerzita

ISBN 978-80-7464-867-0