

ZÁVĚREČNÁ ZPRÁVA

Z CYKLU PILOTNÍCH PROJEKTŮ 2005–2008

HODNOCENÍ VÝSLEDKŮ VZDĚLÁVÁNÍ
ŽÁKŮ 9. TŘÍD ZŠ A ODPOVÍDAJÍCÍCH
ROČNÍKŮ VÍCELETÝCH GYMNÁZIÍ

1A1U1

© CERMAT
září 2008



cermat

Centrum pro zjišťování výsledků vzdělávání

Obsah

1. Úvod.....	4
2. Cíle pilotních projektů	4
3. Obsahové nástroje cyklu pilotních projektů	6
3.1. Test Dovednosti v českém jazyce	7
3.2. Test Matematické dovednosti	9
3.3. Test Obecné dovednosti	10
3.4. Žákovské dotazníky	11
4. Organizační zajištění cyklu pilotních projektů	12
4.1. Přípravná fáze	12
4.2. Tvorba evaluačních nástrojů.....	12
4.3. Přihlašování škol.....	13
4.4. Školení zadavatelů.....	13
4.5. Distribuce testových materiálů	14
4.6. Testování.....	14
4.7. Zpracování dat	15
4.8. Prezentace výsledků	15
5. Účast škol a žáků v cyklu pilotních projektů	16
6. Výsledky testů v cyklu pilotních projektů	18
6.1. Celkové výsledky v testech.....	19
6.2. Testové výsledky podle typu školy	20
6.3. Testové výsledky podle známek	20
6.4. Testové výsledky podle pohlaví žáků.....	22
7. Závěry a doporučení	23

1. Úvod

Závěrečná zpráva informuje o cyklu pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií*, který v letech 2005 až 2008 realizovalo Centrum pro zjišťování výsledků vzdělávání – CERMAT.

Cyklu pilotních projektů předcházela pilotní ročník 2004, v rámci kterého byli testováni žáci základních škol v Karlovarském kraji. Pozitivní zkušenosti z tohoto projektu daly základ pro budoucí rozšíření do čtyřletého cyklu s celorepublikovým rozsahem.

Od roku 2006 byl cyklus pilotních projektů finančně podporován prostředky z Evropských sociálních fondů v rámci systémového projektu Kvalita I (v roce 2005 bylo z prostředků ESF hrazeno vyhodnocení testování). Tento projekt byl v srpnu 2008 ukončen. Závěrečná zpráva tak představuje shrnutí a zhodnocení aktivit, které se v tomto projektu uskutečnily v oblasti testování v uzlových bodech vzdělávací dráhy žáků.

Závěrečná zpráva popisuje ve stručnosti obsah, průběh a výsledky testování v rámci celého cyklu a dále hodnotí jeho úspěšnost ve světle moderních trendů v didaktickém testování i změn v českém vzdělávacím systému. Závěrečná zpráva tak přináší podklady pro případné rozhodnutí o zavedení plošného testování výsledků vzdělávání do českých škol.

Cyklus pilotních projektů byl tvořen čtyřmi ročníky, které se od sebe lišily z hlediska cílů, organizace a částečně i evaluačních nástrojů. Každý ročník byl zakončen samostatnou závěrečnou zprávou, která byla předložena Ministerstvu školství, mládeže a tělovýchovy a následně zveřejněna na webových stránkách Centra pro zjišťování výsledků vzdělávání – CERMAT.¹

2. Cíle pilotních projektů

Cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* sledoval několik skupin cílů. Nejdůležitějšími cíli bylo:

- rozvíjet, ověřovat a postupně připravovat zavedení plošného zjišťování výsledků vzdělávání žáků ukončujících 2. stupeň základních škol a nižší stupeň víceletých gymnázií, a to včetně žáků se speciálními vzdělávacími potřebami a žáků ze škol s polským vyučovacím jazykem
- poskytnout žákům informace o úrovni jejich znalostí a dovedností
- nabídnout školám údaje pro svoji autoevaluaci v kontextu evaluace ostatních škol
- rozšířit údaje využitelné při výběru žáků pro studium na středních školách, případně i vyšších stupních škol
- zmapovat kvalitu školského systému v jednotlivých krajích a v celé České republice pro účely zřizovatelů škol a MŠMT

Ačkoliv každý dílčí projekt vždy naplňoval několik cílů současně, jejich důležitost se v průběhu realizace postupně měnila v návaznosti na to, jak se posouvaly úvahy o povinném testování na základních školách. Účel cyklu pilotních projektů se měnil zejména v návaznosti na probíhající kurikulární reformu a na zásadní změny ve vedení MŠMT.

¹ Souhrnná závěrečná zpráva z dílčích výstupů vychází, nedosahuje však jejich podrobnosti. V případě zájmu je možné vyhledat závěrečné zprávy z dílčích ročníků projektového cyklu na webových stránkách www.esf-kvalita1.cz.

Původně cyklus reagoval na zvyšující se autonomii škol v důsledku územně-správní reformy a kurikulární reformy, která zavádí rámcové vzdělávací programy. Cílem bylo vyvinout tzv. Národní hodnotící zkoušku, která by představovala testování žáků ukončujících první a druhý stupeň základních škol.² Všichni žáci by povinně řešili několik didaktických testů, které by ověřily míru naplnění vzdělávacích výstupů daných učebními osnovami, respektive rámcovými vzdělávacími programy. V dlouhodobém horizontu mohlo dojít k propojení výsledků jednotlivých žáků z 5. třídy, 9. třídy a maturitní zkoušky, což by umožnilo sledování tzv. přidané hodnoty vzdělávání.

Národní hodnotící zkouška by přinášela plošné a srovnatelné údaje o výsledcích vzdělávání jednotlivých žáků a tím minimalizovala potenciální negativní dopady decentralizace českého školství. Výsledky Národní hodnotící zkoušky měly sloužit k autoevaluaci a evaluaci žáků, autoevaluaci škol a výběru žáků v rámci přijímacích zkoušek. Hledání optimální obsahové a organizační podoby Národní hodnotící zkoušky a její postupné zavádění do škol mělo být hlavním cílem cyklu pilotních projektů.

Mezi lety 2006 a 2007 došlo k přehodnocení návrhu Národní hodnotící zkoušky. Myšlenka povinného a plošného testování byla opuštěna. Pilotní projekty tak dále pokračovaly v rozvíjení didaktických testů, které však měly být školám nabízeny jako nástroj dobrovolného ověřování znalostí a dovedností žáků. Těžiště projektu se proto přesunulo od evaluace k autoevaluaci žáků a škol.

Zároveň došlo k posunu v celkovém směřování projektu. Ten byl původně koncipován jako přípravný a měl směřovat k zavedení nového nástroje externí evaluace do českého vzdělávacího systému. Postupně však byla přípravná role projektu vůči zavedení Národní hodnotící zkoušky utlumena a hlavním cílem se stalo okamžité poskytování nástrojů hodnocení pro školy.

Pilotní projekty byly orientovány na několik cílových skupin. Pro každou cílovou skupinu byly připravovány specifické aktivity a výstupy (podrobněji viz kapitola 3). Hlavními cílovými skupinami byly:

- základní školy a víceletá gymnázia, kterým se rozšířily zdroje informací sloužící k vnitřnímu systému řízení kvality škol
- žáci 9. tříd základních škol a odpovídajících ročníků víceletých gymnázií, kteří získali zdroj informací pro sebehodnocení
- střední školy, kterým byly nabídnuty doplňkové nástroje pro hodnocení uchazečů o studium
- krajské úřady, které měly k dispozici hlubší údaje o kvalitě školského systému v kraji
- ostatní sociální partneři (rodiče, zaměstnavatelé, veřejnost atd.), kteří zprostředkovaně získali souhrnné údaje o kvalitě školského systému a případně výsledky konkrétních škol, které se rozhodly pro jejich zveřejnění
- Ministerstvo školství, mládeže a tělovýchovy, které informace o výsledcích vzdělávání mohlo využít při plánování školské politiky

² Vzhledem k tomu, že RVP předpokládají shodné cíle vzdělávání na druhém stupni základních škol a na nižším stupni víceletých gymnázií, došlo od roku 2006 k rozšíření testování ze základních škol i na víceletá gymnázia.

3. Obsahové nástroje cyklu pilotních projektů

Cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* využíval k naplnění výše uvedených cílů několik nástrojů. Jednalo se o:

- didaktický test Dovednosti v českém jazyce (dále jen test DČJ)
- didaktický test Matematické dovednosti (dále jen test MD)
- didaktický test Obecné dovednosti, dříve Studijní dovednosti (dále jen test OD)
- didaktický test Dovednosti v polském jazyce (dále jen test DPJ)³
- žákovský dotazník

Oblasti, které byly testy pokryty, byly zvoleny s ohledem na záměr cyklu (tj. původně příprava Národní hodnotící zkoušky) a obvyklé složení přijímacích zkoušek na střední školy. Dva testy (test MD a test DČJ) byly koncipovány jako předmětové, třetí test OD byl nadpředmětový.

Použití všech tří testů bylo opakovaně hodnoceno pedagogickou veřejností. V roce 2005 se uskutečnil mimořádný seminář, kterého se zúčastnili zástupci základních škol, s nimiž byly diskutovány různé aspekty projektu. Složení a obsahové zaměření testů bylo shledáno jako optimální. Stejně stanovisko ze strany škol zaznívalo i při ostatních příležitostech k vyjádření názorů (předmětové semináře, webové stránky, média atd.).

Jednotlivé testové nástroje byly od začátku projektového cyklu koncipovány tak, aby primárně směřovaly k ověřování dovedností, nikoliv pouze faktických znalostí žáků, a dále aby odpovídaly požadavkům, které jsou dány pedagogickými dokumenty. V případě testů DČJ a MD se proto východisky staly specifické cíle uvedené ve Standardu základního vzdělávání a zároveň i příslušné pasáže Rámcového vzdělávacího programu pro základní vzdělávání.

Díky tomuto směřování mohla zůstat koncepce testů po celou dobu cyklu (tj. i po zavedení RVP do škol) bez zásadních změn. Upravovány byly spíše detaily specifikační tabulky než celkové zaměření testů, což umožňuje vzájemné porovnávání výsledků testů ze všech dílčích ročníků testování. Stručnou obsahovou charakteristiku jednotlivých nástrojů nabízí následující podkapitoly.

Při konstrukci testových nástrojů byly dodržovány standardní principy, které zaručují kvalitu výsledků. Každý test byl sestavován na základě specifikační tabulky, která je stanovena předem vzhledem k účelu testu a určuje počet, obsah a specifický cíl jednotlivých testových úloh. Specifikační tabulka zaručuje, že test sleduje původní cíl, který je navíc transparentní, a dále že testy v jednotlivých letech jsou srovnatelné.⁴

Testy využívaly šest základních typů testových úloh: úlohy s výběrem ze čtyř nabízených odpovědí, svazky dichotomických úloh, přiřazovací úlohy, uspořádací úlohy, otevřené úlohy se stručnou odpovědí a otevřené úlohy se širokou odpovědí. Testy DČJ a testy OD byly sestaveny pouze z uzavřených úloh, které umožňovaly ověřit všechny požadované znalosti

³ Test DPJ byl použit v roce 2007 a 2008, a to výhradně ve školách s polským vyučovacím jazykem. Jeho charakter je shodný s testem Dovednosti v českém jazyce. Vzhledem k tomu, že test DPJ byl odvozován od testu DČJ, byl zadáván pouze ve dvou ročnících projektu a jeho řešení se účastnil nízký počet žáků, nebude mu věnována v předkládané závěrečné zprávě podrobnější pozornost.

⁴ Vypracování a zveřejnění specifikační tabulky je významným pozitivem testů, které připravuje CERMAT. Ostatní testy, které jsou nabízeny českým školám, z tohoto hlediska nesplňují podmínky kvalitních nástrojů.

a dovednosti. Otevřené úlohy byly použity v testech MD. Odpovědi žáků v otevřených úlohách byly vyhodnocovány hodnotiteli proškolenými v CERMATu.

Typům úloh odpovídal obsah, počet bodů a časový rozsah testu. Protože otevřené úlohy jsou časově náročnější, došlo k navýšení časového limitu pro řešení testu MD ze 40 na 60 minut (podrobněji viz 3.2.). Zbývající testy měly po celý cyklus rozsah 40 minut.

Pro každé testování byly testy připravovány ve dvou variantách – A a B⁵, aby bylo zabráněno opisování žáků. Obě varianty byly z hlediska specifikační tabulky totožné a rovněž psychometrické charakteristiky potvrzovaly jejich srovnatelnost. V případě testů DČJ a testů OD byly obě varianty sestaveny ze shodných testových úloh, avšak v odlišném pořadí; varianty testů MD obsahovaly úlohy obsahově rovnocenné, avšak s různými číselnými údaji.

Vzhledem k rozsahu souhrnné závěrečné zprávy není možné podrobně představit obsah a výsledky jednotlivých testů ani pomocí dat dokládat hlavní zjištění. Pro podrobnější informace je možné nahlédnout do dílčích závěrečných zpráv za jednotlivé ročníky.

3.1. Test Dovednosti v českém jazyce

Testy DČJ se soustřeďovaly po celý cyklus na ověřování následujících základních dovedností:

- a) Porozumění obsahu textu i jeho částí
Tato dovednost byla svým charakterem nejbližší testovým úlohám obsaženým i v testech OD. Dílčímu mezitestovému koncepčnímu překrytí nebylo možné se úplně vyhnout, jestliže měla být zachována konzistence testu s kurikulárními pedagogickými dokumenty. I pro test DČJ jde o klíčovou dovednost.
- b) Posouzení významu slov či částí textu
Vždy bylo ověřováno posouzení významu slov (výrazů) v daném kontextu.
- c) Posouzení výstavby textu
Dovednost je charakterizována nalezením závadného místa v textu a jeho opravou, resp. výběrem nejvhodnější úpravy (opravy).
- d) Porovnání textů
Dovednost v sobě zahrnuje několik operací: porozumění odbornému textu, nalezení potřebných informací v tomto textu a jejich následná aplikace v textu neodborném.
- e) Rozeznání základního charakteru textu
Ověřováno bylo rozeznání charakteru textu zejména ze stylistického hlediska.
- f) Ovládání základních pravopisných pravidel
Ověřováno bylo ovládnutí základních pravopisných pravidel v běžných situacích; pozornost byla soustředěna zejména na psaní -i/-y a interpunkci. Velmi často bylo zadání úloh konstruováno tak, aby nebyl specifikován charakter ověřovaného pravopisného jevu. Výsledky ověřování dané dovednosti tak ve srovnání se specifikovaným zadáním ověřovaného jevu skýtaly širší interpretační pole.
- g) Rozeznání nesprávně užitých tvarů slov
Dovednost jdoucí napříč několika oblastmi učiva, zejména morfologií, lexikologií, sémantikou a syntaxí.

⁵ Od roku 2006 byly ke každému testu připravovány také varianty C pro použití v rámci přijímacích zkoušek. Varianta C vycházela vždy ze shodné specifikační tabulky jako test pro řádný termín, a byla tedy s původními variantami A a B srovnatelná. Podrobnější údaje viz kapitola 4.

h) Určení slovnědruhové a větněčlenské platnosti slov

Zařazení dovednosti do testu vyplynulo z tradičního důrazu na osvojování této dovednosti ve školské praxi. Nebylo tudíž možné tuto oblast zcela pominout, nicméně v koncepci testu nehrála tato dovednost nikterak zásadní roli.

V průběhu projektu docházelo k dílčím koncepčním posunům. Poměrně podstatnou změnou v koncepci testu DČJ se stal vznik kumulované dovednosti „Posouzení jazykové správnosti“. Tato dovednost jako ověřovaná charakteristika vznikla parcelací některých z výše uvedených ověřovaných dovedností a z hlediska obsahového se dotýkala několika lingvistických disciplín.

Z uvedeného výčtu je zřejmá orientace na ověřování dovedností směřujících k práci s konkrétními texty, a to při respektování základního charakteru výuky na základních školách podle stávající pedagogické dokumentace. Všechny texty použité v testových úlohách byly autentické, tj. převzaté z každodenních situací, jako jsou běžné komunikační situace mluvené nebo sdělovací prostředky (televize, noviny a časopisy, internet). V testových úlohách nebyly tedy používány žádné vykonstruované jevy, které by vznikly jen pro účely testování. Na začátku projektu byly používány zásadně neumělecké texty, až později byly zařazovány ve funkci výchozích textů i texty umělecké.

Po zavedení tzv. komplexů úloh do vnitřní struktury testů bylo ve výstupech zohledňováno třídění úloh. Testové úlohy byly rozděleny do jednotlivých komplexů podle dvou různých kritérií:

Podle klasifikace úloh	Podle ověřovaných dovedností
komplex úloh bez výchozího textu	komplex jazykové správnosti
komplex úloh s výchozími texty uměleckými	komplex porozumění textu a významu slov
komplex úloh s výchozími texty publicistickými	komplex výstavby textu
komplex úloh s ostatními výchozími texty	

Výsledky žáků v jednotlivých ročnících projektu potvrzovaly některé očekávané skutečnosti. Velké části žáků nečinilo větší problémy řešení úloh v oblasti určování slovnědruhové a větněčlenské platnosti slov, tedy v oblasti, o níž je všeobecně známo, že je jí ve školách dlouhodobě věnována významná pozornost. Průměrných až mírně nadprůměrných výsledků žáci dosahovali při řešení úloh zaměřených na dovednosti související s významem jednotlivých výrazů v textu. Průměrné výsledky vykazovali žáci při řešení úloh ověřujících dovednost rozeznání nesprávných tvarů slov. Dlouhodobě spíše podprůměrné jsou dovednosti porovnání textů, posouzení výstavby textů a obsahu textových celků. Stabilně nejslabších výsledků dosahovali žáci při řešení úloh, jimiž byly ověřovány dovednosti související se znalostí základních pravidel českého pravopisu.

Z hlediska tradičního tematického třídění učiva mají žáci největší potíže s pravopisem, problémy se objevovaly i v různých oblastech morfologie, syntaxe a stylistiky. Po zavedení komplexů úloh byly sledovány další parametry – nejnižší úspěšnosti dosahovali žáci při řešení komplexů úloh s uměleckými výchozími texty, naopak relativně nejméně obtížné pro ně byly úlohy odvozené od textů publicistických.

3.2. Test Matematické dovednosti

Testy MD měly po celý cyklus za cíl prověřit nejen znalosti žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií, ale také postihnout a zhodnotit funkční gramotnost žáků v matematice, tj. schopnost žáků získané poznatky použít.

Testové úlohy obsahově plně respektovaly současné platné pedagogické dokumenty, jejichž výstupy byly zapracovány do specifikační tabulky testů. Po zkušenostech z pilotního ročníku 2004 byly do testových úloh zapracovány výhradně neopominutelné pojmy matematiky na základních školách, což odpovídá očekávaným výstupům daným RVP ZV. Protože testy MD ověřují nejen znalosti, ale i dovednosti žáků, bylo možné během řešení úloh používat kalkulačky i MFCH tabulky.

Úlohy v testech MD byly rozděleny z hlediska obsahu a požadovaných dovedností do tří komplexů, v nichž v posledních dvou ročnících byly počítány průměrné úspěšnosti žáků. Nejlepších výsledků žáci dosahovali v úlohách vyžadujících pouze základní matematické operace. Méně úspěšní byli v dovednostech podmiňujících úspěšné vyřešení úloh zpracovávajících číselné informace. V těchto úlohách nelze postupovat pouze nacvičeným způsobem. Úspěšnost při řešení ovlivňuje společně s nadáním žáků i časový prostor vymezený výuce matematiky. Praxe potvrzuje, že obvyklé školní tempo v matematice je pro nezanedbatelné procento populace stresující, a neumožňuje tak rozvíjet očekávané kompetence.

Nejhorších výsledků dosahovali žáci v úlohách s geometrickou tematikou, která je ve školách (už od prvního stupně ZŠ) stále více opomíjena a pravděpodobně ne vždy metodicky správně zvládnuta. Výraznější pokles se objevil zejména v posledním roce čtyřletého cyklu. Na tento problém opakovaně upozorňovaly všechny závěrečné zprávy za jednotlivé ročníky.

Během projektu se podařilo po pečlivém zvážení prodloužit dobu k řešení testu MD z původních 40 minut na 60 minut. Prodloužení aspoň částečně zohledňovalo možnost žáků se nad řešením některých úloh zamyslet, čímž se zlepšila výpovědní hodnota testu. Časově náročné jsou pro žáky též konstrukční geometrické úlohy. Je ovšem třeba konstatovat, že přes své navýšení umožňoval daný časový limit ověřit znalosti a dovednosti související pouze se zcela základními pojmy z matematiky vyučované na ZŠ.

Další změnou, která pozitivně ovlivnila výsledky žáků, bylo sjednocení dokumentu s testovými úlohami (testový sešit) a záznamového archu (dále ZA) do společného dokumentu. Zvláště v případě otevřených úloh tak žáci nemusejí v jednom dokumentu úlohu číst a v druhém ji řešit, což nenarušuje jejich soustředění a snižuje riziko omylu.

Pokud by se měla i v této zprávě zmínit zřetelná a opakující se zjištění z výsledků cyklu pilotních projektů, mohl by se stát jejich reprezentantem příklad, který v publikaci „Dítě, škola a matematika“ uvádějí prof. M. Hejny a prof. F. Kuřina. Jedná se o $C = H + p \cdot \frac{C}{100}$ příklad lineární rovnice, kde C je cena výrobku.

Absolvent vysoké školy ekonomické dlouhodobě prováděl přeceňování výrobků na základě tohoto vztahu aproximací, aniž by ho napadlo, že proměnnou C lze ze vztahu vyjádřit. Důkazy podobných „nedovedností“ v jednoduchých aplikacích získaných znalostí jsou nalézány na všech stupních vzdělávacího systému (5. třídy, 9. třídy, 4. ročníky SŠ).

Výsledky z jednotlivých ročníků cyklu pilotních projektů umožňovaly seznamovat učitele středních škol v rámci konferencí, školení a oborových setkávání s nutností přihlídnout k významným posunům ve znalostech a dovednostech žáků přicházejících studovat na střední školy. Na střední školy často i do studia ukončeného maturitní zkouškou (dále MZ) přicházejí v posledních letech studovat i ti absolventi 9. tříd, kteří ještě před pár lety pokračovali ve studiu na učitelských oborech.

Poté, co bude v roce 2010 realizována poprvé i společná část MZ, jejíž volitelnou součástí je zkouška z matematiky, bude nutné, aby střední školy ve svých ŠVP přihlédly k reálnému stavu znalostí a dovedností absolventů ZŠ. Matematika jako vyučovací předmět by měla získat adekvátní prostor nutný k doplnění a upevňování učiva ze ZŠ. Bez bezpečné znalosti základů matematiky se může snadno stát rozšiřování učiva na středoškolskou úroveň ryzí formalitou.

3.3. Test Obecné dovednosti

Cílem testů OD bylo ověřovat nespécifické dovednosti práce s informacemi, které nejsou vázány na zvláštní předmětové znalosti. Podmínkou těchto dovedností nicméně je zvládnutí základních matematických operací, čtení a jazykového citu. Mezi testy OD a zbývajícími předmětovými testy tak existovaly určité obsahové překryvy, které však nebyly považovány za nedostatek, neboť výsledky žáků v jednotlivých testech se zajímavě doplňovaly.

V prvních dvou ročnících se test nazýval Studijní dovednosti, posléze došlo k přejmenování na Obecné dovednosti. Nový název lépe vystihoval podstatu ověřovaných dovedností, které však zůstaly bez větší změny.

Základní charakteristikou testu bylo, že úlohy byly rozděleny do bloků po 3–4 položkách. Každý blok byl uvozen výchozími informacemi, které měly podobu odborného, novinového či uměleckého textu, tabulky, grafu či schématu. Všechny výchozí informace vycházely ze situací běžného života a často se jednalo o texty přímo přejaté z denního tisku či časopisů.

Testové úlohy ověřují několik základních dovedností práce s informacemi, a to na několika typech údajů. Obě dimenze tvoří základ specifikační tabulky testů OD. Průnikem obou dimenzí vznikají následující komplexy úloh (tj. skupiny úloh podle blízkosti ověřovaných specifických cílů):

- a) Vyhledání a porovnání dílčí informace slovního charakteru
Úlohy ověřují dovednost vyhledat, reprodukovat a porovnat jeden či více údajů, které jsou obsažené v kratších či delších člancích. Dobré výsledky v úlohách vypovídají o dovednosti pozorného čtení a dovednosti vracet se k textu na základě vnějších otázek, které jsou podmínkou všech pokročilejších operací souvisejících se čtenářskou gramotností.
- b) Globální porozumění informacím slovního charakteru a vysouzení nové informace
Úlohy ověřují dovednost vyhledat a porozumět hlavní myšlence výchozího textu, zaujímat k textu odstup a interpretovat jej jako celek. Na základě porozumění textu pak může být vytvářena nová informace.
- c) Vyhledání a porovnání dílčí informace číselného charakteru
Úlohy s čísly uvedenými v textu či tabulce. Úlohy ověřují, zda žáci dokáží výchozí informace pozorně vnímat a na základě jednotlivých úloh se k nim vracet. Podmínkou k porovnání dílčích údajů jsou jednoduché matematické operace (sčítání, odčítání, násobení, dělení).
- d) Globální porozumění informacím číselného charakteru a vysouzení nové informace
Úlohy ověřují komplexnější operace s čísly. Žáci mají v úlohách za úkol získat orientaci v předkládaných údajích a na základě práce s nimi dospět k informaci nové.
- e) Vyhledání a porovnání dílčí informace grafického charakteru
Úlohy předpokládají jednoduchou práci s grafy či schématy. Žáci v úlohách vyhledávají dílčí údaje a případně je porovnávají mezi sebou či s jinými předloženými údaji.

- f) Globální porozumění informacím grafického charakteru a vysouzení nové informace
Úlohy ověřují dovednost celkové orientace v graficky znázorněných informacích a dále dovednost vysuzovat z těchto informací nové údaje. Významná je zde prostorová představivost a dovednost převádět graficky znázorněné údaje do jiné podoby, například do textu či tabulky.

Průměrná úspěšnost žáků v testech OD se ve všech pilotních projektech pohybovala mezi 50 a 60 %, což se všeobecně považuje za optimální hodnotu vzhledem k rozlišovací schopnosti testu. Rovněž další psychometrické charakteristiky testů a testových úloh byly velmi dobré (reliabilita, diskriminace).

Jako obtížné se tradičně projevovaly úlohy, ve kterých žáci měli pracovat s delšími výchozími texty. Ačkoliv se jednalo o texty atraktivní (jak vyplynulo z dotazování žáků), slabší žáci měli s jejich zvládnutím potíže. Tento závěr koresponduje s výsledky výzkumů PISA v českých školách. Sníženou úspěšnost v řešení vykazovaly také úlohy s graficky znázorněnými informacemi, které nejsou ve školním prostředí pro žáky zcela běžné.

3.4. Žákovské dotazníky

Po celý cyklus pilotních projektů byl standardní součástí testování žákovský dotazník. Ten byl zadáván bezprostředně po vyřešení uvedených tří testů. Jeho účelem bylo získat údaje, které umožní hlubší porozumění a podrobnější prezentaci testových výsledků. Některé údaje byly v žákovském dotazníku zařazeny stabilně, jiné se v průběhu let měnily.

Pro účely základních analýz a prezentace výsledků žákům a školám byly prostřednictvím dotazníků zjišťovány následující údaje:

- pohlaví
- známka z matematiky a z českého jazyka a literatury na posledním vysvědčení
- typ střední školy (případně konkrétní střední škola), na kterou se žák hlásí
- průběh testování
- odhad vlastních výsledků v testech

Na základě uvedených údajů byly tříděny testové výsledky. Vybrané závěry pak byly zahrnuty do výstupních zpráv pro školy či kraje. Další byly využity při zpracování závěrečné zprávy za projekt a vycházel z nich odborný tým CERMATu při přípravě navazujícího ročníku projektu. Kromě uvedených údajů byly ve vybraných ročnících zjišťovány také další informace, které měly prohloubit znalost podmínek, za nichž vzniká testový výsledek žáků. Jednalo se zejména o tyto údaje:

- důvody volby střední školy
- způsoby ověřování znalostí a dovedností v hodinách českého jazyka a v hodinách matematiky
- vzdělání rodičů
- zajímavost testových úloh a výchozích informací pro žáky

4. Organizační zajištění cyklu pilotních projektů

Cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* zahrnoval čtyři ročníky testování, které navazovaly na pilotní ročník 2004. Ten se uskutečnil pouze v Karlovarském kraji a na základě jeho úspěšnosti rozhodlo vedení MŠMT o postupném rozšiřování projektu na celou Českou republiku (s výjimkou Prahy) a prodloužení projektu o další čtyři ročníky. V roce 2005 se testování mohly zúčastnit školy v Libereckém, Pardubickém a Karlovarském kraji, od roku 2006 pak školy ve všech krajích s výjimkou Prahy.⁶

Organizační zajištění spočívalo v plné míře na Centru pro zjišťování výsledků vzdělávání – CERMATu. Přípravu dílčích projektů CERMAT konzultoval s Ministerstvem školství, mládeže a tělovýchovy. Na realizaci projektů CERMATu spolupracoval s krajskými úřady a dále s obcemi s rozšířenou působností a konkrétními školami.

Úzká spolupráce byla navázána již v roce 2005 s krajským úřadem v Liberci. Ta se v dalších ročnících ještě více prohloubila. Liberecký kraj se stal místem pro ověřování některých novinek ještě před tím, než byly použity v celém projektu. Krajský úřad v Libereckém kraji velmi podporoval zapojení svých škol do projektu. V důsledku toho byl podíl přihlášených škol v Libereckém kraji vyšší než v ostatních krajích, a to ačkoliv jejich účast byla stále dobrovolná. Vyšší motivace škol se rovněž odrazila v lepších testových výsledcích žáků.

Jednotlivé ročníky cyklu pilotních projektů byly tvořeny několika samostatnými fázemi. V průběhu čtyř let došlo v každé z nich ke změnám, které byly vedeny snahou zkvalitnit a zefektivnit proces testování a zpracování výsledků. Klíčové změny jsou prezentovány v následujících podkapitolách.

4.1. Přípravná fáze

Příprava dílčích projektů probíhala přibližně od května do září každého kalendářního roku. Začínala po vyhodnocení výsledků předcházejícího ročníku (v roce 2005 po vyhodnocení pilotního ročníku v Karlovarském kraji). Ze strany CERMATu byl každoročně vyhotoven návrh projektové dokumentace zahrnující obsah projektu, harmonogram, podstatné změny a rozpočet. Dokumentace byla předložena MŠMT. Po jejím schválení byla zahájena komunikace s krajskými úřady a případně dalšími partnerskými organizacemi.

Na začátku každého školního roku (říjen či listopad) byl všem základním školám a víceletým gymnáziím zaslán informační dopis, který představoval projekt a nabízel školám dobrovolnou a bezplatnou účast.

4.2. Tvorba evaluačních nástrojů

Pro každý ročník byly vytvářeny testy a žákovský dotazník. Příprava testů je časově náročná. Pravidelně začínala od okamžiku schválení projektové dokumentace, která stanovila, jaké testy budou použity a jaký bude časový limit na jejich řešení.

Následně byly vytvořeny pro každý test specifikační tabulky, které vznikly na základě zhodnocení výsledků z uplynulého ročníku. Specifikační tabulky byly vytvářeny tak, aby testy odpovídaly pedagogickým dokumentům a zároveň umožňovaly částečné srovnání výsledků v čase.

⁶ V roce 2006 byla oslabena účast škol v Královéhradeckém kraji, jehož krajský úřad se odmítl podílet na distribuci testových materiálů. Část škol se přesto do projektu přihlásila. V dalších ročnících již byla účast tohoto kraje bez potíží.

Na základě specifikační tabulky byly vytvořeny testové úlohy. Ty prošly odborným posouzením a pilotním ověřením. Pilotáž testových úloh představuje zadání úloh souboru cca 200 žáků, jejichž odpovědi jsou následně statisticky zpracovány. Na základě získaných psychometrických charakteristik jsou posléze jednotlivé testové úlohy upraveny či vyloučeny. Kvalitní úlohy jsou zařazeny do konečné podoby testu.

Testy byly připraveny ve dvou variantách – A a B, které byly plně srovnatelné (viz kapitola 3.). Následně byla připravena ještě varianta C pro účely přijímacích zkoušek v Libereckém kraji.

4.3. Přihlašování škol

Přihlašování škol do projektu probíhalo v listopadu každého roku. K účasti byly školy vyzvány prostřednictvím adresného dopisu zaslání vedení škol. Součástí dopisu bylo IZO, resp. REDIZO školy a jedinečné heslo pro bezpečný přístup do přihlašovacího elektronického formuláře. Ten byl připraven prostřednictvím webového rozhraní a od roku 2007 pomocí informačního systému CERTIS.

Ředitel školy (pověřená osoba) do přihlašovací aplikace vyplnil jmenný seznam žáků, kteří se testování zúčastní. Na základě těchto údajů byly následně vytištěny v příslušném počtu testové materiály a štítky pro identifikaci konkrétních žáků.

Testování se mohli účastnit rovněž žáci se speciálními vzdělávacími potřebami, pro něž byly připraveny upravené podmínky. Úpravy se odvíjely od typu a stupně postižení, které podle instrukcí určil ředitel školy s přihlédnutím k odborným lékařským a psychologicko-pedagogickým posudkům. Údaje o typu a stupni uzpůsobení byly součástí přihlašovací aplikace (podrobněji viz kapitola 5.).

Přihlašovací aplikace prodělala v průběhu cyklu pilotních projektů změny, které vedly k jejímu zpřehlednění. V prvních dvou letech bylo pro školy obtížné se v přihlašovací aplikaci orientovat, což vyplývalo jednak z toho, že to pro ně byla nová zkušenost, jednak z objektivních nedostatků v aplikaci. Pro zkvalitnění služeb byl vyvinut informační systém CERTIS, který zohledňoval nejčastější potíže v původní aplikaci, a dále byly zřízeny informační linky, kam školy mohly telefonovat v případě nejasností při vyplňování formulářů. Díky těmto úpravám již přihlašování probíhalo bez větších potíží.

4.4. Školení zadavatelů

Testování na školách probíhalo pod dohledem vyškolených zadavatelů. Těmi byli učitelé daných škol, kteří byli jmenováni vedením školy a následně proškoleni v zásadách standardizovaného testování. Jejich úkolem bylo zajistit standardní podmínky po celou dobu testování. Podmínkou pro jmenování do pozice zadavatele bylo, že učitel nebyl aprobevován v českém jazyku a v matematice.

V prvních dvou ročnících (2005 a 2006) se školení uskutečnilo v lednu či únoru příslušného roku přímo v obcích s rozšířenou působností, kam se shromáždili jmenovaní zadavatelé. Školení prováděli lektoři, kteří nejprve prošli instruktáží v CERMATu. Tento postup se ukázal finančně a časově náročným a zároveň po obsahové stránce málo efektivním.

Proto byly v následujících letech materiály ke školení a organizaci vyvěšeny na zabezpečených webových stránkách, kam měly přístup pouze přihlášené školy. Materiály podrobně popisovaly činnosti všech osob účastnících se testování, zvláště pak zadavatelů testů. Případné nejasnosti mohli vyučující konzultovat prostřednictvím e-mailu či telefonické linky. Tento postup byl školami hodnocen jako optimální a rovněž v žakovských dotaznících se ukazovalo, že míra nepřesností v průběhu zadání testů byla srovnatelná s prvními dvěma

ročníky. Celkový počet tříd, v nichž došlo při zadání testů k odchylce od standardních podmínek, byl zanedbatelný.

4.5. Distribuce testových materiálů

Na základě údajů v přihlašovacích databázi bylo v zimě příslušného roku (prosinec – leden) vytištěno potřebné množství testových sešitů, záznamových archů a žákovských dotazníků. Testové materiály byly zkompletovány do zásilek pro jednotlivé školy, respektive třídy. Zásilky byly uloženy v bezpečnostních fóliích, které zaručovaly utajení testů.

V letech 2005–2007 k distribuci využito spolupráce s krajskými úřady. CERMAT zajistil rozvoz materiálů, které byly připraveny pro jednotlivé školy, na krajské úřady. Ty již dále zorganizovaly jejich rozvoz do jednotlivých škol. Za bezpečné uložení zásilek ve škole do doby testování zodpovídali ředitelé.

V posledním ročníku 2008 byla na základě výběrového řízení vybrána firma, která měla zajistit rozvoz připravených zásilek přímo do škol. Ačkoliv ze strany CERMATu byly všechny podklady řádně připraveny, samotný rozvoz proběhl se zásadními chybami, neboť najaté firmě se nepodařilo materiály na všechny školy dopravit včas. Mohlo tak dojít k ohrožení jednotného termínu testování. Promptním zásahem pracovníků CERMATu však bylo riziko odvráceno. Okolnosti selhání externí firmy nicméně byly podrobně zmapovány. Na předmětnou problematiku bude kladen důraz při distribuci materiálů v dalších projektech.

4.6. Testování

Testování probíhalo vždy v jednotném termínu a podle jednotného testovacího schématu tak, aby byla zaručena srovnatelnost žákovských výsledků. Termín byl pravidelně stanovován na přelom ledna a února s ohledem na jarní prázdniny a časovou náročnost zpracování výsledků.

Testování probíhalo v jeden den, během něhož žáci řešili tři testy a vyplňovali žákovský dotazník. Ačkoliv je to pro žáky náročné a některé školy vyzývaly k rozdělení testování do dvou dnů, nebylo možné z organizačních a bezpečnostních důvodů termín změnit. Navíc platí, že rovněž při přijímacích zkouškách na střední školy žáci často řeší tři testy a věrohodnost výsledků tím není zpochybněna. Zpětná vazba ze škol, kterou se CERMAT snažil aktivně získat, ukázala, že větší část škol nepovažuje testování v jeden den za problém.

Testovací schéma bylo každoročně přizpůsobeno žákům se speciálními vzdělávacími potřebami. Jednou z úprav, která se jim nabízí, je prodloužení časového limitu až o 100 %. Protože všichni žáci musí zahájit řešení příslušného testu ve stejný čas, vznikla navýšením časového limitu pro žáky se SVP dlouhá přestávka pro intaktní populaci. Školy ji obvykle dokázaly zajímavě využít pro odpočinek či diskusi o testu.

Testy ve třídách zadávali proškolení zadavatelé. Ti měli zejména zajistit správné administrování testu (tj. nalepení identifikačních štítků, vysvětlení způsobu záznamu odpovědí a dohled na pracovní klid). Po ukončení testování byly podle pokynů vyplněné záznamové archy odvezeny ke zpracování. V prvních letech zajišťovaly svaz materiálů krajské úřady, v posledním ročníku jej organizovala externí firma.

Během testování přijížděli do náhodně vybraných škol pracovníci z krajských úřadů a CERMATu, aby zjistili, jak jsou dodržována pravidla testování a seznámili se s případnými potížemi. Žádné závažné pochybení nebylo nikdy shledáno, z dílčích nesrovnalostí CERMAT vyvodil doporučení pro následující ročník.

4.7. Zpracování dat

Vyplněné záznamové archy a žákovské dotazníky byly svezeny na centrální místo, kde došlo k jejich roztrídění a kontrole. Testy s uzavřenými úlohami byly předány k digitalizaci, kterou prováděla externí firma vybraná na základě výběrového řízení. Testy MD s otevřenými úlohami byly nejprve zpracovány hodnotiteli.

Hodnotitelé byli externí spolupracovníci CERMATu (učitelé matematiky), kteří byli proškoleni v CERMATu. Ve spolupráci s předmětovými koordinátory pro matematiku následně provedli ohodnocení žákovských odpovědí na otevřené úlohy v náhodném vzorku záznamových archů. V prvních dvou letech projektu probíhalo hodnocení úloh ručně, tj. hodnotitelé výsledky zapisovali přímo do záznamových archů. V dalších letech byl použit elektronický postup, kdy nejprve došlo k naskenování záznamových archů a hodnotitelé pak opravovali odpovědi, které se jim zobrazovaly na monitorech. Práci hodnotitelů bylo možné sledovat, průběžně je informovat o případných aktualizacích pokynů k hodnocení, eventuálně je včas upozornit na systémovou chybu, které se dopouštějí. Důležitým posunem byla dále možnost hodnocení úlohy dvěma nezávislými hodnotiteli. Chybovost hodnocení tak byla minimalizována, neboť při případném rozdílu v ohodnocení úlohy dostává úlohu k posouzení správce hodnocení. Elektronické hodnocení navíc přináší možnost zapojení libovolného počtu vyškolených hodnotitelů bez ohledu na jejich bydliště (dosavadním omezením je pouze podmínka kvalitního připojení k internetu).

Výsledkem digitalizace je soubor se všemi žákovskými odpověďmi z jednotlivých testů a žákovského dotazníku. Soubor dat procházel opakovanými kontrolami a následně byl předán ke statistickému zpracování do analytického oddělení CERMATu.

Statistické zpracování dat zahrnuje jednak určení základních psychometrických charakteristik, na jejichž základě se stanoví kvalita testů, a dále spočítání jmenovitých výsledků jednotlivých žáků.

4.8. Prezentace výsledků

Testové výsledky byly zpracovány do několika výstupů, které byly určeny:

- žákům
- školám
- krajským úřadům
- MŠMT

Pro žáky byla vytvořena adresná osvědčení o výsledcích v testech, která byla na školy zaslána na přelomu března a dubna každého kalendářního roku, tedy před konáním přijímacích zkoušek na střední školy. Osvědčení obsahovala výsledky daného žáka, tj. skóre a úspěšnost v jednotlivých testech, a jejich porovnání s průměrnými výsledky žáků dané školy a všech zúčastněných žáků v kraji a v ČR, a to v číselné i grafické podobě. Většina škol osvědčení velmi chválila a předávala je žákům ve slavnostní atmosféře. Část žáků se rovněž osvědčením prezentovala při přijímacích zkouškách na střední školy.

Pro školy byly připraveny výstupní zprávy. Ty obsahovaly jmenovité výsledky jednotlivých žáků a dále souhrnné výsledky žáků dané školy i jednotlivých tříd a jejich porovnání s výsledky všech testovaných žáků v ČR a žáků v daném kraji. Zprávy zahrnovaly výsledky podle známek na vysvědčení či podle typu střední školy, na niž se žáci hlásí.

Zprávy byly školám zasílány v dubnu každého roku. Během prvních dvou ročníků (2005 a 2006) byly zprávy připraveny v tištěné verzi a rozesílány poštou. Následně se přistoupilo

k elektronické verzi, která je efektivnější. Zprávy byly zveřejněny v informačním systému CERTIS, kam mají školy přístup pouze pod svým jedinečným heslem.

Souhrnné zprávy byly každoročně připravovány rovněž pro krajské úřady. Obsah krajských zpráv byl odlišný; obsahoval výhradně agregované údaje, tj. průměrné výsledky žáků v jednotlivých krajích. Pokud krajské úřady chtěly získat adresné údaje, musely oslovit přímo školy. Krajské zprávy byly vyhotoveny obvykle v červnu daného roku a v tištěné verzi rozeslány na krajské úřady.

Posledním výstupem byly závěrečné zprávy za jednotlivé ročníky cyklu pilotních projektů. Zprávy byly primárně určeny Ministerstvu školství, mládeže a tělovýchovy a shrnovaly obsah, průběh a výsledky projektu. Jejich cílem bylo zhodnotit uplynulý ročník a poskytnout podklady pro rozhodnutí o podobě navazujícího ročníku. Po schválení byly závěrečné zprávy vždy zveřejněny na webových stránkách CERMATu, kde se s nimi mohla seznámit široká veřejnost.

CERMAT pravidelně využíval k informování o projektu i média se širokým dopadem. V průběhu každého ročníku vycházely informační články v pedagogickém tisku (Učitelské noviny, Učitelské listy) i v denním tisku (Lidové noviny, Hospodářské noviny, MF Dnes).

5. Účast škol a žáků v cyklu pilotních projektů

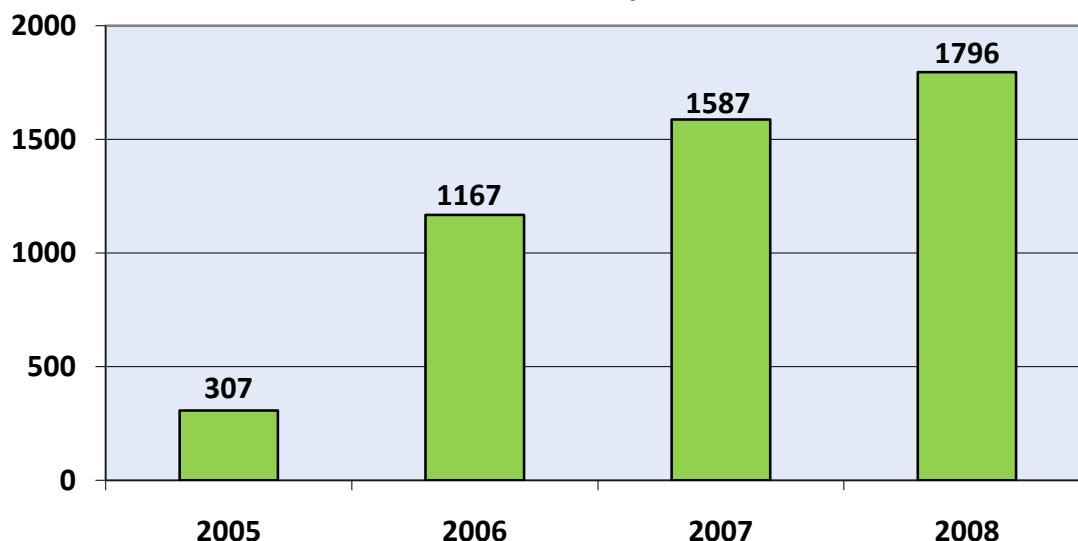
Zájem škol o účast v cyklu pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* se každým rokem výrazně zvyšoval.

Pilotní ročník 2004 se uskutečnil pouze v Karlovarském kraji. Po dohodě s krajským úřadem se tam testování zúčastnily všechny školy. V dalších letech byla zdůrazňována dobrovolnost, a proto se část škol do projektu nehlásila. Důvodem byla jednak malá podpora účasti v projektu ze strany zřizovatelů, obava před náročností a zřejmě i zahlcení jinými aktivitami, včetně testů. Postupně však pilotní projekty získávaly popularitu a důvěru ze strany škol, což se projevovalo nárůstem účasti. To dokládají následující tabulky.

Počet žáků řešících jednotlivé testy

	2005	2006	2007	2008
Test MD	10 969	49 671	58 362	68 649
Test DČJ	10 973	49 708	58 464	68 585
Test OD	10 953	49 717	58 229	68 357

Počet zúčastněných škol



Jak vyplývá z tabulky i grafu, počet testovaných žáků byl ve všech ročnících vysoký a každoročně se výrazně zvyšoval. Podíl testovaných žáků se pohyboval mezi 50 a 97 % ze všech žáků v 9. třídách ZŠ a odpovídajících ročnících víceletých gymnázií v krajích, jež se projektů účastnily (tj. testy řešila nadpoloviční většina všech žáků). Téměř plná účast byla v roce 2005 v prvním ročníku projektu, který se konal ve třech krajích. Posléze došlo k rozšíření projektu na celou Českou republiku s výjimkou Prahy. Vzhledem k dobrovolnosti projektu byla nadpoloviční účast škol považována za dobrou. Postupně se zvyšovala v závislosti na tom, jak narůstalo povědomí a důvěra škol vůči projektu, takže v závěru již dosahovala 73 %.

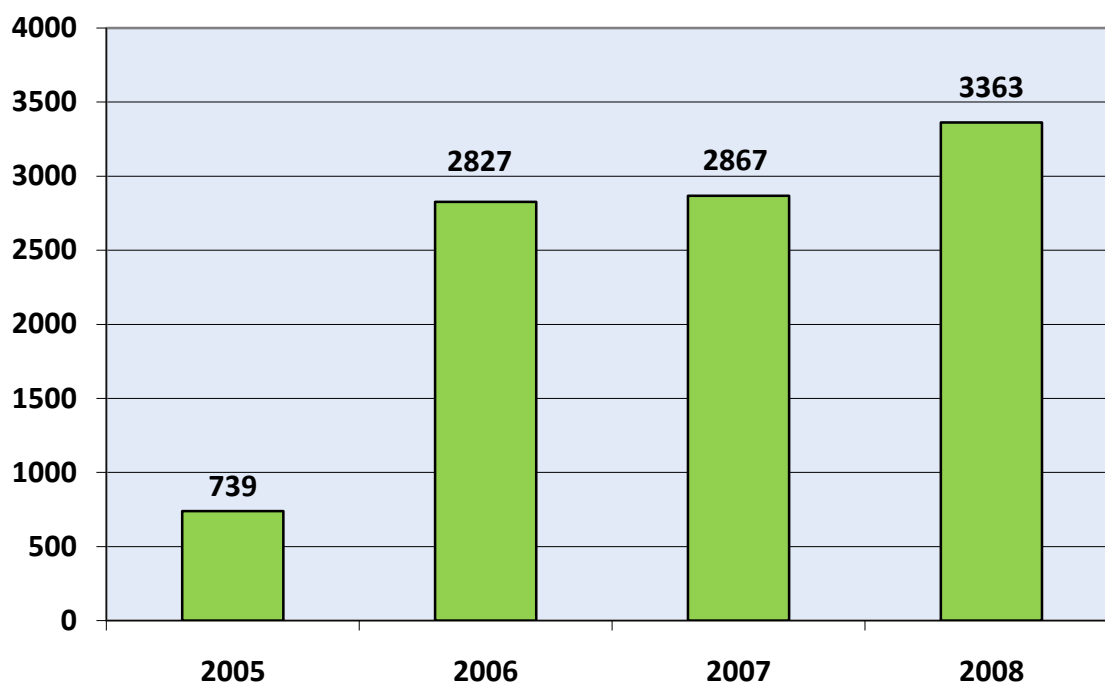
Od roku 2006, tj. druhého ročníku, byly pilotní projekty otevřeny nejen základním školám, ale rovněž víceletým gymnáziím. Zájem víceletých gymnázií v porovnání se základními školami byl podle očekávání relativně nízký. Důvodem je zejména skutečnost, že žáci 9. tříd ZŠ se ocitají v uzlovém bodu své vzdělávací dráhy, zatímco gymnazisté jsou uprostřed plynulé vzdělávací dráhy. Proto mají ZŠ a jejich žáci větší zájem o zhodnocení znalostí a dovedností v momentě zlomu. Gymnazisté mezi všemi testovanými žáky tvořili menšinu – cca 10–15 %. Ačkoliv se jednalo o malé zastoupení, dovolovalo statistické porovnání výsledků obou skupin žáků (viz kapitola 6.).

Účast škol a žáků v jednotlivých krajích nevykazovala žádné odchylky. Podíl zúčastněných škol vůči všem školám byl v jednotlivých krajích přibližně stejný. Výjimku dlouhodobě tvořil Liberecký kraj, s jehož krajským úřadem CERMAT navázal úzkou spolupráci. Krajský úřad školám účast v projektech doporučoval a z vlastních zdrojů hradil například přípravu testů varianty C, které mohly využít střední školy pro přijímací zkoušky u žáků, kteří neprošli řádným testováním.

Pilotních projektů se účastnili vedle žáků z intaktní populace rovněž žáci se speciálními vzdělávacími potřebami. Jim byly při testování nabídnuty upravené podmínky týkající se grafické podoby testových materiálů, časového limitu pro řešení testů a v některých případech obsahu. Žádná z úprav nesnižovala náročnost testu a výsledky žáků se SVP tak bylo možno porovnat s výsledky ostatních žáků.

Počet žáků se SVP, kteří se cyklu pilotních projektů účastnili, se průběžně zvyšoval. Jejich počet uvádí následující graf. Podrobné údaje o organizaci i výsledcích žáků se SVP jsou uvedeny v samostatné zprávě, která shrnuje testování žáků se SVP v letech 2006 až 2008.

Počet žáků se SVP



Počtem žáků, kteří se testování účastnili, byl cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* v českém prostředí unikátní. S výjimkou několika málo šetření, které jednorázově zjišťovaly úroveň znalostí a dovedností žáků v určitých předmětech, neexistuje dosud podobně rozsáhlý soubor informací o výsledcích vzdělávání na českých základních školách, a to jednak z hlediska počtu testovaných žáků, jednak z hlediska porovnání výsledků v čase. Díky počtu testovaných žáků a opakovanému testování ve čtyřech letech jsou výsledná zjištění zajímavá a užitečná pro instituce určující školskou politiku, pro školy samotné a rovněž pro jednotlivé žáky, kteří se mohou porovnat s ostatními.

6. Výsledky testů v cyklu pilotních projektů

Cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* ověřoval znalosti a dovednosti konkrétních žáků. Jejich dílčí výsledky byly prezentovány na osvědčeních. Dále byly zpracovány s ostatními údaji do agregovaných psychometrických charakteristik, které popisují:

- kvalitu testu jako evaluačního nástroje
- úroveň znalostí a dovedností žákovské populace

Druhé jmenované údaje slouží jako zdroj k hlubšímu porozumění stavu vzdělávacího systému a plánování školských politik. Hlavní zjištění týkající se výsledků žáků v testech uvádějí následující tabulky a grafy.

6.1. Celkové výsledky v testech

Základní psychometrickou charakteristikou je průměrná úspěšnost žáků v testech.⁷ Ta představuje v procentech průměr podílů mezi skóre, které jednotliví žáci získali v testech, a celkovým možným počtem bodů. Průměrná úspěšnost tak udává část testu, kterou žáci dokázali správně vyřešit. Ideální hodnota průměrné úspěšnosti se pohybuje zhruba mezi 50 a 60 %, kdy test ověřuje "co žáci umí", a nikoliv, co neumí, a zároveň test dokáže dobře rozlišit mezi žáky s dobrými a slabými výsledky.

Průměrná úspěšnost žáků v testech

	2005	2006	2007	2008
Test MD	42 %	43 %	47 %	36 %
Test DČJ	63 %	66 %	56 %	56 %
Test OD	55 %	61 %	66 %	49 %

Jak vyplývá z údajů v tabulce, žáci dosahovali ve většině testů přiměřené úspěšnosti. Dlouhodobě nejnáročnější byl test Matematické dovednosti, což bylo předpokládáno na základě výsledků jiných šetření zjišťujících znalosti a dovednosti žáků v různých předmětech či oblastech. Matematika patří rovněž mezi školní předměty, v nichž žáci dosahují horších známek.⁸ Zároveň se jedná o test, který využívá otevřené úlohy, v nichž žáci dosahují a priori nižších výsledků než v úlohách uzavřených. Protože obsah testů odpovídá pedagogickým dokumentům, není možné jejich náročnost snížit zahrnutím jednodušších testových úloh.

V případě testu MD došlo v průběhu cyklu k několika úpravám organizačního charakteru, které měly za cíl zvýšit výpovědní hodnotu testu a případně i úspěšnost žáků v řešení. Jednalo se o změnu zápisu odpovědí (sloučení testového sešitu a záznamového archu) a o prodloužení časového limitu na 60 minut, což dovolilo rozšířit ověřované učivo. Zároveň byly vyučujícím poskytnuty podrobné údaje o oblastech, v nichž jsou žakovské znalosti a dovednosti nejslabší, tak, aby se na ně mohli ve výuce soustředit.

Ve všech třech testech došlo ve sledovaném období k mírnému poklesu průměrné úspěšnosti žáků. Patrné je to zvláště na testu OD, kde mezi lety 2007 a 2008 poklesla průměrná úspěšnost o 17 procentních bodů. Příčiny tohoto i ostatních mírnějších rozdílů jsou zejména dvě. První představuje postupné zvyšování počtu a změny charakteru zúčastněných škol. Zpočátku se jednalo o školy silně motivované svým zřizovatelem a o školy velmi kvalitní (a proto se zajímající o výsledky svých žáků), postupně se zvyšovala i účast škol s běžnou žakovskou populací. Druhá příčina se týká záměrné snahy o postupný posun testů k výhradnímu ověřování dovedností (potažmo kompetencí), které jsou pro žáky vzhledem k pedagogické praxi na školách spíše náročnějšími úkoly. Změny jsou patrné ve specifikačních tabulkách jednotlivých testů, které jsou sice mírné, avšak přesto způsobují pokles v úspěšnosti žáků. Tyto zdokumentované a zdůvodněné změny ve specifikačních tabulkách jsou důvodem, proč lze výsledky v jednotlivých ročnících porovnávat pouze částečně.⁹

⁷ Kvalita testů se odvozuje rovněž od reliability, rozptylu odpovědí a diskriminační schopnosti. Všechny testy v těchto parametrech vykazovaly dlouhodobě dobré hodnoty.

⁸ V žakovském dotazníku byly zjišťovány známky žáků na posledním vysvědčení. Při porovnání se opakovaně ukázalo, že žáci získávají signifikantně slabší známky v matematice než v českém jazyku a literatuře.

⁹ Spíše než porovnávat celkové výsledky v testech je vhodné porovnávat vývoj v jednotlivých komplexech úloh.

Jak bylo řečeno v kapitole 5., testování se zúčastnili také žáci se speciálními vzdělávacími potřebami, pro něž byly podmínky řešení testů upraveny. Jejich testové výsledky byly následně porovnávány s výsledky intaktní populace. Podrobné údaje obsahuje samostatná zpráva o testování žáků se speciálními vzdělávacími potřebami za období 2006–2008.

6.2. Testové výsledky podle typu školy

Třech ročníků cyklu pilotních projektů (2006–2008) se vedle žáků 9. tříd základních škol účastnili také žáci odpovídajících ročníků víceletých gymnázií. Ačkoliv byl jejich podíl v celku testovaných žáků malý, jednalo se vždy o dostatečný počet umožňující statistické porovnání. Průměrné úspěšnosti žáků základních škol a žáků víceletých gymnázií uvádí následující tabulka.

Průměrná úspěšnost žáků ZŠ a G v testech

	2006		2007		2008	
	ZŠ	G	ZŠ	G	ZŠ	G
Test MD	40 %	72 %	45 %	73 %	34 %	62 %
Test DČJ	64 %	83 %	55 %	74 %	54 %	77 %
Test OD	60 %	81 %	65 %	84 %	47 %	70 %

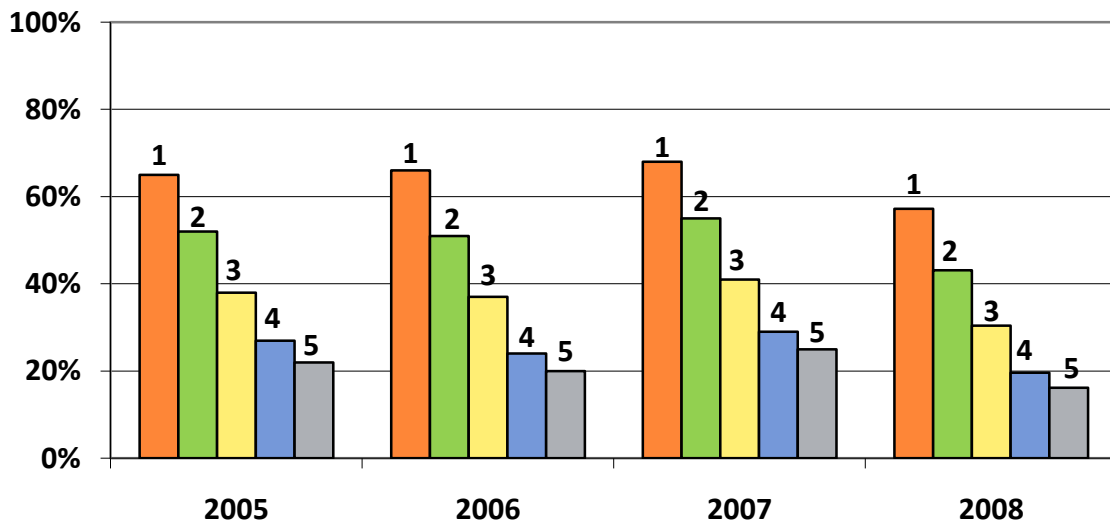
Výsledky žáků víceletých gymnázií byly ve všech testech signifikantně vyšší než výsledky žáků základních škol. Největší rozdíly byly dosahovány v testu MD. Vyšší úspěšnost gymnazistů v testech byla předpokládána, velikost rozdílů však byla překvapující.

Její příčiny spočívají v odlišném složení žakovské populace z hlediska nadání a motivace. Nicméně při podrobnějším pohledu na průměrné výsledky žáků konkrétních škol se ukázalo, že některé základní školy dosahují podobné průměrné úspěšnosti žáků při řešení testů jako některá víceletá gymnázia. Zjištění cyklu pilotních projektů týkající se úspěšnosti žáků základních škol a gymnázií by měla vstoupit do odborné diskuse o vlivu existence víceletých gymnázií na kvalitu výuky na základních školách a vzdělávacích možnostech žáků vzhledem k jejich motivaci, nadání atd.

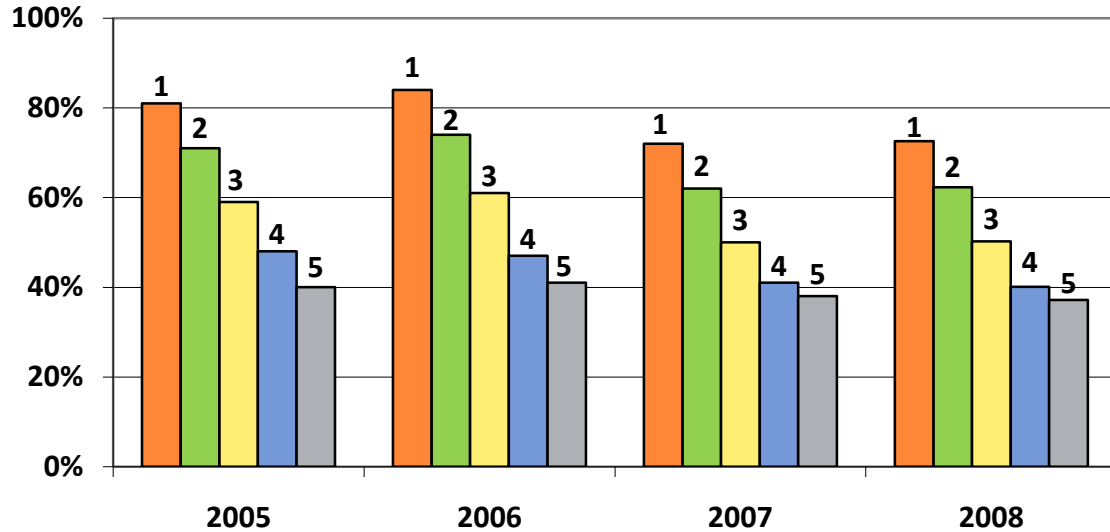
6.3. Testové výsledky podle známek

Po celý cyklus pilotních projektů byla součástí žakovského dotazníku otázka zjišťující známku na posledním vysvědčení. Údaje vyplněné žáky nebyly kontrolovány, což mohlo vést k určitému zkreslení. Analýzy srovnávající výsledky v testech se školním prospěchem žáků byly přesto velmi zajímavé. Žáci byli podle známek na vysvědčení z matematiky a z českého jazyka a literatury rozděleni do 5 skupin a pro každou z nich byla spočítána průměrná úspěšnost žáků v řešení testů MD a DČJ.

**Průměrná úspěšnost žáků
v testu Matematické dovednosti podle známek**



**Průměrná úspěšnost žáků
v testu Dovednosti v českém jazyce podle známek**



Předpokladem porovnání je, že mezi výsledky v testu a školními známkami z daných předmětů by neměl být výrazný rozdíl. Čím lepšího prospěchu žáci dosahují, tím lepší by měly být jejich výsledky v testech. Takový vztah vypovídá o podobnosti nároků, které na žáky má škola a externí testy.

Souvislost mezi testem a školními známkami se ve všech ročnících ukázala jako relativně vysoká. Úspěšnost žáků v testu pravidelně klesala s horšícím se prospěchem. Výjimku tvoří výsledky žáků se známkou 5 na vysvědčení, kteří v testech dopadli podobně jako čtyřkaři (v některých ročnících nebyl rozdíl v jejich výsledcích statisticky signifikantní, tj. výsledky byly shodné).

Analýzy podle známek zjistily velký rozptyl ve výsledcích žáků se stejným prospěchem, a to jak jednotlivě, tak v průměru za celou školu. Na některých školách například jedničkáři dosáhli v testu průměrné úspěšnosti 57 %, na jiné 85 %. Klasifikační normy v jednotlivých školách i třídách jsou velmi málo srovnatelné. Toto zjištění má velký význam pro přijímací zkoušky na střední školy, které obvykle zahrnují prospěch mezi kritéria výběru.

Porovnání úspěšnosti podle známek je zajímavé z hlediska:

- základních škol či víceletých gymnázií – na základě analýz mohou zjistit, jak mají jednotliví vyučující nastaveny klasifikační normy, a porovnat své klasifikační normy s celorepublikovým průměrem
- střední školy – na základě analýz mohou posoudit, zda prospěch jejich uchazečů odpovídá stejně vysokým znalostem a dovednostem jako v celorepublikovém průměru
- MŠMT a další instituce zodpovědné za školskou politiku – na základě analýz mohou učinit rozhodnutí o změnách ve způsobu hodnocení či připravit materiály, které pomohou hodnocení zpřesnit

6.4. Testové výsledky podle pohlaví žáků

Analýzy výsledků se každoročně dělaly také podle pohlaví žáků. Porovnávala se průměrná úspěšnost dívek a chlapců v řešení jednotlivých testů. Počet dívek a chlapců byl ve všech projektech srovnatelný. Výsledky v testech uvádí následující tabulka.

Průměrná úspěšnost dívek a chlapců v testech

	2005		2006		2007		2008	
	dívky	chlapci	dívky	chlapci	dívky	chlapci	dívky	chlapci
Test MD	42 %	43 %	41 %	45 %	46 %	48 %	37 %	38 %
Test DČJ	66 %	61 %	69 %	63 %	60 %	53 %	60 %	55 %
Test OD	55 %	54 %	61 %	61 %	66 %	66 %	49 %	50 %

Rozdíly v dívčích a chlapeckých výsledcích jsou nejmenší v testu Obecné dovednosti. Ve zbývajících dvou testech byly rozdíly téměř každý rok statisticky významné. Přesto však nedosahovaly velké hodnoty. Chlapci byli úspěšnější v testu MD, dívky naopak byly úspěšnější v testu DČJ. Tato zjištění odpovídají výsledkům ostatních pedagogických výzkumů, včetně projektu PISA.

CERMAT si dlouhodobě klade za cíl tvorbu genderově neutrálních testů, v nichž by nedocházelo ke zvýhodňování dívek či chlapců prostřednictvím využívání takových obsahů či kontextů testových úloh, které jsou typické pro dívčí nebo chlapeckou životní zkušenost. Pokud je test neutrální, ale výsledky dívek a chlapců přesto nejsou srovnatelné, znamená to, že genderové nerovnosti existují již ve školním vzdělávání. Na to je potřeba školy upozornit a doporučit jim postup k jejich odstranění.

7. Závěry a doporučení

Cyklus pilotních projektů *Hodnocení výsledků vzdělávání žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií 2005–2008* naplnil své cíle, a může být proto považován za úspěšný. Původní záměr projektu (připravit a zavést Národní hodnotící zkoušku) byl na začátku cyklu změněn, což přineslo řadu obtíží. Projekt byl uzpůsoben novým okolnostem, dále pokračoval a získával stále větší popularitu mezi školami.

Školy cyklus vnímaly jako užitečnou příležitost ověřit znalosti a dovednosti svých žáků a porovnat je s ostatními. Využívání externích evaluačních nástrojů je klíčovou součástí školního hodnocení. Jednotliví žáci mohou na základě kombinace školního prospěchu a testových výsledků získat podrobnější a přesnější představu o svých schopnostech. Jednotlivé školy mohou prostřednictvím testových výsledků svých žáků prohloubit svoji autoevaluaci, neboť lze porovnat pedagogické působení jednotlivých vyučujících i školy jako celku s ostatními školami. V podmínkách decentralizovaného vzdělávacího systému s vysokou autonomií jednotlivých škol je navíc i pro instituce určující školskou politiku důležité mít k dispozici srovnatelné informace o kvalitě škol.

Cyklus pilotních projektů byl v podmínkách českého školství jedinečný a nový. V kontextu probíhající kurikulární reformy se postupně hledala optimální obsahová a organizační podoba testování. Je proto pochopitelné, že měl některé nedostatky. CERMAT chápal celý cyklus jako službu školám a úzce s nimi spolupracoval. Snažil se být proto vůči školám maximálně otevřený, a to jednak ve smyslu zkvalitňování projektu podle potřeb škol, jednak plné transparentnosti. Řídké chyby, ke kterým došlo, byly otevřeně přiznány v dílčích závěrečných zprávách. Nedostatky byly okamžitě napravovány a v žádných dvou po sobě jdoucích ročnících se neopakovaly. Celý cyklus se postupně stával velmi kvalitním, což školy opakovaně chválily. Dokládá to i výrazný nárůst počtu škol, které měly o účast v projektu zájem. V závěru se testování účastnily 2/3 ze všech žáků 9. tříd ZŠ a odpovídajících ročníků víceletých gymnázií v celé České republice (s výjimkou Prahy, která se nemohla zúčastnit).

Vedení škol, vyučující, žáci i zřizovatelé se u různých příležitostí (semináře, ankety, cílená šetření, média atd.) pozitivně vyjadřovali k pokračování projektu. Uvědomují si potřebu externích evaluačních nástrojů, které by se propojovaly s běžným školním hodnocením. Důvěru mají zejména k nástrojům, které jsou vytvářeny standardním postupem a korespondují s požadavky pedagogických dokumentů. Testy vytvářené v CERMATu, na rozdíl od velké části testů nabízených komerčními agenturami takové jsou.

Názory vyučujících například jsou: „*Vy si ty děti nějakou dobu vedete, máte představu o tom, co dokáží, pracujete s nimi. Díky testům pak zjistíte, jak si to dítě stojí vůči ostatním. Tak vlastně vidíte, co za Vámi zůstalo.*“, „*Sto procentně by to mělo pokračovat.*“, „*Pro děti je to taková generálka – ve známém prostředí a bez velkého stresu si vyzkoušejí, co v nich je a jak si na tom stojí v porovnání s ostatními.*“, „*Je to objektivní měřítko pro porovnání škol pro tu školu, pro její kantory a pro rodiče – to považuji za velice dobré.*“ Výpovědi pocházejí z Případové studie, která byla součástí závěrečné zprávy z projektu za rok 2007.

Ze zhodnocení cyklu pilotních projektů vyplývají následující doporučení:

- pokračovat v tvorbě externích evaluačních nástrojů s garantovanou kvalitou, tj. nástrojů, které reprezentují nároky dané pedagogickými dokumenty, mají dobré psychometrické charakteristiky, umožňují autoevaluaci žáků a škol, zaručující srovnatelnost testových výsledků atd.
- podporovat komplexní pedagogické hodnocení, jehož jsou testy pouze jednou z částí
- zachovat dobrovolnost účasti škol na testování, ale podpořit je v zapojení do testování
- pokračovat ve využívání elektronických způsobů testování (např. počítačem adaptované testy – CAT) a hodnocení, které snižují finanční a časové nároky
- orientovat testování na tzv. přidanou hodnotu, tj. ověřovat znalosti a dovednosti v uzlových bodech vzdělávací dráhy žáků

Možnosti využití zkušeností z cyklu pilotních projektů

Novela školského zákona uvolnila limit počtu přihlášek uchazečů ke studiu na střední školu, z čehož vyplývá, že na místo jedné přihlášky mohou žáci podávat přihlášku až na tři střední školy.

Novelou tohoto zákona, která je bezesporu přínosem pro žáky, vzniká však poměrně riziková a komplikovaná situace pro střední školy, spojená zejména s extrémním zvýšením administrativní zátěže, zvýšením časových nákladů na realizaci přijímacích zkoušek související s nutností vypsání více termínů konání a v konečném důsledku celkovým zvýšením nákladů nutných vynaložit na konání přijímacích zkoušek.

Motivací vedení Centra pro zjišťování výsledků vzdělávání je snaha snížit negativní dopady novely školského zákona do agendy SŠ související s organizací přijímacího řízení.

Navržené řešení, které vyplývá ze studie zpracované v srpnu 2008, využívá synergického efektu vyplývajícího z provázání s paralelně probíhajícími projekty, které CZVV realizuje, a to zejména s právě končícím cyklem pilotních projektů Národní hodnotící zkoušky a projektem reformy maturitních zkoušek.

Vytvoření jednotných standardizovaných evaluačních nástrojů pro vyhodnocení přijímacího řízení na střední školy a sjednocení koncepce organizace je tedy jednou z cest využití zkušeností z cyklu pilotních projektů Národní hodnotící zkoušky.