

Apendix: Doporučené zdroje

Téma hodnocení a datové politiky ve vzdělávání má šířku i hloubku, které nelze pokrýt v jednom přehledovém textu. Existuje řada knih, odborných článků a jiných dokumentů, které se různým prvkům této politiky věnují. Některé představují široké úvodní přehledy a doporučení pro veřejnou politiku, jiné hluboké koncepční či technické materiály. V následujícím seznamu lze najít výběr relevantních materiálů pro různé tematické okruhy.

Co číst	Náročnost	Popis
Přehledové texty o otázkách testování ve vzdělávání		
Black, P. (1997). <i>Testing: Friend or Foe?: Theory and Practice of Assessment and Testing</i> . RoutledgeFalmer, Taylor & Francis Group.	Základní	Přístupný úvod do teorie a praxe hodnocení, psaný bez zbytečného odborného žargonu. Věnuje se klíčovým tématům jako jsou spolehlivost, validita, ověřovací a srovnávací testy. Vhodný jako první četba pro každého, kdo se s problematikou teprve seznamuje. Každá kapitola nabízí diskusní otázky a tipy na další literaturu.
Stobart, G. (2008). <i>Testing times: The uses and abuses of assessment</i> . Routledge.	Základní	Aktuálnější kritický přehled toho, jak hodnocení formuje učení i sebepojetí žáků. Stobart rozebírá pět velkých témat: testování inteligence, různé styly učení, akontabilitu, „diplomovou nemoc“ a formativní hodnocení. Kniha ukazuje, proč dobré úmysly hodnocení nezřídka vedou k učení „na test“. Dobrý výchozí bod pro diskusi o smyslu a dopadech testování.
Isaacs, T., Zara, C., Herbert, G., Coombs, S. J., & Smith, C. (2013). <i>Key concepts in educational assessment</i> . SAGE Publications Ltd.	Základní	Odborný slovník klíčových pojmů a konceptů hodnocení vzdělávání – od validity přes reliabilitu po standardizaci a zpětnou vazbu. Slouží jako rychlá referenční

		příručka při práci s odbornými texty nebo při tvorbě datové politiky. Každé heslo je stručně, ale přesně vysvětleno s ohledem na praktické i teoretické aspekty. Vhodné mít po ruce jako terminologický průvodce.
Mislevy, R. J. (2018). <i>Sociocognitive Foundations of Educational Measurement</i> . Routledge.	Pokročilá	Přelomová syntéza psychometrie, kognitivní psychologie a filosofie měření v jednom svazku. Mislevy propojuje klasickou teorii testů, IRT modely a bayesovské sítě s moderním sociokognitivním pohledem na to, co testování skutečně měří. Klíčová četba pro ty, kdo chtějí rozumět budoucnosti vzdělávacího hodnocení v éře digitálních a adaptivních testů. Vyžaduje solidní metodologické základy.
Národní a nadnárodní organizace		
AERA, APA, & NCME. (2014). <i>Standards for educational and psychological testing</i> . American Educational Research Association.	Střední	Závazná referenční norma pro vývoj, hodnocení a použití vzdělávacích a psychologických testů vydaná třemi předními americkými odbornými asociacemi. Pokrývá celý životní cyklus testu: od konstrukce a validace přes skórování až po interpretaci výsledků a etické otázky. V českém kontextu slouží jako mezinárodní srovnávací standard a východisko pro tvorbu vlastních metodických pravidel. Čtěte selektivně podle aktuálního tématu.
Santiago, P., Gilmore, A., Nusche, D., & Sammons, P. (2012). <i>OECD Reviews of Evaluation and Assessment in Education: Czech Republic 2012</i> . OECD.	Základní	Systematická diagnóza stavu hodnocení ve vzdělávání v České republice zpracovaná týmem OECD. Dokument identifikuje slabiny tehdejšího systému – fragmentaci odpovědnosti, absenci zpětné vazby na úrovni škol i nedostatečná data

		pro řízení vzdělávání – a nabízí konkrétní doporučení. Ačkoliv je starší, většina identifikovaných problémů zůstává aktuální. Dobrý výchozí bod pro pochopení kontextu české vzdělávací politiky.
Kellaghan, T., Greaney, V., & Murray, T. S. (2009). <i>Using the Results of a National Assessment of Educational Achievement</i> (Vol. 5). World Bank.	Základní	Pátý svazek série Světové banky o národních hodnoceních se věnuje otázce, co s výsledky dělat poté, co je sběr dat hotový. Pokrývá celý řetězec práce s informacemi o výsledcích od interpretace dat přes komunikaci různým skupinám (vláda, školy, média, veřejnost) až po institucionální podmínky, které umožňují nebo blokují skutečné využití dat v politice. Klíčové zjištění je, že nevyužití výsledků hodnocení je téměř vždy problém organizační a politický, ne technický. Nejrelevantnější pro úředníky zodpovědné za to, jak jsou výsledky CERMAT nebo ČŠI komunikovány a co se s nimi dál děje.
Raudonytė, I. (2019). <i>Use of Learning Assessment Data in Education Policy-Making</i> . IIEP-UNESCO Working Paper.	Základní	Přehledový working paper UNESCO/IIEP mapující, co literatura říká o tom, jak (a zda vůbec) data z hodnocení ovlivňují tvorbu vzdělávací politiky. Ústřední zjištění: data nejčastěji ovlivňují nastavení agendy a monitoring, ale zřídka samotnou formulaci konkrétních politik. Autorstvo identifikuje typické bariéry jako nedostatečné kapacity k analýze, slabou koordinaci mezi agenturami a chybějící zpětnovazební smyčky mezi úrovněmi systému. Text je stručný a přístupný.
Newton, P., Baird, J., Goldstein, H., Patrick, H., & Tymms, P. (2008).	Pokročilá	Rozsáhlá kolekce výzkumných studií zpracovaná pro britský regulátor kvalifikací (QCA), která

<p>Techniques for monitoring the comparability of examination standards - Publications - GOV.UK. Gov.uk.</p>		<p>mapuje metody zajišťování srovnatelnosti výsledků zkoušek v čase i mezi různými poskytovateli. Srovnatelnost je přitom klíčovým, ale v ČR dosud systematicky neřešeným problémem – bez ní nelze výsledky testů zodpovědně interpretovat. Technicky náročné, ale bezprostředně relevantní pro tvorbu národní testovací politiky.</p>
<p>International Test Commission (2014). ITC Guidelines on Quality Control in Scoring, Test Analysis, and Reporting of Test Scores. <i>International Journal of Testing</i>, 14(3), 195–217.</p>	<p>Střední</p>	<p>Mezinárodně uznávaná doporučení pro zajišťování kvality při skórování testů, analýze výsledků a reportování. Pokrývá celý řetězec od sběru dat přes jejich zpracování až po komunikaci výsledků veřejnosti a tvůrcům politik. Pro datové politiky ve vzdělávání jde o praktický normativní rámec, na který lze odkazovat při návrhu procesů a kontrolních mechanismů. Text je přehledně strukturovaný a relativně stručný.</p>
<p>EMIS</p>		
<p>van Wyk, C., & Crouch, L. (2020). Efficiency and Effectiveness in Choosing and Using an EMIS: Guidelines for Data Management and Functionality in Education Management Information Systems (EMIS). UNESCO Institute for Statistics.</p>	<p>Střední</p>	<p>Praktický průvodce výběrem a provozem vzdělávacích informačních systémů (EMIS) vydaný UNESCO. Zaměřuje se na to, jaká data systémy potřebují sbírat, jak je spravovat a jak je efektivně využívat pro rozhodování. Relevantní pro každého, kdo se podílí na návrhu nebo reformě datové infrastruktury ve školství. Doporučujeme hledat části o datové architektuře, procesech validace dat a vazbě na vzdělávací politiku.</p>
<p>OECD (2023), OECD Digital Education Outlook 2023: Towards an Effective</p>	<p>Základní</p>	<p>Mezinárodní srovnání digitálních ekosystémů ve vzdělávání napříč zeměmi OECD. Přehledně ukazuje, jak různé státy budují datové a technologické infrastruktury pro</p>

<p><i>Digital Education Ecosystem</i>, OECD Publishing, Paris.</p>		<p>vzdělávání, a identifikuje oblasti, kde Česká republika zaostává za mezinárodním standardem. Užitečné jako kontext pro argumentaci ve prospěch investic do digitalizace vzdělávání. Zpráva je bohatě vizualizovaná a čtivá i pro nespecializované publikum.</p>
<p>Abdul-Hamid, H. (2014). <i>What Matters Most for Education Management Information Systems: A Framework Paper</i>. World Bank Group.</p>	<p>Základní</p>	<p>Analytický rámec Světové banky pro budování efektivních vzdělávacích informačních systémů. Definuje klíčové funkce EMIS, popisuje typická selhání systémů a nabízí doporučení pro jejich návrh a implementaci. Vhodné jako vstupní literatura pro každého, kdo začíná přemýšlet o architektuře datových systémů ve vzdělávání. Srozumitelné i bez technického zázemí.</p>
<p>Poisson, M., & Jorgoni, E. (2025). <i>Open School Data in European Education Systems</i>. Council of Europe Publishing.</p>	<p>Základní</p>	<p>Nejnovější srovnávací přehled praxe otevřených školních dat ve členských státech Rady Evropy, zpracovaný společně s IIEP-UNESCO. Mapuje, která data státy zveřejňují, jak jsou řízeny datové platformy a jaký mají vliv na transparentnost a odpovědnost vzdělávacích systémů. Identifikuje typické bariéry – nejednotné standardy, zastaralou infrastrukturu a obavy ze soukromí – i inovativní přístupy. Příímý srovnávací kontext pro českou diskusi o otevřených datech ve školství.</p>
<p>Teoretická a kritická literatura</p>		
<p>Au, W. (2022). <i>Unequal By Design: High-Stakes Testing and the Standardization of Inequality</i> (2. vyd.). Taylor & Francis.</p>	<p>Základní</p>	<p>Kritická analýza standardizovaného testování jako nástroje reprodukce sociálních a ekonomických nerovností. Au sleduje historické kořeny testování v eugenice a hnutí</p>

		za vědecké řízení školy a ukazuje, jak dnešní testovací praxe posiluje třídní a rasové nerovnosti. Americký kontext, ale argumenty přenositelné i do českého prostředí.
Baird, J.-A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? <i>Assessment in Education: Principles, Policy & Practice</i> , 24(3), 317–350.	Pokročilá	Klíčová studie od předních autorů oboru, která rozebírá hlubokou propast mezi psychometrickou tradicí testování a teorií vzdělávání. Autoři argumentují, že statistické nástroje psychometrie vycházejí z předpokladů, které jsou v zásadním rozporu s tím, co víme o tom, jak se lidé učí. Relevantní pro každého, kdo navrhuje systémy hodnocení a chce pochopit jejich skryté limity. Klíčová četba pro kritické myšlení o testování.
Biesta, G. J. J. (2016). Good education in an age of measurement: Ethics, politics, democracy. Routledge, Taylor & Francis Group.	Střední	Filosofická kritika toho, jak posedlost měřením vytlačuje otázku po smyslu a účelu vzdělávání. Biesta ukazuje, že důraz na měřitelné výstupy neodpovídá tomu, co od vzdělávání skutečně chceme, a navrhuje tři vzájemně provázané cíle: kvalifikaci, socializaci a subjektivaci. Důležitá četba pro každého, kdo tvoří vzdělávací politiku a nechce redukovat kvalitu vzdělávání na čísla. Čtivé i bez filosofického zázemí.
Markus, K. A., & Borsboom, D. (2024). Frontiers of test validity theory: Measurement, causation, and meaning (2. vyd.). Routledge.	Pokročilá	Nejnovější a nejkomplexnější přehled teorie validity – tj. otázky, co vlastně testy měří a zda to měří správně. Borsboom patří k nejvlivnějším současným autorům v oblasti filosofie měření a psychometrie. Kniha propojuje technické, kauzální a sémantické aspekty validity do ucelené teorie. Nezbytná četba pro odborníky, kteří chtějí do hloubky rozumět tomu, co testové výsledky skutečně říkají

		a co ne.
Newton, P., & Shaw, S. (2014). <i>Validity in educational & psychological assessment</i> . SAGE Publications Ltd.	Pokročilá	Systematický přehled konceptu validity od klíčových autorů oboru, sledující vývoj teorie od jednoduchých definic po moderní argumentační přístupy. Kniha ukazuje, proč validita není jen technická vlastnost testu, ale komplexní argumentační záležitost zahrnující účel, kontext a důsledky měření. Vhodné pro odborníky, kteří chtějí porozumět, proč se validita nedá redukovat na korelační koeficienty. Náročnější četba, ale přehledně strukturovaná.
Downing, S. M., & Haladyna, T. M. (Ed.). (2006). <i>Handbook of test development</i> . L. Erlbaum.	Pokročilá	Rozsáhlá odborná příručka pokrývající celý proces tvorby testů – od konceptuálního rámce přes tvorbu položek, pilotování a analýzu až po standardizaci a reportování výsledků. Prakticky orientovaný manuál, na který se v oboru odkazuje jako na jeden ze základních referenčních zdrojů. Ideální pro specialisty zapojené do přímé tvorby nebo evaluace testovacích nástrojů. Čtěte selektivně podle potřeby.
Reckase, M. (2023). <i>The Psychometrics of Standard Setting: Connecting Policy and Test Scores</i> (1st ed.). Chapman and Hall/CRC.	Pokročilá	Specializovaná monografie věnovaná statistické teorii a metodám stanovování hranic výkonu (cut scores) pro interpretaci testových výsledků. Standard setting – tedy rozhodnutí, od jakého skóre je žák „dostatečně dobrý“ – je klíčovým a politicky citlivým krokem, který v ČR dosud postrádá systematickou metodologii. Kniha spojuje technické aspekty s normativními politickými rozhodnutími. Nezbytná četba pro odborníky zodpovědné za nastavení standardů národního testování.

