

JAK MĚŘIT METAKOGNICI (NEJEN) U NADANÝCH DĚTÍ

Ondřej Straka

MASARYKOVA
UNIVERZITA

JAK MĚŘIT METAKOGNICI (NEJEN) U NADANÝCH DĚTÍ

Ondřej Straka

Masarykova univerzita
Brno 2021

Publikace vznikla v rámci projektu GA17-14715S: Vývoj metakognitivních schopností u nadaných dětí, podpořeného Grantovou agenturou České republiky

Knihu recenzovali:

Mgr. Eva Klimecká, Ph.D.; Centrum výzkumu FHS, Univerzita Tomáše Bati ve Zlíně

doc. Mgr. Jiří Mudrák, Ph.D.; Psychologický ústav AV ČR



Tato práce je licencována pod licencí

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

© 2021 Masarykova univerzita

ISBN 978-80-210-9905-0

<https://doi.org/10.5817/CZ.MUNI.M210-9905-2021>

OBSAH

■ ■ ■	Úvod 5
■ ■ ■	Vymezení metakognice 9
■ ■ ■	Metakognice a mimořádné nadání 12
■ ■ ■	Základní rozdělení metod pro zjišťování metakognice 16
■ ■ ■	On-line metody 17
	Analýza TAP 17
	Analýza postupů registrovaných pomocí počítače 25
	Sledování očních pohybů (eye-tracking) 30
■ ■ ■	Off-line metody 34
	Dotazníky a posuzovací škály 34
	Metakognitivní interview 38
	Metoda stimulovaného vybavení 39
■ ■ ■	Specifika měření metakognitivního monitorování obtížnosti úlohy 41
■ ■ ■	Míry absolutní přesnosti 50
	Index absolutní přesnosti 51
	Index kalibrace s rozdělením posuzovací škály do kategorií 52
	Bias index 53
	Indexy používané při práci s kategoriickými druhy metakognitivních odhadů 54
	Prostý koeficient shody (simple matching coefficient) 56
	Hamannův koeficient 56
■ ■ ■	Míry relativní přesnosti 58
	Pearsonův korelační koeficient 58
	Index rozlišení 60
	Koeficient gamma 61
	Porovnání Hamannova koeficientu a koeficientu gamma 64
■ ■ ■	Metriky založené na teorii detekce signálu 70
	AUROC2 80
	Meta-d' 81
■ ■ ■	Závěrem 84
■ ■ ■	Summary 87
■ ■ ■	Literatura 89



ÚVOD

Metakognice – tedy schopnost jedince reflektovat svoje vlastní poznávací procesy, případně i poznávací procesy jiných lidí, a využívat tyto reflexe při učení a řešení problémů, je pojem, který se za poslední zhruba čtyři desítky let pevně etabloval v rámci pedagogiky, psychologie a dalších příbuzných oborů. V zahraniční, zejména v anglicky psané odborné literatuře, vychází obrovské množství původních teoretických i výzkumných studií (např. rešerše provedená s pomocí plnotextové databáze EBSCO odhalí jen za posledních pět let více než 5000 publikovaných statí obsahujících v názvu klíčové slovo „metacognition“ či „metacognitive“). Publikovány jsou obsáhlé sborníky, v nichž je téma metakognice zpracováno z mnoha různých pohledů – jako několik příkladů z mnoha uvedme projevy metakognitivních schopností v různých školních předmětech či studijních oborech (Zohar & Dori, 2011; Dori et al., 2018), vývoj metakognitivních schopností a jejich specifika v jednotlivých životních etapách (Metcalf & Dunlosky, 2009), možnost rozvíjet metakognici žáků prostřednictvím cílených programů a efekt takovýchto intervencí na školní výsledky a na prevenci studijní neúspěšnosti (Waters & Schneider, 2010; Azevedo & Aleven, 2013; Peña-Ayala & Cárdenas, 2015) a v neposlední řadě i vztah metakognice a mimořádného intelektového nadání (Shaughnessy et al., 2008). Existuje dokonce impaktovaný časopis, který se specializuje výhradně na publikování prací týkajících se tématu metakognice (jedná se o časopis *Metacognition and Learning* vydávaný od roku 2006).

V české odborné literatuře je téma metakognice zastoupeno o poznání skromněji. V uplynulých zhruba 10–12 letech bylo publikováno několik teoretických či přehledových studií a několik původních výzkumů. Lokajíčková (2014) podává širší výklad celého konceptu, představuje některé taxonomické systémy a zamýšlí se nad úlohou metakognice ve výuce. Tématu metakognice se na různých místech věnuje také obsáhlá učebnice pedagogické psychologie, jejímž autorem je Mareš (2013a). Přehledová studie Straky a Portešové (2019) se zabývá problematikou metakognice ve vztahu ke specifické populaci mimořádně nadaných dětí. Vztahem metakognice a mimořádného nadání se zabývá také výzkumná práce Málkové a Jirotkové (2019). Další empirické studie publikované v posledních letech se zaměřují např. na možnosti rozvoje metakognitivních schopností v předškolním vzdělávání (Hrbáčková, 2011), na intervence usilující o rozvoj metakognice u žáků základních škol (Foltýnová, 2009), na vztah mezi úrovní vědeckého myšlení a metakognicí u studentů učitelství

(Nováková et al., 2018) či na vztah alternativních výukových metod v matematice a metakognice u takto vzdělávaných žáků (Chytrý et al., 2019).¹

Je tedy vidět, že téma u nás sice není úplně opomíjené, nicméně – přinejmenším v kvantitativním ohledu – je odstup za vývojem ve světě zjevný. Tato situace je nepříznivá především s ohledem na omezené možnosti využívat poznatků o metakognici v pedagogické a psychologické praxi. Pokud jde o teoretická zjištění týkající se metakognice, je na místě se domnívat, že můžeme s určitou opatrností čerpat z výsledků zahraničních autorů, přestože i v této oblasti se samozřejmě může projevit vliv kulturních rozdílů. Z toho, jak je metakognice uchopena v pracích Flavella (1979), Brownové (1987) a jejich následovníků, ji však lze chápat jako obecně lidskou dispozici a její charakteristiky, zkoumané a popsané v anglosaském prostředí, je možné s rozumnou obezřetností zobecnit na všechny lidi, nebo přinejmenším na lidi žijící v „industriálních“ kulturách. Zabýváme-li se ovšem *aplikací* těchto poznatků², vstupuje do hry organizace a struktura školského systému v té které zemi, rozsah, uspořádání a závaznost kurikula a mnoho dalších rozdílů, které mechanické přejímání zahraničních výsledků komplikují. Bylo by proto nepochybně prospěšné, kdyby se původních výzkumných studií zabývajících se metakognicí, které by zároveň zohledňovaly lokální specifika našeho školství, realizovalo podstatně víc.

Analogicky by velký přínos představovala rovněž mnohem intenzivnější práce s metakognicí v rámci individuálního pedagogicko-psychologického poradenství. Obrací-li se dnes na školní či poradenské psychology a speciální pedagogy rodiče nebo učitelé žáků, kteří ve výuce selhávají, dokáže aktuálně poradenský systém v České republice standardně a ve většině případů snad i kvalitně diagnostikovat a zohlednit takové příčiny, jakými jsou snížený intelekt, poruchy učení, případně i různé osobnostní, rodinné a sociální faktory. Jak však ukazuje zahraniční zkušenost, existují i žáci, jejichž rozumové schopnosti jsou průměrné nebo i nadprůměrné, kteří nejsou znevýhodněni žádným z právě uvedených faktorů a kteří přesto ve výuce selhávají, protože svoje rozumové schopnosti nedokáží adekvátně používat (McGuire, 2015). Individuální práce zaměřená na rozvoj a trénování metakognitivních dovedností může takovým žákům výrazně pomoci a zejména u starších žáků a studentů může někdy představovat poslední šanci, jak vůbec pokračovat ve studiu. Aby takovou pomoc mohli poradenská pracovníci efektivně poskytovat, potřebovali by také mít k dispozici ověřené nástroje a postupy pro posouzení úrovně metakognitivních schopností toho kterého žáka. Takové nástroje ovšem – až na nepočtené výjimky – v českém jazyce a standardizované pro českou žákovskou populaci v současnosti dostupné nejsou.

¹ Zmínit můžeme i některé starší práce, které se v českém kontextu tématu metakognice věnovaly (např. Hrbáčková, 2004; Krykorková & Chvál, 2001).

² Např. pokud se snažíme naplánovat a ověřit efektivitu kurzu pro rozvoj metakognitivních schopností u žáků, nebo nás zajímá, do jaké míry jsou studijní neúspěchy žáků v konkrétním předmětu způsobeny na jedné straně nedostatky v jejich metakognitivních schopnostech a na druhé straně dalšími faktory – jakými mohou být nevhodný způsob prezentace učiva, klima školy apod.

Příčin tohoto stavu je nepochybně více. Zde bych se nicméně chtěl blíže zaměřit na jednu z nich. Za tu lze považovat přetrvávající nízké povědomí o *metodách* měření metakognice. V tomto směru nejsou příliš nápomocné ani výše uvedené přehledové studie, které se primárně zaměřují především na samotné vymezení konstruktů, případně na jeho zasazení do širšího pedagogického kontextu, a méně již na otázky diagnostiky a měření. Jinak řečeno, zabývají se tím, co je to metakognice, jak se vyvíjí, jakou má vnitřní strukturu, jak souvisí s dalšími pedagogickými či psychologickými jevy, a spíše opomíjí otázku, jak úroveň metakognitivních schopností zjišťovat a případně i vyčíslit u konkrétního jedince. Toto tvrzení neplatí zcela absolutně, např. výše uvedená práce Lokajíčkové (2014) obsahuje stručný popis několika v zahraničí používaných dotazníků určených pro zjišťování metakognice, některé dotazníkové metody s metakognitivním přesahem (tj. vedle metakognice zároveň měřící i jiné jevy) zmiňuje ve své publikaci také Mareš (2013a). Nováková, Chytrý a Říčan (2018) vysvětlují v úvodu ke své výzkumné studii význam některých indexů používaných ke kvantifikaci metakognitivního monitorování (index absolutní přesnosti, bias a stručně i některé další). Pokud je mi však známo, žádný autor české odborné literatury se zatím nepokusil o souhrnný přehled základních metod používaných pro měření metakognice, které by byly popsány na jednom místě, vzájemně porovnaný a kriticky zhodnoceny. Zaplnění této mezer je hlavním cílem předkládané publikace.

Tato práce byla původně zamýšlena jako časopisecký článek v podobě literárního přehledu (Mareš, 2013b). Téma se však v průběhu přípravy ukázalo natolik obsáhlé a zároveň nosné na to, aby si po mém soudu zasloužilo formát ucelené knižní monografie. Přesto tato práce ani nyní nemá ambici popsat a zhodnotit všechny nástroje, které byly ve světě vytvořeny a publikovány. Taková snaha by narážela na fakt, že konkrétních diagnostických nástrojů vzniklo v posledních dekadách v zahraničí skutečně mnoho – např. přehledová studie zaměřená pouze na dotazníky a podobné metody (sebeuposuzovací škály apod.), kterou realizovala Craigová a kolegové (2020), cituje celkem 36 metod, k nimž existuje dostupná a adekvátně zpracovaná anglicky psaná dokumentace. Dotazníkové metody přitom, jak ještě uvidíme dále, nejsou zdaleka jediným způsobem, jak zjišťovat a měřit metakognitivní schopnosti. Uvést byt i stručný popis každého individuálního nástroje by vyžadovalo bez nadsázky stovky stran textu, což by – alespoň se tak domnívám – z pohledu čtenáře nebylo účelné. Namísto toho se tedy pokusíme popsat základní *kategorie* metod (např. dotazníky, aplikace pro registraci a analýzu záznamů o činnosti účastníka atd.), kriticky zhodnotit jejich silné a slabé stránky a pro ty kategorie, kde to dává smysl, pak jako ilustrační příklad představit několik málo konkrétních diagnostických nástrojů či standardních postupů. Nejprve budou prezentovány metody, které registrují potřebná data souběžně s metakognitivní aktivitou jedince (tzv. on-line metody). Ve druhé části knihy se budeme zabývat metodami, které zjišťují metakognitivní schopnosti jedince s určitým časovým odstupem od úkolů či činností, jež od něj vyžadují metakognitivní aktivitu (tzv. off-line metody). Třetí a nejobsáhlejší část této publikace bude věnována specifikům měření metakognitivního monitorování, s důrazem na zpracování výsledků tohoto měření a na jejich vyjádření prostřednictvím několika různých metrik či indexů.



Tato studie byla vypracována v rámci širšího výzkumného projektu zaměřeného na problematiku metakognice u nadaných dětí. Tomu je přizpůsoben zejména výběr výše zmíněných ilustračních příkladů – zařazeny jsou do něj přednostně nástroje, které buď byly použity v zahraničí publikovaném výzkumu týkajícím se metakognice u nadaných, nebo které můžeme považovat za potenciálně vhodné pro účely praktické diagnostiky. Zároveň je ale nutné zmínit, že – pokud je autorovi známo – neexistuje žádný test či jiný diagnostický nástroj pro měření metakognice, který by byl vyvinut a používán výhradně pro tuto specifickou populaci. Četné empirické studie také ukazují, že rozdíly v metakognitivních schopnostech mezi nadanými dětmi a běžnou populací sice existují, ale nejsou zdaleka tak výrazné, jak by se mohlo z laického pohledu zdát (Alexander et al., 1995; Greene et al., 2008; Veenman, 2015). Z toho vyplývá, že zde prezentovaná klasifikace nástrojů pro měření metakognice a jejich zhodnocení má v zásadě obecnou platnost a může z ní čerpat i čtenář, který se zajímá o metakognici u průměrně nadaných dětí, u dětí s lehčími výukovými deficity, případně i u dospělých osob.

VYMEZENÍ METAKOGNICE

Jak bylo uvedeno výše, podrobný popis toho, co to je metakognice, jaké kategorie metakognitivních jevů rozeznáváme a jaké z toho vyplývají závěry pro vzdělávání, je již dostupný i v české odborné literatuře, a proto se mu zde nebudeme obsáhle věnovat. Bude nicméně užitečné dříve, než se pustíme do samotné prezentace diagnostických a měřících metod, zde tyto informace alespoň stručně zrekapitulovat.

Termín metakognice navrhli na přelomu 70. a 80. let minulého století krátce po sobě J. Flavell a A. Brownová. Oba autoři definovali tento pojem poměrně stručně a velmi obecně. Zatímco Flavell (1979, s. 906) definoval metakognici jako „znalosti a poznávání kognitivních jevů“, pro Brownovou (1987) představuje metakognice „znalost a regulaci vlastního kognitivního systému“. Pokusíme-li se tyto definice nějak parafrázovat, můžeme říci, že metakognice je schopnost jedince uvědomovat si vlastní poznávací schopnosti – myšlení, vnímání, paměť apod. – a využívat výsledky tohoto uvědomění při učení, řešení problémů a při dalších aktivitách, do kterých se tyto poznávací schopnosti zapojují.

Skutečnost, že člověk může „poznávat vlastní poznávání“, byla pojednávána např. v různých filozofických směrech již dávno v minulosti, přínos Flavella a Brownové ovšem spočíval v tom, že k tomuto tématu obrátili pozornost psychologie a pedagogiky a že z něj učinili konstrukt, který je možné měřit a experimentálně studovat. S tím souvisí další podstatný rys současného pojetí metakognice, a sice fakt, že podobně jako v případě „klasických“ poznávacích schopností se jedná o vlastnost, v jejíž míře se lidé navzájem liší a která tedy může být u různých osob rozvinuta na mimořádně vysoké, průměrné či podprůměrné úrovni. Zároveň existují dobré důvody předpokládat, že úroveň metakognice nemusí být plně determinována úrovní rozumových schopností a že konkrétní osoba tedy může být například vysoce inteligentní, a přitom mít zároveň jen průměrnou či dokonce podprůměrnou úroveň metakognitivních schopností (Veenman et al., 2006). Někteří autoři pak dospívají k závěru, že pro predikování školní úspěšnosti je metakognice dokonce důležitějším faktorem, než samotné rozumové schopnosti (Veenman, 2011; Cooper et al., 2008; Helms-Lorenz & Jacobse, 2008).

Důvod, proč metakognice souvisí se školní či obecně studijní úspěšností, je zjevný a můžeme si ho přiblížit na následujícím příkladu. Uvažme žáka, který chybně podceňuje relativní (vzhledem ke svému nadání) náročnost látky, kterou se musí naučit na nadcházející test,

protože má nerealisticky vysokou představu o svých rozumových schopnostech. Takový žák, i když může být motivován získat dobrou známku, nevěnuje přípravě dostatek času a úsilí a v testu pravděpodobně selže. Důležité je ovšem zdůraznit, že selže nikoli proto, že by jeho rozumové či paměťové schopnosti na úkol nestačily, ale proto, že tyto schopnosti nedokázal adekvátně využít. Analogicky žák, jehož poznání vlastní kognice je zkreslené v opačném směru, tedy ten, který svoje schopnosti podceňuje, bude často věnovat studijním úkolům zbytečně mnoho času a více úsilí, než jaké od něj zvládnutí dané látky reálně vyžaduje. To by mohl mezitím zaměřit jinam, např. na další učivo, takže ani tato situace není výukově optimální. Výhodu získá naproti tomu žák, který svoje rozumové schopnosti dokáže relativně k danému úkolu posoudit realisticky a adekvátně se podle toho zachovat. Společně se Schrawem (2001) můžeme shrnout, že *metakognice se liší od kognice tím, že vyžaduje porozumět tomu, jak je určitý úkol proveden, zatímco kognice je to, co je nutné pro prosté provedení úkolu.*

Z výše uvedeného je patrné, že metakognice je pojem značně široký, a proto se pochopitelně objevily pokusy metakognitivní jevy nějak podrobněji klasifikovat. Různých taxonomických systémů bylo navrženo více, zde se omezíme pouze na jeden, který přímo vychází ze základatelských prací Flavella a především Brownové, který je velmi rozšířený, srozumitelný a často se promítl i do konstrukce představovaných měřících nástrojů (Veenman et al., 2006). Podle tohoto pojetí rozlišujeme dvě základní kategorie metakognitivních jevů. První z nich, označovaná jako *metakognitivní znalosti, znalosti kognice* (knowledge of cognition), či *deklarativní formy metakognice*, zahrnuje konkrétní znalosti, které má subjekt o poznávacích schopnostech, zejména vztažených k jeho vlastní duševní aktivitě. Patří sem například povědomí jedince o tom, že když se učí nepřetržitě déle než půl hodiny, jeho paměťový výkon po uplynutí této doby rapidně klesá, nebo přesvědčení o tom, že si při jednorázové expozici zapamatuje více nových slov v cizím jazyce, pokud jsou mu prezentovány v mluvené formě, než kdyby je viděl napsané. Za zdůraznění stojí fakt, že metakognitivní znalosti konkrétního jedince mohou být správné a realistické, ale i chybné (tj. žák si například může myslet, že k poklesu paměťového výkonu u něj dochází až po půl hodině, přestože objektivní test či experiment by ukázaly, že k tomu dochází již po 20 minutách apod.).

Součástí metakognitivních znalostí jsou také určité informace o fungování poznávacích schopností u lidí obecně – např. že snadněji si zapamatujeme kratší seznam nesouvisejících prvků než delší seznam, že opakované učení vede k lepšímu zapamatování, či že myšlení povede ke správným výsledkům pravděpodobněji v klidném prostředí než v prostředí plném rušivých podnětů. Dospělým osobám se takové informace mohou jevit jako naprosto banální, z hlediska vývojové psychologie tomu ale tak není – je jasné, že i takto základní znalosti si dítě může osvojit až v určitém věku a empiricky se ukázalo, že se děti v načasování a ve snadnosti, s jakou si tyto metakognitivní znalosti osvojují, vzájemně poměrně dost liší (Alexander et al., 1995; Schneider & Pressley, 2013). To má pochopitelně významné důsledky pro výuku.

Druhá kategorie metakognitivních jevů se zpravidla označuje jako *procedurální formy metakognice*, případně jako *regulace kognice* (regulation of cognition), a někdy též jako *metakognitivní dovednosti* (metacognitive skills). Patří sem jednak schopnost *metakognitivního monitorování* – tedy schopnost člověka, který musí řešit úkol vyžadující zapojení kognitivních schopností, posoudit, jak pro něj bude tento úkol obtížný a/nebo jaká je pravděpodobnost, že se mu jej podaří správně vyřešit. A dále sem patří regulace a kontrola strategií, které se uplatní při učení či řešení problémů za účelem dosažení nejlepšího možného výsledku. Podstatou těchto strategií je účelně a v optimálním pořadí využívat dostupné poznávací schopnosti v průběhu řešení úkolu. Obě procedurální složky jsou většinou ve své činnosti úzce propojené – např. žák vyhodnotí látku, kterou se má naučit, jako příliš obtížnou na to, aby to dokázal na první pokus (monitorování). Rozdělí si proto text do několika logických celků, které se učí odděleně, a naplánuje si určitý počet a určitou sekvenci opakování (strategie). Jiný příklad může představovat situace, kdy student čte nový, neznámý text a díky metakognitivnímu monitorování si uvědomuje, že některá tvrzení autora a závěry, které z nich vyvozuje, mu nedávají smysl. Uplatní tedy strategii, která spočívá v cílené aktivaci v minulosti osvojených znalostí (prior knowledge), které mu pro tuto novou látku poskytnou potřebný kontext a hlubší souvislosti. Některé další metakognitivní strategie ještě zmíníme později, v části věnované metodě protokolové analýzy.

METAKOGNICE A MIMOŘÁDNÉ NADÁNÍ

Pojetí mimořádného nadání a představa, koho lze vlastně považovat za nadané dítě, byly dlouhou dobu zásadním způsobem ovlivněny zakladatelskými pracemi L. M. Termana, které se datují do 20. let minulého století. Terman (1926) využil tehdy relativně nových testů určených k měření inteligence a zahájil rozsáhlý longitudinální výzkum, jehož cílem bylo vybrat vzorek mimořádně nadaných dětí (ve smyslu výrazně vysokých hodnot IQ) na počátku jejich vzdělávací cesty a dále dlouhodobě sledovat jejich studijní a pracovní úspěšnost. Termanova východiska, která se zdála být zpočátku potvrzována průběžnými výsledky jeho studií, lze shrnout do několika základních bodů: 1. lidské intelektové schopnosti jsou převážně vrozené a lze je jen minimálně ovlivnit v průběhu života jedince; 2. inteligenci je možné chápat jako kompaktní, jednodimenzionální schopnost, jejíž míru lze proto vyjádřit jediným číslem (intelligenčním kvocienem). Odtud zákonitě vyplývá předpoklad, že jedinec by měl mít tendenci být zhruba stejně úspěšný (např. nadprůměrně, průměrně, či podprůměrně) ve všech typech aktivit, do nichž se inteligence promítá – např. v matematice, ve verbálním projevu, v prostorových schopnostech a podobně. Schopnost řešit rozdílné typy problémů tak představuje pouze různé projevy téhož společného faktoru – obecné inteligence; 3. úroveň inteligence u konkrétního člověka je možné přesně a spolehlivě změřit pomocí vhodných psychodiagnostických nástrojů.

V tomto termanovském pojetí bylo tedy dítě považováno za nadané tehdy a právě tehdy, pokud ve standardizovaném intelligenčním testu dosáhlo skóru, který byl rovný nebo vyšší než určitá hraniční hodnota (angl. cut-off score). Ta byla přitom stanovena v podstatě zcela arbitrárně. Sám Terman původně pracoval s kritériem 140 bodů IQ, setkat se ale můžeme i s jinak stanovenými prahy – později se např. jako obvyklé kritérium prosadila spíše hodnota IQ 130 (Portešová, 2011). Toto pojetí mělo určitou vnitřní přitažlivost, která vyplývala z jeho jednoduchosti: kritérium bylo jednoznačné a srozumitelné, v každém konkrétním případě bylo díky tomu možné rozhodnout, zda se o nadané dítě jedná, či nikoli. Těžištěm odborného zájmu pak bylo jednak zkoumání určitých osobnostních či temperamentových rozdílů mezi nadanými dětmi a běžnou populací, především byl ale důraz kladen na možnosti plošného vyhledávání nadaných žáků. Cílem této identifikace bylo primárně poskytnout nadaným

žákům adekvátní (tj. odpovídající jejich schopnostem, nikoli nutně jejich chronologickému věku) vzdělávací příležitosti – například prostřednictvím obohacování učiva, přeskočení ročníku či jiné formy akcelerace. Naopak vnitřní struktura kognitivních schopností u nadaných dětí byla jako odborné téma spíše přehlížena, což bylo s ohledem na předpoklad jednodimenzionality inteligence pochopitelné.

Důvodů, proč byla termanovská tradice postupně překonána a dnešní pohled na nadání je mnohem komplexnější, bylo více. Jako velmi důležité můžeme zmínit dva z nich. Za prvé, celkový pohled na inteligenci (tj. nejen u dětí, a nejen ve vztahu k mimořádnému nadání) v rámci psychologie se v posledních několika dekádách proměnil s nástupem teorií a modelů zdůrazňujících vícedimenzionalitu rozumových schopností. Například CHC teorie – zřejmě nejdůležitější a nejvlivnější z těchto přístupů – sice stále pracuje s faktem, že jednotlivé dílčí kognitivní schopnosti (jako jsou verbální, prostorové, matematicko-logické a další) spolu vzájemně korelují a vytvářejí tedy jeden společný g-faktor, tyto korelace jsou nicméně méně významné, než se dříve soudilo (McGrew, 2009). Je proto opodstatněné posuzovat každou z těchto schopností i samostatně, optimálním výsledkem testování inteligence by tedy již nemělo být jediné číslo, ale celý *profil* hodnot dílčích kognitivních schopností. Proces identifikace nadání se tím poněkud komplikuje. Z určitého pohledu dává například smysl považovat za nadaného takového člověka, který má na zcela mimořádné úrovni rozvinutou jednu dílčí schopnost – např. prostorovou představivost, a dosahuje díky tomu v konkrétních činnostech i vynikajících výsledků (Gridley et al., 2003). Ve všech ostatních testovaných kognitivních schopnostech mohou být přitom jeho výsledky pouze mírně nadprůměrné či průměrné, a ani jeho celkové IQ proto nemusí být nijak mimořádně vysoké. Identifikační kritéria pro nadání při práci s celými profily schopností musí být zákonitě složitější, než když se inteligence chápala jako jednodílná schopnost.

Za druhé se zároveň zproblematizoval i předpoklad, že extrémně vysoká hodnota celkového IQ sama o sobě stačí jako kritérium mimořádného nadání (Portešová, 2011; Missett & McCormick, 2014). Jak ukázaly některé pozdější analýzy vycházející ze samotného Termanova longitudinálního výzkumu, které s velkým časovým odstupem realizovali jeho následovníci, řada jedinců, identifikovaných v první fázi studie jako nadané děti, později v dospělosti nedosáhla nijak mimořádných či pozoruhodných výkonů. Různí autoři se pokoušeli vyřešit tento rozpor rozpracováním nových teoretických přístupů, které v konceptualizaci nadání zohledňovaly i jiné, mimointelektové faktory (pro detailnější přehled viz Pfeiffer, 2013; v české odborné literatuře např. Havigerová, 2011). Zřejmě neúspěšnější byl v tomto ohledu J. Renzulli (2005), jehož známý tříkruhový model zahrnuje kromě samotných intelektových předpokladů také tvořivost a motivaci. Lze tedy shrnout, že v současném odborném pohledu na (intelektové) nadání hraje i nadále zásadní roli mimořádně vysoká úroveň jedné či několika kognitivních schopností. Vypůjčme-li si ovšem frázi používanou matematiky, jedná se o podmínku nutnou, nikoli ale postačující.

Výše nastíněné zjednodušené a dnes již překonané termanovské pojetí nadání bohužel stále do určité míry přetrvává v povědomí laické, a částečně i pedagogické veřejnosti – možná právě pro jeho jednoznačnost a intuitivní srozumitelnost. Tím se zároveň udržuje při životě i velmi odolný mýtus, že totiž dítě s mimořádným nadáním musí jaksí z definice excelovat ve všech oblastech, že jeho učení nemůže být jiné než maximálně efektivní a že takové dítě v principu nemůže zažívat výukové obtíže (Moon, 2009). Odtud je již krůček k myšlence, že i metakognitivní schopnosti dítěte s nadprůměrným intelektem musejí být rovněž rozvinuty na nadprůměrné úrovni, jinak řečeno, že intelekt plně determinuje a predikuje metakognici (Elshout & Veenman, 1992)³. Vztah metakognice a nadání ovšem zdaleka není tak přímočarý, jak by se mohlo na první pohled zdát. Existuje totiž i konkurenční model, který tvrdí, že metakognice a intelekt jsou dvě samostatné, na sobě zcela nezávislé schopnosti (Allon et al., 1994). Četné empirické studie ovšem nasvědčují tomu, že nejvíce se blíží realitě třetí, v určitém smyslu kompromisní model, který se v literatuře obvykle označuje jako tzv. *smíšený model* (mixed model; Veenman, 2008). Podle tohoto pojetí spolu sice inteligence a metakognice do určité míry souvisí, jejich korelace je ovšem relativně slabá, a není proto vyloučené najít žáky, kteří jsou vysoce inteligentní a úroveň jejich metakognitivních schopností je přitom nápadně nižší (přestože takových dětí bude relativně méně než těch, u kterých jsou obě schopnosti rozvinuty zhruba na stejné úrovni).

Přijmeme-li při výzkumu a pedagogické práci s nadanými dětmi smíšený model jako východisko, vyplývají z toho minimálně dva důležité závěry. Za prvé, při mnoha výzkumech týkajících se mimořádného nadání by bylo na místě zahrnovat do analýz rovněž úroveň metakognitivních schopností jako jednu z proměnných, protože pokud tento faktor nezohledníme, můžeme získat dost zkreslený obraz o fungování takového dítěte ve školním prostředí i mimo ně. Např. jedním z dlouhodobých a dosud uspokojivě nezodpovězených výzkumných problémů v této oblasti je již výše nastíněná otázka, proč určitá část dětí, které byly v dětství spolehlivě a často i opakovaně identifikovány jako mimořádně nadané, později v dospělosti nedosáhne ničeho mimořádného a ve svém profesním uplatnění se zásadně neliší od svých někdejších průměrných spolužáků. Jednou z pravděpodobných hypotéz – jejíž ověření by ovšem vyžadovalo dlouhodobé a náročné longitudinální studie – je ta, podle níž rozhodujícím faktorem je právě úroveň metakognitivních schopností. Pouze ti nadaní žáci a studenti, kteří mají zároveň vysoce rozvinutou metakognici, dokáží svoje nadání využít k tomu, aby se v některé oblasti stali experty (Carr & Taasoobshirazi, 2008; Veenman, 2008).

³ Veenman et al. (2005) dále tvrdí, že s tímto pojetím implicitně souzní i vlivná teorie inteligence, kterou zformuloval a zastával Robert Sternberg (2005). Sternbergův model pracuje s několika dílčími komponentami, jejichž společná a koordinovaná činnost je podle něj podstatou lidského intelektu. Jednou z těchto složek jsou tzv. *metakomponenty*, což je v podstatě pouze jiný název pro metakognitivní schopnosti.

S tím úzce souvisí druhý závěr vyplývající ze smíšeného modelu – a sice že metakognici je potřebné zohlednit i při individuálním psychologickém poradenství a pedagogické práci s nadanými žáky a studenty. Máme-li například žáka, jehož mimořádné intelektové schopnosti byly prokázány standardizovaným a spolehlivým inteligenčním testem, a tento žák přitom nedosahuje ve srovnání se svými spolužáky o nic lepších studijních výsledků, případně se potýká přímo se studijním neúspěchem, je namístě se ptát, čím je tento stav zapříčiněn. Pochopitelně se může jednat o nedostatek motivace, o souběh nadání s různými smyslovými či kognitivními hendikepy či o nepříznivé a nepodnětné prostředí, v němž žák vyrůstá. Vyloučíme-li ovšem všechny tyto příčiny a žakovy výsledky stále zaostávají za jeho potenciálem, zbývají jako alternativní vysvětlení problémy v oblasti jeho metakognice. K jejich odhalení ovšem potřebují psychologové a další poradenští pracovníci patřičné diagnostické nástroje a odpovídající znalosti pro jejich používání a interpretaci.

ZÁKLADNÍ ROZDĚLENÍ METOD PRO ZJIŠŤOVÁNÍ METAKOGNICE

Způsobů, jak klasifikovat metody pro zjišťování metakognice existuje samozřejmě více, v této publikaci budeme vycházet z jednoduchého a velmi rozšířeného schématu, které používá např. Veenman (2011). Základem tohoto dělení je odpověď na otázku, zda daná metoda registruje projevy metakognitivní aktivity přímo v jejich průběhu (tj. zpravidla v době, kdy posuzovaná osoba řeší nějaký kognitivní úkol, který u ní vyvolá aktivaci metakognitivního monitorování a/nebo metakognitivních strategií), nebo zda je registruje s určitým časovým odstupem. Postupy, při kterých se projevy metakognitivní aktivity registrují bezprostředně, se označují jako tzv. *on-line metody*. Pokud jsou naopak informace o metakognici získávány jindy než přímo v okamžiku, kdy jedinec řeší zadaný úkol, hovoříme o tzv. *off-line metodách*. Pro off-line metody je – s určitými výjimkami – typické, že jde ze strany diagnostikované osoby o určitou formu sebeposouzení (angl. self-report). To může mít buď písemnou (dotazníky, sebeposuzovací škály apod.), nebo ústní (interview) podobu. Dále je u off-line metod potřebné rozlišovat, kdy se – relativně vzhledem k podnětovému úkolu – posuzované osoby dotazujeme. Zejména v případě jednorázových experimentálních designů tak mohou být informace zjišťovány jednak prediktivně (např. student ještě před samotným vyplňováním znalostního testu či písemné práce odhaduje, jak bude v tomto testu úspěšný, případně uvádí, jaké metakognitivní strategie bude při řešení používat). Alternativně – a to je zřejmě častější případ – mohou být zjišťovány postdiktivně (posuzovaná osoba odpovídá až po dokončení úkolu). V případě standardizovaných dotazníků je obvyklé ptát se posuzovaných osob na obecné tendence k používání určitých metakognitivních strategií bez vazby na konkrétní úkol, i když i z tohoto pravidla existují některé výjimky.⁴

⁴ Viz např. dotazník MAESTRA, který ještě krátce zmíníme dále v textu.

ON-LINE METODY

Analýza TAP

Zcela zásadní postavení si ve výzkumu metakognice získala metoda, která se v anglicky psané literatuře označuje jako *think-aloud protocol analysis* (či TAP analysis). Označení představuje určitý problém z hlediska české terminologie, protože doslovný překlad, tedy analýza protokolů myšlení nahlas, zní značně krkolomně, a proto se s ním také setkáváme pouze zřídka. V dalším textu budeme proto používat zkrácené označení *protokolová analýza*, případně *analýza TAP*. Vzhledem k významu této metody se jí budeme věnovat ve srovnání s dalšími metodami poněkud podrobněji.

Na úvod je třeba zdůraznit, že analýza TAP není metodou, která by se používala výhradně pro zjišťování metakognice, a ani za tímto účelem nebyla původně vyvinuta. Vznik metody souvisí s renesancí zájmu o studium lidských poznávacích schopností, které nelze spolehlivě studovat prostřednictvím pozorování zjevného chování, a s etablováním kognitivní psychologie jako svébytného a vlivného psychologického směru. Badatelé hlásící se k tomuto paradigmatu chtěli studovat poznávací procesy, které byly nepřístupné metodám klasického behaviorismu. Zároveň se ovšem chtěli vyvarovat spoléhání se na vlastní introspekci jakožto na zdroj informací o studovaných poznávacích procesech vzhledem k jejich subjektivnosti, nespolehlivosti a nemožnosti je replikovat (Hunt, 2015).

Jedním z východisek měla být právě metoda analýzy TAP, která tak byla zpočátku používána pro výzkum čistě kognitivních procesů (např. jak přemýšlejí při hře amatérští či profesionální šachisté, jaké postupy volí lidé při řešení matematických příkladů či různých logických problémů apod.; Ericsson & Simon, 1993). Později se ovšem ukázalo, že metoda je velmi vhodná i pro detekci metakognitivní aktivity v průběhu řešení podobných úloh (tedy např. k zachycení skutečnosti, že subjekt průběžně monitoruje obtížnost úlohy a/nebo úspěšnost svého dosavadního postupu řešení apod.), a začala proto být za tímto účelem hojně používána (Afflerbach, 2000; Schellings et al., 2013).

Metoda spočívá v tom, že posuzovaná osoba je předtím, než se pustí do samotného řešení problému, instruována, aby v průběhu řešení průběžně verbalizovala své myšlení (tedy „myslela nahlas“). Klíčovým rysem metody je fakt, že tato verbalizace má probíhat zcela

spontánně a samotný slovní projev nemá být ze strany posuzované osoby nijak plánován či promyšlen s ohledem na srozumitelnost vnějšímu pozorovateli (např. experimentátorovi). Je tedy zcela v pořádku, pokud proband používá při „myšlení nahlas“ např. gramaticky chybné či nedokončené věty, případně pouze jednotlivá slova a slovní spojení, pokud verbalizovaná sdělení nijak dále nevysvětluje, neposkytuje k nim kontext, ani neuvažuje o tom, zda se v nich neodkazuje ke skutečnostem, které experimentátor nemůže znát a podobně. Optimální rovněž je, aby v průběhu řešení úkolu doprovázeného myšlením nahlas probíhala v co nejmenší míře interakce mezi probandem a experimentátorem či diagnostikem. Úlohou posuzujícího je verbalizaci probanda pouze pasivně sledovat a pouze v situacích, kdy proband přestává verbalizovat (což se při aplikaci této metody stává poměrně často), jej nenásilně povzbudit k pokračování ve slovním projevu – instrukcí typu „Prosím, stále myslíte nahlas“ a podobně. Takto podněcený verbální projev posuzované osoby je nahráván, audionahrávka je následně přepsána a získaný záznam (think-aloud protocol, protokol myšlení nahlas) je dále analyzován, přičemž konkrétní způsob skórování prvků, které jsou v protokolu přítomny, a jejich další analýzy se již mohou napříč různými výzkumy či jinými případy použít metody navzájem dost lišit.

Lišit se také mohou úkoly, které jsou probandovi zadány za účelem vyvolání metakognitivní aktivity. Mezi časté podnětové materiály patří matematické či logické příklady, úlohy vyžadující naplánování určitého sledu kroků pro dosažení stanoveného cíle, v pedagogickém výzkumu je také časté využití čtených učebních textů. V posledně uvedeném případě čte proband nahlas text, z něhož se má naučit určité informace. V případě, že jej ke čtenému napadá něco dalšího, pak čtení přeruší a i tyto myšlenky verbalizuje. V přepsaném protokolu je pak nutné odlišit tyto verbalizace od samotného podnětového textu, včetně případného označení úseků, které proband při čtení vynechal – i toto vynechání může nést určitou informační hodnotu.

V závislosti na druhu, obtížnosti a délce podnětových úkolů mohou některé protokoly myšlení nahlas od jednoho subjektu zahrnovat až desítky stran. Pro ilustraci si ovšem metodu přiblížíme pouze na dvou krátkých ukázkách. První příklad pochází z čistě kognitivního výzkumu, sledoval se zde postup probandů při řešení matematických příkladů z paměti:

36 krát 24... 4 ... přenesu ..., ne počkat .. 4 ... přenesu dvojku...14 144 ... 0 ... 36 krát 2 je ... 12 ... 6 ... 72 ... 720 plus 144 ... 4 ... hm hm ...6 8 ... hm ... 864. (Ericsson & Simon, 1993, s. 347)

Druhý příklad již pochází z článku, který se přímo zabýval tématem souvisejícím s metakognicí (Cromley & Azevedo, 2006). Participanti se v tomto výzkumu učili z poměrně obtížného textu o americké válce za nezávislost. Text byl záměrně vybrán tak, aby od čtenářů vyžadoval četné vyvozování informací, které nebyly explicitně uvedeny a znovuvybavení dříve osvojených znalostí, jež poskytovaly potřebný kontext. Přepis (resp. český překlad přepisu) je uveden ve formátu, v jakém s ním pracovali samotní autoři, tedy čtení ze samotného zdrojo-

vého textu je zvýrazněno podtržením, doprovodná verbalizace participanta je text bez podtržení. Pokud určitá část textu byla následně kódována jako projev konkrétní metakognitivní aktivity (např. aktivace předchozích znalostí, sumarizace aj.), je vždy uvedena na samostatném řádku.

Protesty proti zdanění bez zastoupení přicházely nicméně především od obchodníků z Nové Anglie, které nejvíce zasáhl Zákon o cukru. To mě připomíná, prostě, myslím, že někdo někdy řekl, že během otrokářství...

Žádné zdanění bez zastoupení

Zákon o kolkovném

Aby získal ještě větší příjmy na úhradu obrany kolonistů, schválil parlament Zákon o kolkovném v roce 1765. Snažili se dostat peníze od lidí v koloniích, aby platili za svoje vlastní věci. (Cromley & Azevedo, 2006, s. 237)

Protokoly v obou uvedených příkladech jsou kvalitní v tom smyslu, že probandi byli zjevně instruováni, aby se nesnažili svoje spontánní verbalizace před vyslovením ještě dodatečně uspořádat či upravovat, a tuto instrukci se jim podařilo dodržet. Zároveň je z toho patrné, že pro člověka, který není s technikou obeznámen, se může typický protokol zdát obtížně srozumitelným a chaotickým a vyhodnocení tohoto typu dat vyžaduje od posuzovatelů určitou zkušenost a cvik. Na druhou stranu právě díky absenci dodatečného plánování a reformulací poskytují právě takové protokoly teoreticky „nejčistší“ a nejméně zkreslená data o skutečném průběhu (meta)kognitivní aktivity u sledovaných probandů.

U čtenáře, který se s protokolovou analýzou setkává poprvé, by mohl snadno vzniknout mylný dojem, že se jedná o metodu vhodnou a užívanou pouze pro kvalitativní výzkum či obecně kvalitativní práci s daty. Je pravdou, že mnohé práce využívající analýzu TAP k výzkumu metakognice, zejména v prvních letech po zavedení tohoto konstruktů, měly skutečně výrazné kvalitativní zaměření, tj. soustředily se zejména na to, aby v protokolech identifikovaly konkrétní typy metakognitivních strategií, které lidé při učení či řešení problémů používají. Tato úvodní fáze nepochybně významně přispěla k hlubšímu poznání metakognice; sám Flavell (1979) ve svém originálním článku vymezil pouze 3 strategie jako potenciální předmět dalšího empirického zkoumání – a to konkrétně plánování, monitorování a evaluaci. Teprve další výzkumy, převážně realizované v 80. a na počátku 90. let minulého století, tento inventář rozšířily do podoby, kterou např. zachytili ve své přehledové studii Schellingsová a kolegové (2013). Tento seznam můžeme chápat jako určitý vzor či základní rámec, ze kterého vychází mnoho obdobných taxonomií, se kterými se setkáváme v pozdějších výzkumech.

Jakmile ovšem existuje předem daný inventář metakognitivních aktivit společně s kritérii, jak jednotlivé výroky či jiné aspekty (např. intonační zdůraznění určité části čteného textu či naopak její vynechání atd.) v protokolech myšlení nahlas posuzovat, může být protokolová analýza snadno použita jako zdroj dat pro další statistické zpracování. Konkrétní seznam kategorií a konkrétní způsob skórování se mezi jednotlivými autory dost liší a nepochybně by bylo možné navrhnout i alternativní způsoby, které zatím v žádném publikovaném výzkumu nikdo nepoužil. Protokol konkrétní posuzované osoby lze například skórovat podle toho, kolik jednotlivých metakognitivních aktivit se v něm podařilo identifikovat (bez ohledu na to, jestli se výrok signalizující danou aktivitu objevil pouze jednou, nebo mnohokrát). Alternativně lze naopak bodovat každý konkrétní výskyt každé konkrétní aktivity. Skórování jednotlivých výskytů je dále možné různým způsobem odstupňovat. Například ve studii Veenmana a van Cleefové (2019) byly ty prvky protokolů, které podle posuzovatelů signalizovaly, že daná metakognitivní aktivita byla „jednoznačně přítomna“, skórovány dvěma body, zatímco prvky signalizující zahájení, ale nedokončení aktivity, byly hodnoceny jedním bodem. Nepřítomnost metakognitivní aktivity byla pochopitelně skórována nulou.

Cromley a Azevedo (2006) v článku, z něhož pochází výše uvedená ukázka TAP při učení se z textu pojednávajícího o historii, použili kódovací schéma, které je v úplnosti popsáno jinde (Azevedo et al., 2004). V tomto schématu hodnotí posuzovatel při každém výskytu signálu určité metakognitivní aktivity také to, zda tato aktivita byla v dané situaci použita účelně (signováno +), či nikoli (signováno –). Například výrok probanda o otrokářství v uvedené ukázce byl hodnocen jako PKA- (prior knowledge activation, aktivování předchozích znalostí), protože tato znalost s daným úsekem čteného textu nesouvisí a nemohla podle posuzovatele přispět k lepšímu porozumění a naučení. Tato a mnohá další kódovací schémata mají jednu společnou výhodu – je možné ověřit jejich užitečnost stanovením základních psychometrických ukazatelů (shoda posuzovatelů, vnitřní konzistence, případně i test-retest reliability aj.) a získané metriky dále porovnávat s jinými kvantitativními daty.

Několik příkladů kategorií metakognitivních aktivit a prvků v TAP, ze kterých byla přítomnost těchto aktivit vysouzena, ukazuje Tabulka 1. Jak jsme již viděli, taxonomie metakognitivních aktivit, které lze registrovat v TAP, nejsou u různých autorů zcela jednotné. Tabulka 1 vychází z kategorizace použité v práci Meijera a kolegů (2012), která je pro ilustrační účely výhodná tím, že je relativně stručná. V dalších pracích, na nichž se vesměs podíleli autoři z téhož odborného týmu (Meijer et al., 2006; Schellings et al., 2013), se můžeme setkat s výrazně rozsáhlejší taxonomií, která obsahuje různé podkategorie, kategorie specifické pouze pro určitý typ úkolu (tj. určité metakognitivní aktivity se typicky vyskytují například pouze při výpočtu matematického příkladu, ale již ne při učení se z odborného textu atp.), s ohledem na stručnost zde ovšem tyto komplexní klasifikace uvedeny nejsou a případné zájemce lze odkázat přímo na původní zdroje. Konkrétní prvky (výroky) z protokolů myšlení nahlas, použité v Tabulce 1, pocházejí částečně z výše uvedených prací, částečně z vlastní experimentální zkušenosti (např. Portešová et al., 2021).

Tabulka 1.

Přehled základních metakognitivních aktivit a jejich indikátorů v TAP

Metakognitivní aktivita	Podrobnější charakteristika metakognitivní aktivity	Příklady projevů metakognitivní aktivity v TAP
Orientace	Ujasnění si a pochopení podstaty úkolu; identifikace podstatných informací; aktivace dříve nabytých znalostí	<p>„...jo voni tady chtějí vlastně jen změnu rychlosti, takže ta rotace mě vlastně vůbec nemusí zajímat...“</p> <p>„...ono je to v mílich... i ten výsledek asi budou chtít v mílich...ty jo kolik je míle... necelý 2 kilometry myslim...“</p> <p>[jako indikátor orientace může sloužit i opakované přečtení zadání, případně čtení určité části textu se zdůrazněnou intonací atd.]</p>
Plánování	Stanovení dílčích cílů; záměrné vyhledávání konkrétní informace v textu či v zadání	<p>„...tohle spočítám až potom, tady si radši teď nakreslím graf...“</p> <p>„...tady jako napřed si musím spočítat objem té levé části bazénu a pak té hlubší ... pak to teprve vynásobit tou hustotou...“</p>
Vlastní vykonání úkolu (executing) ⁵	Samotná kognitivní aktivita (čtení, počítání atd.)	<p>„...tak 12 krát 100 je 1200 a ještě k tomu přičíst 300...takže dohromady je to 1500...“</p>
Monitorování	Detekce inkonzistencí; uvědomění si, že určité informace subjektu nedávají smysl; posouzení snadnosti úlohy	<p>„... že forma vlády byla autokratická... tady vůbec nechápu, co tím mysleli...“</p> <p>„...in-ter-pre-ta-ce ... to je nějaký těžký slovo ... asi to znamená jako co si vo tom kdo myslí...“</p>
Evaluace	Průběžná kontrola mezivýsledků; porovnávání dosaženého s cílem	<p>„...tak ta rychlost na konci musí být stejná jako na začátku a na začátku to je 30 kilometrů a na konci 60 takže tam je něco blbě...“</p>
Elaborace	Parafrázování, vyvozování	<p>„...když tady píšou že byla krize, tak to taky znamená, že asi byla taky nezaměstnanost...“</p>

⁵ Striktně vzato nepředstavuje tato konkrétní kategorie metakognitivní aktivitu, ale „jenom“ aktivitu kognitivní. Zároveň je ale zřejmé, že v každém skórovacím systému využívajícím TAP pro výzkum metakognice musí být nějaká kategorie tohoto druhu obsažena – při optimálním plnění úkolu není možné kognitivní aktivity pouze plánovat či monitorovat, ale je také nutné je reálně vykonat.

Z velkého množství zahraničních prací využívajících analýzu TAP zde zmíníme pro ilustraci tři konkrétní příklady:

Dentonová a kolegové (2015) provedli výzkum, v němž zaznamenali celkově vyšší sklon k užívání metakognitivního monitorování u vysoce úspěšných studentů, ve srovnání se studenty méně úspěšnými.

Výskyt metakognitivních strategií ve vztahu k úrovni studijního výkonu byl také obsahem studie, kterou realizovali Bråten a Strømsø (2003). Na jejich výsledcích je zajímavé to, že pozorovali nejen (očekávanou) vyšší frekvenci metakognitivních strategií, jako je monitorování či organizování u studijně úspěšnějších žáků, ale i určité strategie, které byly čtenější naopak u žáků s horšími studijními výsledky. Tyto strategie – např. memorování a elaborace, ačkoli mnohdy mohou být velmi prospěšné, byly zřejmě v případě konkrétního sledovaného úkolu nedostatečně efektivní. Zajímavý byl také časový průběh, což je další typ informací, který dokáže protokolová analýza zachytit a zpracovat; v případě této studie se ukázalo, že zatímco studijně úspěšní žáci měli v průběhu řešení vzrůstající tendenci používat metakognitivní monitorování, u kontrolní skupiny bylo naopak monitorování relativně častější na počátku úkolu a v jeho průběhu jej účastníci používali stále méně.

Vedle srovnání nadaných a běžných žáků (obdobně lze ovšem samozřejmě porovnávat i mnoho jiných specifických skupin) zmiňme i využití protokolové analýzy pro ověřování efektivity mentorských a intervenčních programů zaměřených na rozvoj metakognitivních schopností u žáků. Za tímto účelem realizovali výzkum s využitím analýzy TAP např. Azevedo et al. (2008) – s pozitivním výsledkem, tj. u žáků po intervenci byly v protokolech účastníků patrně častější projevy využívání metakognitivních strategií.



Máme-li zhodnotit protokolovou analýzu jako celek, nemůžeme se vyhnout základní metodologické otázce týkající se toho, nakolik jsou data, které analýza TAP poskytuje, validní. Hu a Gao (2017) používají v této souvislosti dva pojmy: *reaktivita* (reactivity) a *přesnost* (veridicality). Reaktivita označuje skutečnost, že samotný akt měření, tj. nahrávání za účelem TAP, může ovlivňovat (meta)kognitivní děje, které se pokoušíme sledovat. Přesnost je pojem, který v tomto kontextu znamená, že v získaném protokolu jsou zachyceny všechny děje, které se v mysli probanda v průběhu sledování skutečně odehrály, a naopak že žádný prvek protokolu nesignalizuje chybně mentální děj, ke kterému ve skutečnosti nedošlo. Oba pojmy nelze zaměňovat, protože protokol může např. zcela věrně zachycovat postup řešení zadané úlohy, který by ovšem proběhl jinak, kdyby úlohu posuzovaná osoba řešila v soukromí a bez doprovodné verbalizace. Naopak, samotný proces snímání TAP nemusí způsob řešení ovlivnit, ale v protokolu mohou chybět důležité prvky, třeba proto, že posuzovaná osoba v některých úsecích přestala verbalizovat a examinátor ji k tomu včas znovu nepodnítil.

Otázkou reaktivity se intenzivně zabývali již původní významní představitelé a propagátoři analýzy TAP jakožto metody (Ericsson & Simon, 1993). Tito autoři vycházeli z předpokladu, že pokud je tentýž úkol zadán dvěma skupinám probandů, které se liší pouze tím, že u jedné zároveň probíhá snímání TAP, a obě skupiny se přitom neliší ve výsledcích, kterých při plnění úkolu dosáhnou, svědčí to o přijatelně nízké úrovni reaktivity analýzy TAP jakožto použité metody. Tito autoři pak srovnali velké množství různých (kognitivních) úkolů a došli k závěru, že u tzv. verbalizací 1. a 2. typu (zjednodušeně řečeno, jedná se o způsob verbalizování uvedený výše, kdy sledovaná osoba spontánně vyslovuje své myšlenky, aniž by se je snažila nějak utřídit, vysvětlovat či uspořádat s ohledem na gramatiku) nemá samotná aplikace metody žádný podstatný vliv⁶. Bannertová a Mengelkamp (2008) pak dospěli k podobným závěrům v případě úkolů cílených záměrně na výzkum metakognice. Pro úplnost je ovšem třeba dodat, že některé novější studie tento závěr zpochybňují (Hu & Gao, 2017). Ukazuje se, že jedním z důležitých intervenujících faktorů by mohl být věk probandů, a to v tom smyslu, že u mladších dětí by mohla být míra reaktivního ovlivnění sledované mentální aktivity větší než u adolescentů či dospělých osob. Dále je třeba dodat, že i v případě, kdy snímání TAP neovlivní celkové výsledky vykonávané úlohy, téměř vždy ovlivní čas řešení. Protokolovou analýzu proto není vhodné používat ve výzkumných designech, kdy čas řešení vystupuje jako sledovaná závislá proměnná (Bowles, 2010).

Pokud jde o přesnost, může tuto charakteristiku protokolů myšlení nahlas ovlivnit několik faktorů, které je nutné při plánování výzkumu či individuální diagnostiky brát v úvahu. Za prvé je to otázka užití analýzy TAP při plnění úkolů, které se obecně obtížně verbalizují (např. geometrické úlohy, úlohy vyžadující vizuální představu ve 3 dimenzích, mentální rotace předmětů atd.). Autoři zabývající se metodologií protokolové analýzy nedospěli v tomto ohledu k jednoznačným závěrům a existují i studie, kdy byla analýza TAP úspěšně použita např. při řešení geometrických úloh (Ericsson & Simon, 1993). Menší problém to samozřejmě představuje, pokud pro všechny prvky úlohy existují příslušné pojmy – např. přepona, odvěsna, kolmý atd. Na základě dalších zjištění lze nicméně doporučit, aby pokud je základním důvodem použití analýzy TAP studium metakognice a vizuálně náročný úkol nesouvisí přímo s podstatou výzkumné otázky, badatelé záměrně volili spíše takové typy podnětových úkolů, které lze verbalizovat relativně snadno.

Ještě důležitějšími faktory jsou ovšem úroveň verbálních schopností a relativní obtížnost úkolu použitého ke stimulaci metakognitivní odezvy ve vztahu k rozumovým schopnostem posuzované osoby. Obecně je pro uspokojivou přesnost TAP potřebná velmi dobrá úroveň

⁶ Přestože intuitivně by se mohlo zdát, že důsledkem verbalizace 3. typu bude narušení a snížení výkonu, bylo prokázáno, že tento typ verbalizace paradoxně výkon někdy naopak zvyšuje. Např. pokusné osoby, které průběžně verbalizovaly svoje myšlení takovým způsobem, aby byl jejich projev koherentní a pro experimentátora srozumitelný, dosahovaly lepších výsledků v kognitivně-psychologických zkouškách, jakými jsou např. Hanojská věž či Wasonova úloha (Ericsson & Simon, 1993).

verbálních schopností, a protokolovou analýzu proto nelze korektně použít u velmi malých dětí a u osob s jazykovými deficity, problematické je také to, má-li myšlení nahlas probíhat v jiném než mateřském jazyce, či pokud jsou v cizím jazyce předkládány podnětové úkoly. Nedostatečná jazyková úroveň může vést jednak pochopitelně k tomu, že proband si nevybaví vhodný výraz pro popis aktuálně probíhající mentální aktivity, což se projeví krátkými či „prázdnými“ protokoly. Neméně důležitý je i druhý efekt, který spočívá v dodatečné zátěži, kterou u osoby s nedostatečnou jazykovou kompetencí vyvolá nutnost verbalizovat a která se přidává k zátěži vyvolané samotným úkolem. Pravděpodobným výsledkem je v takové situaci zvýšení reaktivity, tedy ovlivnění probíhající mentální aktivity samotným procesem pozorování.

Pokud jde o podnětový úkol, je při výzkumu metakognice nutné, aby byl zhruba středně obtížný vzhledem ke kognitivním schopnostem a zkušenostem posuzované osoby (Veenman, 2011). Příliš snadné úkoly bývají obvykle zcela zautomatizovány a lidé je dokáží řešit velmi rychle a bez nutnosti používat k tomu jakékoli metakognitivní strategie. Ukázkou nevhodného stimulu je použití následujícího příkladu u dospělých probandů (nebo i u velmi malých, ale výrazně nadaných dětí): „*Turista ušel napřed 2 km a potom ještě 3 km. Kolik ušel dohromady kilometrů?*“. Pokud je naopak experimentální úkol na samotné hranici schopností posuzované osoby, hrozí, že se nutnost verbalizace, přestože sama o sobě představuje pouze nízkou zátěž, stane tím posledním, kritickým vlivem, který povede např. k zahlcení pracovní paměti a k selhání v úloze. V obou případech budou opět výsledkem převážně krátké, a navíc nepřesné protokoly nevhodné k další analýze.

Tento „faktor střední obtížnosti“ může být obzvlášť zrádný ve výzkumných designech, kdy jsou porovnávány nadané děti s běžnou populací. Pokud totiž obě skupiny pracují se stejnými podnětovými materiály, může se snadno stát, že zadaný úkol je optimálně náročný pro kontrolní skupinu, ale příliš jednoduchý pro nadané žáky. Analýza TAP pak může vést k nesprávnému závěru, že nadaní v menší míře používají metakognitivní strategie ve srovnání se svými vrstevníky, přestože takový závěr je pouze výzkumným artefaktem.

Pokusíme-li se protokolovou analýzu na závěr krátce zhodnotit, pak hlavní výhodu lze spatřovat v tom, že jde o metodu propracovanou, etablovanou a s dlouhou výzkumnou tradicí. S pomocí analýzy TAP bylo v zahraničí realizováno velké množství výzkumů, kterými se lze inspirovat a na které lze navazovat. Dále k protokolové analýze existují i čistě metodologické publikace, z nichž lze čerpat doplňující informace o různých variantách použití, o alternativních vyhodnocovacích schématech apod. (zejména je možné doporučit sice již starší, ale stále velmi užitečnou knihu Ericssona a Simona, 1993). Z pohledu výzkumu metakognice umožňuje analýza TAP přístup k jevům, které nelze přímo pozorovat a ke kterým by se jinak badatelé alternativně dostávali jen obtížně, nebo vůbec. Jak už ale bylo naznačeno výše, metoda rozhodně není všespasitelná a existují situace, kdy její použití vhodné není. Za další nevýhody je nutné zmínit také značnou časovou náročnost celé procedury, umocněnou tím, že získávání dat musí prakticky vždy probíhat v individuálních sezeních.

Analýza postupů registrovaných pomocí počítače

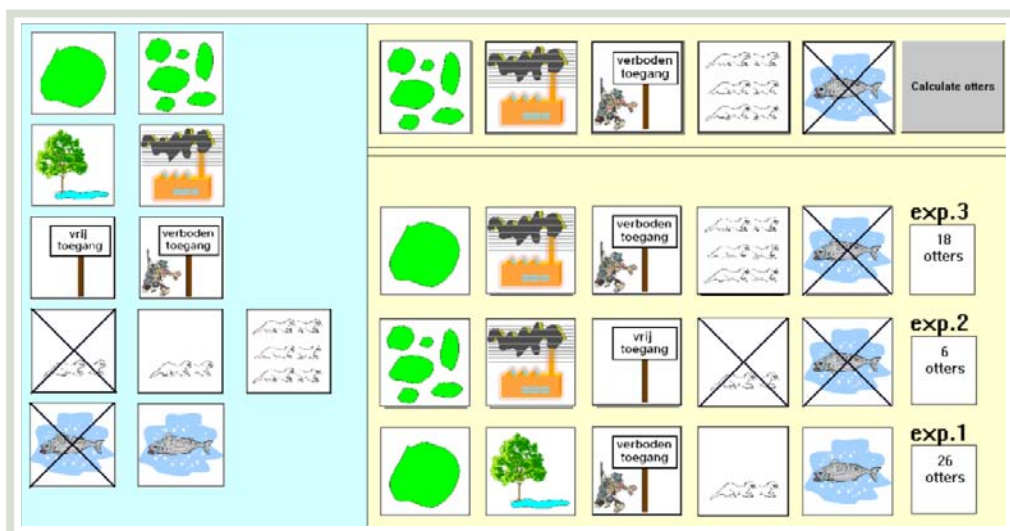
Další třída on-line metod pro měření metakognice nemá nějaký ustálený a všeobecně užívaný název. Například Veenman et al. (2014a, b) používají při popisu své metody, která spadá do této kategorie, termínu *analýza počítačových záznamů* (analysis of computer logfiles), setkat se ale můžeme i s označením *sledování aktivit* (tracing; Kinnunen & Vauras, 2010), i s dalšími. Označení, které zde budeme dále používat, má za cíl vystihnout hlavní rys této skupiny metod, které mohou být administrovány výhradně za použití výpočetní techniky a speciálně vyvinutého či upraveného softwaru. Jejich základní princip je jednoduchý: posuzované osobě je prezentován určitý problém k vyřešení či určitý textový materiál k nastudování, které vždy probíhá ve standardním, předem připraveném softwarovém prostředí. Tím je jednak kontrolováno, jaké dílčí kroky mohou probandi při řešení vůbec realizovat a jakými konkrétními informacemi mohou disponovat (buď automaticky, nebo na vyžádání), především ale příslušný program zaznamenává, jaké konkrétní aktivity, v jakém čase a v jakém pořadí ten který proband v průběhu řešení provedl.

Takto získaný záznam poté může být „ručně“, častěji však automaticky, skórován a vyhodnocen a z profilu zaznamenaných aktivit je možné odvodit míru zapojení předpokládaných metakognitivních dovedností. Například na nízkou míru metakognitivního monitorování můžeme usuzovat tehdy, je-li probandovi prezentováno v řadě napřed několik jednoduchých úkolů následovaných úkolem výrazně těžším a ze záznamu je patrné, že proband nevěnoval řešení těžší úlohy o nic více času, než předcházejícím problémům jednodušším (obzvlášť pokud těžší úlohu zároveň vyřeší chybně). Naopak pokud proband narazí na nejednoznačný problém, pro nějž je možné nalézt dvě nebo i více řešení, z nichž každé je z určitého pohledu správné, a tento rozpor jej vede k aktivnímu vyhledávání dalších informací dostupných v rámci dané aplikace, lze to interpretovat jako projev adekvátně použité metakognitivní strategie. Diagnostických či experimentálních aplikací tohoto druhu bylo v zahraničí vytvořeno poměrně dost, zde si pro ilustraci přiblížíme tři konkrétní příklady.

První z představených nástrojů vyvinuli původně Wilhelm, Beishuizen a van Rijn (2005), později metodu upravil tým odborníků kolem M. Veenmana (Veenman et al., 2014a, b). Modifikace spočívala především ve zvýšení obtížnosti podnětových úloh, díky čemuž je metoda lépe využitelná při zjišťování metakognice u nadaných žáků a studentů. Z pohledu testované osoby má aplikace podobu výukového prostředí, kde proband provádí řadu jednoduchých „experimentů“ a sleduje jejich vliv na hodnotu jedné závislé proměnné. Metoda má více různých variant, ve verzi popsané ve výše odkazovaných publikacích je sledovanou proměnnou velikost populace vyder, které obývají určité území (v alternativních verzích se sleduje např. působení různých faktorů na růst rostlin, na zdraví člověka apod.).

V rámci jednoho experimentu proband libovolně volí hodnoty (ze dvou, případně tří možností) pěti nezávislých proměnných: *habitat* (zda může sledovaná populace žít a rozmnožovat se na jednom souvislém území, nebo zda je oblast výskytu rozdělena na jednotlivé „ostrovky“), *úroveň znečištění okolí*, *přístupnost lokality* (lokalita je volně přístupná vs. pro veřejnost

platí v lokalitě zákaz vstupu) a fakt, zda jsou v průběhu roku do lokality navíc *uměle vysazovány k již přítomné populaci další páry vyder*. Posledním faktorem, jehož hodnotu může proband měnit, je *přikrmování versus nepřikrmování ryb* ve sledované lokalitě v průběhu zimy. Každý „experiment“ spočívá v tom, že testovaná osoba prostřednictvím odpovídajících piktogramů navolí kombinaci všech pěti faktorů (tedy např. zvolí, že vydří lokalita je rozdělena do několika izolovaných oblastí, je znečištěná, veřejnost do ní má zakázaný přístup, uměle zde byly navíc vysazeny 3 páry vyder a nedochází zde k přikrmování ryb během zimy). Aplikace následně vygeneruje a sdělí údaj o velikosti vydří populace za těchto konkrétních podmínek. Kombinace navolené v již uskutečněných experimentech spolu s výslednou hodnotou sledované závislé proměnné jsou zachovány v pravé části aplikačního okna (viz Obr. 1) a pokud testovaná osoba provedla již více než 5 experimentů, může si starší výsledky prohlížet prostřednictvím scrollování nahoru či dolů.



Obrázek 1.

Ukázka experimentálního prostředí z metody používané a prezentované Veenmanem a spolupracovníky (Veenman et al., 2014a, b)

Úkolem, který je posuzované osobě explicitně zadán, je odhalit, jaký vliv mají jednotlivé faktory na celkovou velikost vydří populace. Aplikace je přednastavena tak, že některé faktory mají vliv přímý, ať již pozitivní nebo negativní (habitat, míra znečištění), některé jsou irelevantní a ve skutečnosti nemají vliv žádný (přístup veřejnosti) a některé projevují svůj vliv ve vzájemné interakci (např. přikrmování ryb má efekt pouze v interakci s mírou znečištění). Mnohem důležitějším prvkem diagnostického procesu je ovšem skutečnost, že aplikace

registruje několik dílčích ukazatelů, u kterých se předpokládá přímá souvislost s úrovní a kvalitou metakognitivních dovedností zapojených při řešení těchto úloh. Úplný výčet a podrobnější popis a zdůvodnění použití těchto indikátorů uvádí např. Veenman et al. (2014a), pro ilustraci zde zmíníme např. indikátor *Votat.pos*, což je počet přechodů (tj. bezprostředně po sobě následujících experimentů), při nichž proband změnil hodnotu pouze jedné proměnné, či indikátor *Thinktime*, což je součet časů (v rámci celého testovacího sezení), které uplynou mezi zobrazením výsledku předchozího experimentu a prvním krokem (tj. volbou hodnoty první proměnné) v experimentu následujícím. Dále se registruje např. scrollování za účelem zobrazení výsledků předchozích experimentů, celkový počet experimentů či počet experimentů s unikátní kombinací zvolených faktorů.

Výběr a způsob hodnocení sledovaných metakognitivních indikátorů proběhl na základě studií, v nichž byly souběžně s administrací této metody snímány a analyzovány protokoly myšlení nahlas a metoda byla vůči výsledkům analýzy TAP validizována (Veenman 2014a, b). Rozložení vyhodnocovaných indikátorů je buď normální, nebo je v některých případech (indikátory vztahující se k využívání scrollování) na normální převedeno odmocninovou transformací. Celková úroveň metakognitivních dovedností posuzované osoby je pak vypočtena jako průměr dílčích indexů. Vzhledem k tomu, že metoda byla zamýšlena primárně k výzkumnému použití a počítá se s tím, že její administrátoři se dobře orientují v základních psychometrických pojmech, je výstupem diagnostiky prostý z-skór, který se dále nepřevádí na žádnou další, laikům lépe srozumitelnou škálu (typu T-skór či steny).

Kromě společných výhod všech metod postavených na analýze počítačově registrovaných záznamů je další předností představeného diagnostického nástroje skutečnost, že i při relativně vysoké obtížnosti podnětových úkolů nevyžaduje od probandů konkrétní znalosti v rámci nějakého vědního oboru či školního předmětu a vystačí si s poměrně krátkou a stručnou instrukcí. Nevýhodou může být určitá zastaralost grafického prostředí a z pohledu českých odborníků také skutečnost, že metoda není ani volně šiřitelná, ani běžně dostupná na komerční bázi. Její hlavní doménou jsou především výzkumy, ve kterých byl zapojený někdo z jejích autorů či jejich spolupracovníků. V Nizozemí byla také určitý čas používána při přijímacím řízení na některé přírodovědné vysokoškolské obory (Veenman, 2016). Metoda je na druhou stranu natolik dobře popsána a její uživatelské rozhraní je natolik jednoduché, že pro odborníky se znalostí aplikovaných programovacích prostředí typu PsychoPy by zřejmě nepředstavovalo velký problém vytvořit si nástroj podobného druhu ve vlastní režii.

Na velmi podobném principu je založena i druhá z metod, kterou si zde krátce představíme, a sice metoda IMMEX (Interactive Multimedia Exercises; Cooper et al., 2008). Podobně jako v předchozím případě, i zde pracuje testovaná osoba v určitém výukovém prostředí, provádí v něm virtuální experimenty a její aktivita je přitom zaznamenávána a následně vyhodnocena. Základní rozdíl spočívá v tom, že v případě IMMEXu jsou podnětové úkoly navázány na určitý vědní obor, konkrétně na chemii. Nástroj proto není tak univerzálně použitelný jako předchozí metoda, a má-li vést ke spolehlivým a relevantním výsledkům, vyžaduje to od probandů znalosti chemie alespoň na středoškolské úrovni, např. včetně principů, podle nichž se laboratorně

určují chemické látky apod. Posuzované osobě je postupně předkládána řada úkolů, které spočívají v identifikaci neznámé látky. Za tímto účelem může proband provádět jednak jednotlivé zkoušky (např. zjišťovat barvu plamene, rozpustit látku a zjišťovat vlastnosti roztoku atp.), jednak si může přečíst některé doplňující informace v „knihovně“, která je také součástí výukového prostředí. Provedení každé z těchto aktivit je zaznamenáváno a na základě těchto záznamů proběhne posléze vyhodnocení metakognitivních dovedností posuzovaného studenta.

Od nástroje rozvíjeného Veenmanem a spolupracovníky se metoda IMMEX liší také celkově sofistikovanějším způsobem vyhodnocení. V pozadí aplikace stojí umělá neuronová síť, která na principu skrytých markovovských modelů (Hidden Markov Model, HMM) zpracovává jednotlivé záznamy. Architektura a konkrétní technické provedení tohoto vyhodnocovacího mechanismu přesahuje rozsah této knihy, zájemce o podrobnější informace je možné odkázat na publikaci Stevense a kolegů (Stevens et al., 2005). Z pohledu uživatele – pedagoga či psychologa – lze nicméně poměrně srozumitelně popsat finální výstup diagnostiky prováděné pomocí metody IMMEX. Souhrn činností každého probanda je zařazen do jednoho z pěti *strategických stavů*, které jsou chápány jako projev nízkých, středních či vysoce rozvinutých metakognitivních dovedností zapojených do řešení úloh. Stavů jsou jednoduše označeny arabskými číslicemi a jejich charakteristika a interpretace je následující:

Stav č. 1 – proband použil celkově málo položek dostupných v aplikaci (tj. experimentů nebo informací z „knihovny“) a postupoval přitom nesystematicky a náhodně → signalizuje nízkou úroveň metakognice

Stav č. 2 – proband použil relativně větší množství prvků, zhruba rovnoměrně mezi nimi byly zastoupeny výsledky vlastního experimentování a informace připravené v „knihovně“ → signalizuje střední úroveň metakognice

Stav č. 3 – charakteristické je využití extrémně vysokého množství prvků nabízených aplikací, výrazně většího, než bylo nezbytné pro splnění úkolu (tj. testovaná osoba např. prováděla „zbytečné“ pokusy v době, kdy už by jí informace, které zatím získala, postačovaly pro identifikaci neznámé chemické látky apod.) → signalizuje nízkou úroveň metakognice

Stav č. 4 – proband prováděl velké množství vlastních experimentů (testů) a v malé míře využíval doplňující informace z „knihovny“ → signalizuje střední úroveň metakognice.

Stav č. 5 – vyznačuje se efektivním využitím relativně malého množství položek, které aplikace nabízí. Na rozdíl od stavu č. 1 jsou tyto položky vybírány systematicky a tento výběr vede k úspěšnému splnění úlohy → signalizuje vysokou úroveň metakognice.

Další výhoda IMMEXu spočívá v tom, že pro jednotlivé úkoly je k dispozici údaj o jejich obtížnosti, a pro tutéž posuzovanou osobu je tedy možné určit zároveň úroveň její metakognice i úroveň jejích schopností (ve smyslu správnosti řešení zadaných podnětových úloh). Následně je také možné sledovat, zda se metakognitivní strategie u posuzované osoby liší u úloh, které jsou pro ni relativně jednoduché, od strategií uplatněných při úkolech relativně obtížných.

Určitým omezením metod, které mají povahu umělého prostředí, v němž testované osoby experimentují či obdobným způsobem řeší nějaký úkol, je skutečnost, že tyto metody z velké míry pomíjejí čtení delších textů. Tato forma učení (se) přitom představuje zásadní součást vzdělávání, minimálně u starších dětí, adolescentů a mladých dospělých. Jako třetí ilustraci proto uvedeme metodu, kterou vyvinuly Kinnunenová a Vaurasová (2010) a která se naopak právě na tuto činnost cíleně zaměřuje.

Metoda tichého registrovaného čtení (traced silent reading) využívá jednoduchý software, do kterého je možné nahrát podnětový text. V průběhu samotného čtení je následně většina textu maskována a probandům se zobrazuje pouze aktuálně čtený prvek (v závislosti na nastavení mohou jako základní jednotky vystupovat buď jednotlivá slova, nebo celé věty). Maskování ostatního textu probíhá tak, že každé jednotlivé písmeno je nahrazeno jiným znakem, nahrazující znak je totožný pro celý text (např. se může jednat o číslici 0). Například věta:

Tento druh ptáků si nestaví vlastní hnízda.

Vypadá v určité fázi procesu čtení takto:

00000 0000 ptáků 00 0000000 0000000 000000.

Probandi se v textu posunují pomocí klávesových šipek doleva či doprava, např. při jednoduchém stisknutí šipky doprava se zobrazí následující slovo (či věta), zatímco předcházející, dosud viditelný prvek se naopak zamaskuje. Veškeré interakce probandů s textem jsou opět registrovány, takže je mimo jiné zaznamenána celková rychlost čtení, délka zobrazení každého prvku (slova či věty), to, kolikrát a kam se proband v textu vracel, aby si znovu ověřil již dříve přečtenou informaci atd. Autorky prostřednictvím tohoto nástroje studovaly například čtení textů obsahujících různé typy diskrepancí, jako je přítomnost neznámého či obsahově se nehodícího slova ve větě, skutečně či zdánlivě si protirečící tvrzení v různých částech téhož textu, informace, které jsou v rozporu s obecnými znalostmi o světě, jaké můžeme předpokládat u většiny probandů apod. Za projev adekvátně rozvinutého metakognitivního monitorování se považuje to, reaguje-li na tyto diskrepance posuzovaná osoba jinak, než na méně problematické části textu (např. je čte opakovaně a tedy déle, častěji se k nim vrací nebo ji podnítí k opětovnému ověření si jiné části textu atd.)⁷.

⁷ Metod tohoto typu existuje samozřejmě více. Uvedené tři příklady mají navíc společné to, že se jedná o čistě „diagnostické“ nástroje, tj. nástroje které slouží pouze ke zjišťování metakognitivních schopností a nemají žádný další účel. Velmi zajímavá je ovšem i jiná skupina softwarových aplikací, které slouží primárně jako výukové systémy (tutoring systems) – jejich hlavním cílem je tedy rozvíjet metakognitivní dovednosti, nebo rozvíjet znalosti určitého vědního oboru s využitím metakognice. Zároveň ovšem dokáží tyto nástroje také registrovat činnost studenta, a i ony tak mají určitý diagnostický přesah. Problematika těchto metod je příliš komplexní na to, abychom se jim zde mohli věnovat, zájemce lze odkázat např. na následující zdroje: Schwartz et al. (2009), Lajoie et al. (2015) či Poitras (2015).

Sledování očních pohybů (eye-tracking)

Metoda sledování očních pohybů, z jejichž záznamů se pak zpětně usuzuje na duševní procesy probíhající ve chvíli sledování u pozorovaného jedince, je stará více než sto let. Teprve koncem 70. let minulého století se nicméně přístrojové vybavení nezbytné pro záznam očních pohybů stalo běžně komerčně dostupným a finančně dosažitelným, což vedlo následně k velkému nárůstu empirických studií využívajících tuto metodu nejen v psychologii, ale i v mnoha dalších oborech (např. lingvistika, kartografie, marketing aj.; Holmqvist et al., 2011). Toto období se z velké části kryje s obdobím intenzivního zájmu o studium metakognice a bylo by tedy logické předpokládat, že mezi oběma trendy vznikne určitá synergie. Z více či méně zjevných důvodů k tomu ovšem nedošlo, a jak konstatují van Gog a Jarodzka (2013), zatímco eye-tracking byl v nedávné minulosti intenzivně využíván ke studiu mnoha kognitivních jevů, jeho využití pro výzkum metakognice bylo nepoměrně skromnější. Alemdag a Cagiltay (2018) dospívají k podobnému závěru i v článku z konce minulé dekády, takže můžeme soudit, že tento stav přetrvává i do současnosti. Existují samozřejmě důležité výjimky – o některých konkrétních se ještě stručně zmíníme v dalším textu – celkově ovšem bylo metakognitivně zaměřených studií využívajících eye-tracking provedeno a publikováno poměrně málo, takže z nich ani nemohla být provedena nějaká autoritativní metodologická syntéza (ve smyslu doporučení, které metriky dostupné prostřednictvím eye-trackingu jsou vhodné jako indikátory těch kterých metakognitivních jevů, které metriky se naopak neosvědčily, jakých interpretačních zkreslení je nutné se především vyvarovat atp.).

Typický experiment využívající eye-tracking probíhá tak, že zkoumaný jedinec sleduje vymezené zorné pole, zatímco jeho oči (případně pouze jedno oko) jsou snímány speciální infračervenou kamerou a jejich poloha je průběžně registrována takovým způsobem, aby bylo možné zpětně určit, na kterou konkrétní část podnětového pole proband v daném čase zaměřoval svůj zrak. Ve velké části eye-trackingových výzkumů (a téměř výhradně ve studiích zaměřených na studium metakognice) tvoří podnětové pole obrazovka počítače, přestože – zpravidla za cenu nižší přesnosti – lze sledovat oční pohyby i v přirozeném prostředí (v místnosti, v otevřené krajině atd.). Při následném zpracování je obvykle podnětové pole rozčleněno na jednotlivé části (tzv. oblasti zájmu, areas of interest, AOI) a data registrovaná prostřednictvím eye-trackingu jsou analyzována ve vztahu k těmto oblastem (např. badatelé zajímá pouze ta část obrazovky, v níž se nachází odstavec s důležitým textem, část obsahující ilustraci či graf apod.).

Lidské zrakové vnímání probíhá zpravidla jako sled dvou vzájemně se střídajících stavů. Termín *fixace* označuje stav, kdy je oko téměř nehybné a jeho poloha je taková, aby obraz fixované oblasti zorného pole dopadal na část sítnice vyznačující se nejostřejším viděním (žlutá skvrna). Rychlý, skokový pohyb mezi dvěma po sobě následujícími fixacemi se označuje jako *sakáda* a během tohoto pohybu oko prakticky nevnímá a nepřijímá žádné vizuální informace. Trvání fixací a trvání i délka sakád se mohou lišit, mimo jiné v závislosti na druhu sledovaných podnětů, na povaze úkolu, který byl probandovi zadán, na charakteru duševních

procesů probíhajících během plnění úkolu apod. Typicky trvá jedna fixace v řádu nižších stovek milisekund, trvání sakády je obvykle několik desítek milisekund (Holmqvist et al., 2011). Odtud je patrné, že se jedná o velmi rychlé procesy, které probíhají v zásadě automaticky a subjekt si je není schopen uvědomovat, ani je cíleně vědomě kontrolovat. Při měření očních pohybů je zaznamenána přesná poloha místa, které se v daný okamžik promítá na žlutou skvrnu, na osách X a Y (pokud je podnětovým polem obrazovka počítače, jsou souřadnice obvykle vyjádřeny v pixelech na základě rozlišení monitoru). Četnost těchto záznamů je daná nastavenou frekvencí zobrazovací aparatury, např. přístroj s nastavenou frekvencí 250 Hz registruje pozici očí v intervalech po 4 milisekundách. Existuje několik algoritmů, kterými může speciální software vyvinutý pro zpracování eye-trackingových dat vypočítat z těchto hrubých dat výskyt fixací a sakád a z těchto údajů pak mohou být odvozeny další metriky pro následné statistické zpracování (Popelka, 2018).

Pro studie využívající eye-trackingu v kognitivní psychologii a v nepsychologických oborech bylo vyvinuto více než 100 různých metrik, které mohou být předmětem analýz (Holmqvist et al., 2011). Zmíníme zde pouze několik, které jsou používány nejčastěji (Popelka, 2018) a jejichž povaha je velmi srozumitelná a přímočará. Jde o *celkový počet fixací* (fixation count) – buď v celém podnětovém poli, nebo pouze ve vymezených oblastech zájmu. Dalšími metrikami týkajícími se fixací jsou *průměrné trvání fixace* (mean fixation duration) a *celkový součet trvání všech fixací* (total dwell time). Tyto dvě metriky na sobě nemusí být nutně závislé, protože pro určitého jedince nebo pro určitý druh podnětu může být charakteristický velký počet relativně krátkých fixací. Průměrné trvání fixace může být v takovém případě nízké, zatímco celkový součet trvání fixací může být naopak nadprůměrný. Dále se často sleduje *celková délka všech sakád* (overall saccade length, případně scanpath length) či *délka průměrné sakády*.

Uvedené metriky, stejně tak jako mnohé další z nich odvozené, závisí na fyzikálních vlastnostech podnětového pole, kromě toho ale byla prokázána jejich souvislost s charakterem kognitivního zpracování vnímaných informací. Například pokud jsou informace přijímané zrakem (ať už jde o text, či např. o grafiku) pro subjekt nějakým způsobem nejednoznačné či matoucí, projeví se to v poklesu celkového množství fixací a v nárůstu délky průměrné fixace (Holmqvist et al., 2011). Jako další příklad lze uvést očekávaný vyšší počet fixací v té oblasti zorného pole, která obsahuje důležitý informační prvek. Odtud zbývá již jen malý krok k myšlence využít některé metriky jako indikátory kvality zapojených metakognitivních procesů. Ve studiích, které z tohoto předpokladu vycházely, se jako potenciálně významný indikátor metakognitivní aktivity ukázaly jednak metriky související s počtem a délkou fixací, jednak *tranzice* (přechody) mezi jednotlivými oblastmi zájmu (tj. situace, kdy je fixace lokalizována v jedné oblasti zájmu a další fixace, která po ní bezprostředně následuje, je lokalizována v oblasti jiné). Další nálezy naznačují, že by mohla existovat souvislost mezi úrovní metakognice, a to zejména metakognitivního monitorování, a specifickým typem sakád označovaných jako *regrese*. Jedná se o sakády, které se objevují převážně při čtení a které mají opačný směr oproti běžnému levoprávnímu pohybu očí po řádku. Účelem

těchto zpětných pohybů je zřejmě znovu fixovat již přečtené úseky textu, pokud z nich při první fixaci subjekt nezískal s dostatečnou jistotou všechny potřebné informace (Holmqvist et al., 2011).



Z poměrně omezeného množství studií využívajících eye-tracking pro zjišťování metakognice můžeme zmínit pro ilustraci následující: Kinnunenová a Vaurasová (1995) srovnávaly v již starším výzkumu oční pohyby dětí s nadprůměrnými a podprůměrnými rozumovými schopnostmi. Oční pohyby sledovaly při čtení textů obsahujících různé typy inkonzistencí, např. neznámé slovo, vzájemně si protirečící informace atd. (jednalo se tedy o podobné inkonzistence, jako ve výše zmíněné studii tichého registrovaného čtení, kterou realizovaly tytéž autorky). U všech dětí byl zaznamenán zvýšený počet regresí a delší fixační čas v případech neznámých slov, ale pouze nadprůměrné děti reagovaly podobně i na jiné typy inkonzistencí (např. protirečící si informace).

Hyönä, Lorch, a Kaakinenová (2002) popsali odlišné typy lidí, pokud jde o strategie při čtení výukového textu. Zatímco nejméně rozvinuté úrovni metakognice odpovídala strategie, při níž sledovaná osoba pouze rychle a lineárně čte prezentovaný text, jedinci s nejvyspělejší strategií se často vraceli k důležitým částem textu a vyznačovali se častějším a delším fixováním nadpisů či posledních vět v odstavcích (které určitým způsobem stručně shrnovaly hlavní myšlenku celého odstavce). Jako poslední příklad uvedeme výzkum, který prováděli Masonová, Tornatora, a Pluchino (2013), a v němž žáci základní školy četli výukový text z fyziky (týkající se vlastností plynů), doprovázený ilustračním obrázkem umístěným vedle textu. Míra naučených informací byla následně posouzena pomocí krátkého písemného testu. Pro žáky, kteří v testu dosáhli nejlepších výsledků, byl ve fázi učení charakteristický vyšší počet tranzicí mezi čteným textem a ilustrací, ve srovnání s hůře skórujícími žáky.

Je tedy vidět, že eye-tracking má pro výzkum metakognice určitý potenciál a v budoucnu se snad dočkáme rozsáhlejších studií, které lépe objasní vztahy mezi určitými vzorci očních pohybů a konkrétními metakognitivními procesy. Velkou výhodou tohoto přístupu představuje především fakt, že se tak lze dostat k informacím, které jsou při využití alternativních metod hůře přístupné, případně i zcela nedostupné. Pro příklad můžeme použít zjištění zmíněná v předchozích dvou odstavcích – tedy vyšší počet regresí při čtení či častější tranzice mezi ilustrací a textem jako potenciální projev intenzivnějšího zapojení metakognitivních strategií. Jak již bylo zmíněno dříve, oční pohyby probíhají velmi rychle a nejsou pod vědomou kontrolou subjektu. Je tedy velmi nepravděpodobné, že by v pozadí stojící procesy mohly být vyjádřeny např. souběžnou verbalizací. Můžeme proto předpokládat, že při určitém experimentálním uspořádání odhalí eye-tracking efekty nezachytitelné prostřednictvím protokolové analýzy, a už vůbec ne prostřednictvím off-line dotazníkových metod. Zároveň je ale vhodné zmínit, že eye-tracking má jako nástroj pro zjišťování metakognice i četné nevýhody, které částečně i vysvětlují, proč nebyl za tímto účelem zatím používán častěji.

Za prvé, samotný výchozí předpoklad, tedy že charakter očních pohybů přesně odráží aktuálně probíhající duševní procesy sledovaného člověka (v eye-trackingové literatuře se tento předpoklad často označuje jako *eye-mind hypothesis*, doslovně přeloženo tedy „hypotéza oko-mysl“; Holmqvist et al., 2011), může být v mnoha experimentálních uspořádáních zacílených na studium metakognice dost problematický. Těsnou souvislost mezi očními pohyby a souběžně probíhajícími (meta)kognitivními procesy je možné opodstatněně předpokládat při úlohách spočívajících v bezprostředním získávání informací (např. čtením, prohlížením si grafického znázornění nějakého problému apod.). Pokud ale zkoumaná osoba při plnění úlohy přemýšlí, aktivně si vybavuje informace z dlouhodobé paměti nebo se snaží udržet informace v paměti pracovní, vazba mezi pohybem očí při pohledu na podnětové pole a probíhající duševní aktivitou může být velmi volná, případně až nulová.

Za druhé, samotný proces měření je z pohledu zkoumané osoby stále do určité míry rušivý – je nutné zachovávat pokud možno nehybnou polohu hlavy a celého těla ve zhruba konstantní vzdálenosti od monitoru, před samotným úkolem je nutná někdy i několikrát opakovaná kalibrace, aby byla zajištěna dostatečná přesnost měření atd. Navíc i přes nejlepší snahu se v záznamu téměř každého probanda objeví úseky, kdy aparatura nedokázala jeho oční pohyby registrovat, což vede nevyhnutelně ke zkreslení dat.

Za třetí, přesnost aparatury podmiňuje i podobu experimentálních úkolů – třeba u přístroje, který umožňuje použití mimo specializovanou laboratoř, např. přímo v domovské škole žáků, kteří se účastní výzkumu, je zpravidla nutné volit abnormálně veliké písmo, mají-li být různé charakteristiky očních pohybů registrovány a analyzovány na úrovni jednotlivých slov. Fakt, že experimentální podněty se pak výrazně liší od materiálů, s nimiž se žáci setkávají v běžné výuce, pak pochopitelně snižuje zobecnitelnost takto získaných závěrů.

OFF-LINE METODY

Dotazníky a posuzovací škály

Jak bylo zřejmé z popisu výše uvedených on-line metod, velmi často se jedná o postupy, které jsou poměrně náročné na administraci a/nebo na vyhodnocení. Dotazníkové metody představují v tomto určité protipól, jsou to nástroje velmi efektivní z pohledu „cena – výkon“, tedy poměrně snadno lze získat velké množství dat, alespoň teoreticky užitečných a vhodných k dalším analýzám. Nepřekvapí proto, že v zahraničí bylo různých dotazníků určených pro měření metakognice vyvinuto mnoho, přes několik desítek. Nedávno publikovaná přehledová studie Craigové et al. (2020) uvádí 36 jednotlivých dotazníkových metod, vytvořených mezi lety 1982 a 2018, které byly vyhledány pomocí systematické rešerše a dále hodnoceny. Ani tento počet není konečný, protože autoři se omezovali pouze na nástroje, jejichž vývoj či použití byly popsány v anglicky psané publikaci (dotazník samotný přitom mohl být, a v několika případech také byl, vyvinut a administrován v jiném než anglickém jazyce). Přestože některé z existujících dotazníků jsou známější a jsou používány celkově častěji než jiné, Craigová a kolegové zároveň konstatují, že žádná konkrétní metoda si nezískala takové uznání, aby se stala jakýmsi „oborovým standardem“ pro měření metakognice. Pro ilustraci si zde proto představíme tři vybrané metody, které lze z různých důvodů považovat za důležité či zajímavé.

Dotazník metakognitivního uvědomění (MAI, Metacognitive Awareness Inventory) je metodou, kterou lze označit přízviskem „tradiční“, tento nástroj vytvořili Schraw a Dennisonová již v roce 1994, používán je ovšem do současnosti (Harrison & Vallin, 2018). Autoři se při jeho tvorbě opírali o klasifikaci metakognitivních jevů podle Brownové (1987), kterou pro svoje potřeby v některých jednotlivostech modifikovali. Dotazník tedy obsahuje položky zaměřené jak na deklarativní formy metakognice (metakognitivní znalosti), tak na různé procedurální aspekty, které autoři dále člení do několika podkategorií (např. plánování, strategie pro práci s informacemi apod.). Příkladem „deklarativních“ položek jsou např. tvrzení „*Dokážu dobře posoudit, jak moc něčemu rozumím*“ či „*Vím, které rozumové schopnosti jsou mou silnou či slabou stránkou*“. Mezi „procedurální“ položky pak patří např. výroky: „*Než odpovím, zvážím několik alternativ řešení problému*“ (monitorování), nebo „*Když narazím na důležitou informaci, zvolím*“ (strategie pro práci s informacemi). V původním článku od Schrawa a Dennisonové

(1994) je v příloze uvedeno úplné znění všech položek, včetně vymezení, jaký konkrétní aspekt metakognice by měla ta která položka měřit. Dotazník MAI je určen pro dospělou populaci a standardizován byl na vzorku 197 mladých dospělých – vysokoškolských studentů nespécifikované univerzity na středozápadě USA. Následně ještě autoři ověřovali metodu na druhém vzorku, sestávajícím ze 110 studentů. Finální verze dotazníku sestává z 52 položek, odpověď na každou položku vyjadřuje posuzovaná osoba na grafické, stomilimetrové škále, jejíž levý konec vyjadřuje absolutní souhlas, pravý konec absolutní nesouhlas.

Vzhledem ke stáří původní metody stojí za zmínku nedávný výzkum, který provedli Harrison a Vallinová (2018), který je cenný také tím, že autoři pracovali s podstatně větším výzkumným vzorkem – 622 osob – než zahrnovala původní studie Schrawa a Dennisonové. Autoři také provedli jiné a v některých ohledech pokročilejší statistické analýzy (vycházející z IRT – teorie odpovědi na položku), než jaké byly realizovány při vytváření dotazníku, a došli na jejich základě k závěru, že kompletní verze o 52 položkách je psychometricky málo vyhovující. Jako alternativu navrhli zkrácenou variantu o 19 položkách, která zachovává 2 základní faktory (deklarativní a procedurální formy metakognice), ovšem již nikoli jednotlivé subškály, jako např. plánování, strategie pro práci s informacemi atd.

Na zkušenosti s dotazníkem MAI navázali Sperlingová a kolegové (2002) vytvořením upravené verze určené pro děti a žáky na počátku adolescence. Takto vzniklý nový dotazník nese název *Dotazník metakognitivního uvědomění pro mládež* (Junior Metacognitive Awareness Inventory, JMAI, případně i Jr. MAI). Dotazník byl navržen a standardizován pro poměrně široké věkové rozpětí, od 3. po 9. ročník základní školy. Protože v takto vymezeném období dochází k velkým vývojovým změnám, vytvořili autoři dvě separátní verze dotazníku JMAI: verze A je určena pro děti ve 3., 4. a 5. ročníku, verze B pro děti od 6. do 9. ročníku. Obě varianty se liší jednak způsobem odpovídání na položky – zatímco u verze A vybírá žák u každého výroku jednu ze 3 možných odpovědí (nikdy, občas nebo vždy), ve verzi B má odpověď na jednotlivé položky podobu pětistupňové Likertovy škály. Kromě toho se obě varianty liší také délkou, verze A sestává z 12 položek, ve verzi B jsou všechny tyto položky rovněž použity, kromě toho ovšem obsahuje ještě 6 dodatečných položek, u nichž se předpokládá, že by na ně mladší a méně zkušení žáci vesměs nedokázali smysluplně odpovědět. Jednotlivé položky byly převzaty z původního dotazníku MAI, v některých případech ovšem bylo jejich znění přeformulováno do podoby přístupnější dětem a mladším adolescentům (např. původní výrok z MAI „*Dokážu se motivovat k učení, když je to potřeba*“ byl pozměněn na „*Dokážu se přimět k učení, když je to potřeba*“).

Standardizační studie proběhla na dvou souborech – první čítal 344 žáků, druhý 416 žáků. JMAI zachovává rozdělení položek do dvou dimenzí (deklarativní – znalost kognice; procedurální – regulace kognice), vzhledem k výrazně redukované délce ovšem na rozdíl od dospělé verze již neobsahuje žádné subškály v rámci těchto dvou dimenzí. Autoři provedli v rámci standardizační studie i faktorovou analýzu, z níž vyplynulo, že kromě tohoto dvoufaktorového řešení lze strukturu dotazníku přijatelně popsat i pomocí pěti faktorů, stabilních

napříč oběma verzemi. S ohledem na možné interpretační nejasnosti ovšem nakonec pro praktické využití doporučují používat pouze zmíněné dvě škály.

Ve standardizační studii byla rovněž sledována souvislost mezi úrovní metakognice, vyjádřenou skórem v dotazníku JMAI, a školním výkonem. Toto srovnání bohužel odhalilo skutečnost, která problematizuje praktické využívání JMAI např. pro školní poradenství a podobné účely: korelace mezi oběma proměnnými je celkově slabá a s věkem navíc klesá, u starších žáků pracujících s verzí B byly zjištěné korelace nevýznamné a v některých případech i záporné. Youngová a Worrell (2018) potvrdili tato zjištění, tedy jak dvoufaktorou strukturu dotazníku, tak prakticky neexistující vztah mezi jeho výsledky a mírami studijního výkonu, v nedávné studii zaměřené cíleně na populaci nadaných žáků. Slabý vztah mezi naměřenou úrovní metakognice a výkonovými ukazateli je zřejmě obecnějším problémem metakognitivních dotazníků, jak bude ještě zmíněno níže.

Jako třetí příklad uveďme dotazník, který má původ ve frankofonním prostředí – jeho autory jsou odborníci působící částečně v Québecu, částečně ve francouzsky hovořící části Švýcarska. Jímí vytvořený *Dotazník metakognitivních znalostí a regulace mentální činnosti* (COⁿnaissances MÉ^tacognitives et de la Gestion de l'Activité Mentale; COMÉGAM) konceptuálně vychází, podobně jako MAI i JMAI, z klasifikace metakognitivních jevů podle Brownové a podobně jako MAI se pokouší oba základní faktory ještě dále strukturovat pomocí podrobnějších subškál. Strukturu dotazníku si zde s ohledem na omezený prostor již popisovat blíže nebudeme, kompletní znění položek, včetně jejich přiřazení k příslušným subškálám, je k dispozici ve statí popisující vývoj a použití metody (Richer et al., 2004).

V témže textu autoři zároveň popisují poměrně originální způsob využití svého nástroje, který stojí za zmínku a který by se dal zřejmě analogicky použít i v případě jiných dotazníků. Zatímco většina metakognitivních dotazníků se používá buď pro výzkumné účely, nebo při vzdělávací či poradenské činnosti přímo s žáky, autoři využili metodu COMÉGAM při vzdělávání studentů pedagogiky – budoucích učitelů. V rámci univerzitního kurzu studenti dotazník nejprve vyplnili a poté, aniž k němu dostali k dispozici jakékoli další informace, měli za úkol pokusit se formou práce v malých skupinkách odhalit jeho zaměření, vnitřní strukturu (tj. jaké dílčí faktory nástroj měří) a svoje vlastní výsledky. Teprve následně autoři studentům metodu představili a poskytli jim klíč k jejímu vyhodnocení. Studenti měli tuto zkušenost, tedy například nakolik se jim podařilo předem rozpoznat měřené dimenze metakognitivních znalostí a dovedností, následně reflektovat a takto získané povědomí o metakognici pak využít při navazující pedagogické praxi, kde již pod supervizí sami vyučovali žáky (Richer et al., 2004)⁸.

⁸ Některé další zahraniční metakognitivní dotazníky, které zde nejsou zmíněny, popisuje v české odborné literatuře Lokajčková (2014).

Craigová et al. (2020) docházejí v již zmíněné přehledové studii k některým obecnějším závěrům, které již byly naznačeny konkrétně v souvislosti s dotazníky MAI a JMAI. Za prvé se zdá, že základní hrubá klasifikace metakognitivních jevů na deklarativní a procedurální formy metakognice je dobře opodstatněná a že toto rozdělení dokáže až na výjimky zachytit většina vytvořených dotazníků v podobě dvou oddělených faktorů, které mohou být nezávisle skórovány. Jako problematická se naproti tomu ukázala snaha mapovat pomocí dotazníkových metod strukturu metakognice na jemnější, detailnější úrovni. Mnohé metody byly původně navrženy tak, aby vedle základního dělení na dvě dimenze obsahovaly v rámci každé z nich ještě další, podrobnější subškály (zaměřené např. na jednotlivé typy metakognitivních znalostí, na jednotlivé druhy strategií atp.). Následné analýzy ovšem často nedokázaly tuto předpokládanou strukturu potvrdit. Ještě větší problém představuje skutečnost, že faktorová struktura se často liší u téhož nástroje napříč různými studii a výzkumnými vzorky. Například v případě dotazníku MAI, který byl výzkumně používán relativně nejčastěji, došli různí autoři k faktorovým řešením, sestávajícím ze tří, čtyř, pěti, šesti či osmi komponent.

Tyto nesrovnalosti vedly Craigovou a kolegy (2020) k závěru, že současné metakognitivní dotazníky zřejmě nejsou schopny odhalit komplexnější vztahy mezi metakognitivními jevy prostřednictvím používaných subškál. To může mít řadu různých příčin. Za prvé je možné, že lidé obecně, nebo přinejmenším děti a mladší adolescenti, nejsou na tak detailní úrovni svoje metakognitivní charakteristiky vůbec schopni spolehlivě posoudit. Druhé vysvětlení může souviset s relativně malou délkou používaných subškál; pokud si má metakognitivní dotazník zachovat nějakou „rozumnou“ délku a zároveň má obsahovat šest či osm subškál, musí zákonitě každá jednotlivá subškála sestávat pouze z několika položek, což snižuje její reliabilitu. Určitým řešením do budoucna by tak mohlo být vytváření úžeji zaměřených, specializovaných dotazníků určených pro posuzování pouze určitých dimenzí metakognice. Teoreticky by například neměl být problém vyvinout metodu, která bude zcela opomíjet oblast metakognitivních znalostí a soustředí se pouze na metakognitivní strategie, přičemž každá jednotlivá strategie bude posuzována subškálou, která bude sestávat z relativně většího množství různorodějších položek.

Třetí vysvětlení nabízí hypotéza, kterou navrhla Ningová (2016). Ta tvrdí, že posoudit vlastní metakognitivní schopnosti dostatečně detailním způsobem závisí na celkové úrovni rozumových předpokladů. Posuzování metakognice prostřednictvím podrobnějších subškál tedy nebude fungovat u relativně malých dětí, ale jeho spolehlivost bude narůstat současně s věkem. Zároveň je ale nutné brát v úvahu i diferenciaci schopností v jednotlivých věkových kohortách – mimořádně nadané dítě tedy bude pravděpodobně schopnější posoudit svoji metakognici na úrovni detailních subškál spíše než chronologicky starší, ale podprůměrný adolescent apod. Přístup Ningové, v němž je u participantů zohledněna také celková úroveň jejich kognitivních schopností, je zatím spíše ojedinělý (Craig et al., 2020) a k jeho potvrzení by nepochybně přispěly další výzkumy zaměřené cíleně na metakognici u nadaných. Podobných studií, jako je např. již dříve zmíněná práce Youngové a Worrella (2018), vychází

však v poslední době málo a tato konkrétní práce příliš jasno do celé věci nevnese vzhledem k tomu, že v ní byl použit JMAI, který byl od počátku navržen jako dvoufaktorový a celkově jednoduchý nástroj.

Dalším důležitým zjištěním, ke kterému dospěli Craigová se spolupracovníky (2020) a které je v souladu i se závěry prací dalších autorů, je fakt, že výsledky (dosažené skóry) v metakognitivních dotaznících nemají zpravidla téměř žádnou souvislost (ve smyslu naměřených korelací) s indikátory metakognitivní aktivity získanými prostřednictvím on-line metod. Vzhledem k tomu, že on-line metody sice také neumožňují přímou detekci metakognitivních procesů, alespoň však projevy těchto procesů registrují bezprostředně, představuje tento nesoulad vážný problém pro validitu metakognitivních dotazníků. Veenman (2012, s. 28) to charakterizuje takto: „Zdá se, že lidé nedělají to, co říkají, že provedou, ani si přesně nevzpomínají na to, co skutečně udělali“ (míněno při řešení úloh, které vyžadují zapojení metakognitivních dovedností).

I pro tento nesoulad bylo navrženo několik možných a realistických vysvětlení. Jedním z nich je jednoduše omezená kapacita jak pracovní, tak dlouhodobé paměti. Toto omezení by vedlo ke zkreslení a nepřesnostem při následných výpovědích posuzované osoby samo o sobě. Často navíc může docházet k tomu, že formulace položek v metakognitivních dotaznících navozují v probandech dodatečně falešný pocit známosti a vedou je k přesvědčení, že si vzpomínají na vykonání určité metakognitivní aktivity, kterou ovšem ve skutečnosti nepoužili. Roli dále může hrát značná obecnost ve znění mnoha dotazníkových položek a překvapivě také efekt určité sociální žádoucnosti: žák, který vyplňuje dotazník, chápe, že určité odpovědi na předložené otázky naznačují jeho aktivní, cílevědomý přístup k učení a že tento přístup je ze strany pedagogů, rodičů či dalších dospělých obvykle oceňován. To u něj může posílit tendenci odpovídat způsobem, který v souhrnu vede k vysokým metakognitivním skórum bez ohledu na to, jak při učení či řešení úloh postupuje ve skutečnosti. Ať už je správné kterékoli z těchto vysvětlení, případně ještě úplně jiné, představuje problematická validita metakognitivních dotazníků zásadní nevýhodu jak ve výzkumu, tak zejména při jejich použití v rámci pedagogického poradenství.

Metakognitivní interview

Vedle klasických „papírových“ dotazníků jsou pro zjišťování metakognice v menší míře používány i další sebe-posuzovací metody, které mají zpravidla podobu standardních strukturovaných či polostrukturovaných rozhovorů (Gascoine et al., 2017). Oproti dotazníkům chybí těmto postupům výhoda snadné a rychlé hromadné administrace, alespoň teoreticky ovšem mají zároveň několik předností, které mohou tento nedostatek vykompenzovat. Za prvé může rozhovor tvář v tvář navodit kvalitnější raport mezi posuzovanou osobu a pedagogem či psychologem a motivovat tak particpanta k promyšlenějším a hlubším odpovědím. Za druhé může být interview použito i u relativně mladších dětí, které by vzhledem k aktuální úrovni

ni čtení a vlastního psaného projevu ještě nebyly s klasickým dotazníkem schopny pracovat (viz např. Alexander et al., 2006). Třetí potenciální výhoda pak souvisí s výše zmíněným problémem paměťového zkreslení. Jestliže formulace otázky může v probandech vyvolávat falešnou vzpomínku na mentální děje, které u nich ve skutečnosti neproběhly, bylo by žádoucí volit otázky méně konkrétní, pokud možno s otevřenou odpovědí. U klasického dotazníku by to s sebou ovšem přinášelo riziko nejasných, obtížně skórovatelných odpovědí. Metakognitivní interview umožňuje examinátorovi, aby na podobné nejednoznačné odpovědi zareagoval bezprostředně a ujistil se o tom, jak přesně byly myšleny (Schmitt & Sha, 2009). U volných odpovědí je pak obecně kladen větší důraz na úsudek examinátora, což by mělo spolehlivosti celé procedury teoreticky prospět vzhledem k tomu, že examinátor může být v tomto ohledu vyškolen, po určitém množství administrací se může opírat o svou zkušenost atd. Jedná se ovšem pouze o předpoklad, pokud je autorovi známo, žádná rozsáhlejší studie, která by cíleně porovnávala validitu dotazníkových a rozhovorových metod pro zjišťování metakognice nebyla ani v zahraničí dosud provedena (asi nejbližše se tomuto cíli blíží studie Gascoineové a spolupracovníků z roku 2017, která má ovšem širší a obecnější záběr).

S přihlédnutím k tomu, že se jedná o relativně méně častý přístup ke zjišťování metakognice, představíme si ilustračně pouze jeden příklad konkrétní metody tohoto druhu. *Interview o metakognitivním povědomí* (Interview about metacognitive awareness, IMA) vytvořily autorky Schmittová a Sha (2009). Jedná se o krátký nástroj, sestávající pouze z 8 standardních otázek, zacílených na využívání metakognice v procesu čtení. Probandi odpovídají na každou položku volnou odpovědí, examinátor následně při vyhodnocení přiděluje jednotlivým odpovědím 0, 1 nebo 2 body podle toho, nakolik pokročilou úroveň metakognice podle jeho úsudku odpověď vyjadřuje. Autorky uvádějí i konkrétní příklady skórování. Například u položky „Co obvykle uděláš před tím, než začneš číst nový příběh nebo knihu“ skórují 0 bodů odpovědi „Podívám se, jestli je to zajímavé“, nebo „Přečtu si konec“. Odpověď „Zjistím, o čem to je“ je hodnocena jedním bodem. Dvěma body je pak hodnocena odpověď „Podívám se na nápis a zkusím odhadnout, co se v příběhu stane“.

Metoda stimulovaného vybavení

Jako poslední si představíme off-line metodu, která není založena na sebe-posouzení jedince a která se tak od předchozích dvou poněkud liší. *Metoda stimulovaného vybavení* (stimulated recall) je poměrně málo známým a málo používaným postupem, který se pokouší obejít některé nevýhody klasické protokolové analýzy (O'Brien, 1993; Edwards-Leis, 2006). Propagátoři metody stimulovaného vybavení považují snímání protokolu myšlení nahlas za příliš invazivní postup a domnívají se, že zejména u dětí to vede ke zkresleným a nespolehlivým výsledkům. Alternativa, kterou nabízejí, spočívá v tom, že posuzovaná osoba je v průběhu řešení určité úlohy (při níž po ní verbalizaci myšlení nevyžadujeme) nahrávána na video a teprve při následném přehrání videozáznamu popisuje posuzovaná osoba – která záznam sleduje společně s examinátorem – postup svého myšlení, k jehož vybavení by jí nahrávka

měla být nápomocná. Proceduru lze různými způsoby upravovat, v klasickém pojetí, které popsal a doporučil O'Brien, má proband plnou kontrolu nad přehráváním – zejména může nahrávku zastavit, kdykoli potřebuje více času na to, aby popsal, jak a o čem v okamžiku zachyceném na nahrávce přemýšlel. Probandi jsou také na začátku druhé (přehrávací) fáze instruováni, aby svoje myšlenkové postupy popisovali samostatně a spontánně, examinátor ovšem může klást v případě nutnosti otázky. Ty by měly být pokud možno co nejméně specifické (např. „*O čem jsi právě teď přemýšlel?*“), je potřebné vyvarovat se otázek hodnotících či návodných (např. „*Tady jsi se usmál, co tě v tu chvíli napadlo?*“).

I z tohoto letmého popisu je zjevné, že metoda stimulovaného vybavení je reálně použitelná pouze u omezeného typu podnětových úkolů. Pokud má např. žák vyřešit matematický problém, což typicky zahrnuje velké množství dílčích kroků, mezivýpočtů či poznámek, jejichž zápis na papír může být nahráván, lze následné stimulované vybavení provést relativně snadno. Naopak metoda nebude příliš uplatnitelná při činnosti, která zahrnuje například pouze déletrvající učení se ze čteného textu. Oproti protokolové analýze má ovšem tento přístup jednu výhodu navíc – je možné jej využít i při skupinové práci více žáků (příčemž oddálené vybavení se poté provede s každým zvlášť). Lamová (2008) např. takto zkoumala metakognitivní strategie při skupinové výuce cizího jazyka.

Vzhledem k tomu, že technika stimulovaného vybavení představovala zatím spíše okrajový přístup, nebyla pro ni ani vypracována nějaká komplexnější metodologie. Pokud jde o podrobnější klasifikaci metakognitivních jevů, mělo by být v zásadě možné bez nutnosti větších změn použít schémata vypracovaná pro protokolovou analýzu. Podrobnější srovnání obou metod ovšem i v zahraniční literatuře chybí.

SPECIFIKA MĚŘENÍ METAKOGNITIVNÍHO MONITOROVÁNÍ OBTÍŽNOSTI ÚLOHY

Zbývající část této publikace bude věnována specifickým postupům, které se uplatňují při diagnostice metakognitivního monitorování v užším smyslu – tj. při subjektivním odhadu obtížnosti úlohy, případně subjektivním odhadu vlastní úspěšnosti při řešení úlohy. U dosud představených metod hrál kognitivní úkol, použitý k vyvolání a následnému zaznamenání metakognitivní aktivity, pouze převážně instrumentální roli, i když jak jsme viděli, některé metody – např. IMMEX – umožňují i souběžnou analýzu úspěšnosti v samotné kognitivní úloze. V případě studia metakognitivního monitorování je právě (ne)úspěšnost v řešení samotné kognitivní úlohy a její vztah k subjektivnímu metakognitivnímu odhadu⁹ zásadní. Prakticky se postupuje tak, že posuzované osobě zadáme určitou sadu problémových úloh – např. znalostní či didaktický test, přičemž u těchto úloh (položek testu) je možné jednoznačně určit, zda byly či nebyly vyřešeny správně. Zároveň posuzovanou osobu žádáme, aby stanoveným způsobem odhadla, nakolik si je správností svého řešení jistá, případně nakolik jí položka přišla snadná/obtížná. Stupeň souladu mezi tímto subjektivním odhadem a reálně dosaženými výsledky představuje hledanou míru kvality metakognitivního monitorování.

⁹ Pokud proband posuzuje míru své jistoty, že určitý úkoly vyřešil správně, a vyjádří tuto informaci na nějaké číselné škále, používá se pro tento indikátor v anglofonní odborné literatuře obvykle termín „confidence rating“. Protože doslovný překlad, tedy spojení typu „ohodnocení sebedůvěry“ či „rating sebedůvěry“ by působil v češtině poněkud neohrabaně a mohl by být i významově matoucí, budeme v dalším textu pro tento indikátor používat termín „odhad úspěšnosti“ – míněn je tedy subjektivní odhad probanda, nakolik byl při řešení konkrétní úlohy/položky úspěšný. Termíny „metakognitivní odhad“ či „metakognitivní hodnocení“ budou používány jako obecnější označení, které kromě výše uvedeného zahrnuje i případy, kdy proband například pouze hodnotí (opět zpravidla na několikabodové škále) subjektivně vnímanou obtížnost úlohy, bez explicitního posouzení vlastního výkonu.

Z praktického hlediska je velkou výhodou tohoto typu diagnostiky především fakt, že ji lze provozovat s využitím různých znalostních, výkonových či podobných testů, které lze pro tento účel rychle a lehce upravit a které jsou většinou snadno dostupné. Tyto nástroje přitom původně vůbec nemusely být určeny k měření metakognice (v našich podmínkách lze např. uvažovat o využití starších a veřejnosti zpřístupněných testů státní části maturitních zkoušek apod.). Nutná úprava obvykle spočívá pouze v doplnění prvků, jejichž prostřednictvím proband vyjadřuje svůj subjektivní odhad, např. ve formě několikabodové číselné škály, binárního hodnocení ANO/NE, SNADNÉ/OBTÍŽNÉ, či grafického prvku – úsečky, která umožňuje odstupňovat odpověď vyznačením místa v různé vzdálenosti od krajních bodů a podobně. Tyto prvky mohou být doplněny buď za každou položkou, případně za určitými logicky ucelenými sadami položek (jeden subtest apod.), nebo pouze souhrnně na jednom místě na závěr celého testu.

Zároveň má tento druh diagnostiky pochopitelně i některé významné nevýhody. Za prvé pro něj zdaleka nelze využít jakýkoli typ podnětových kognitivních úloh, použité úlohy musí mít, jak už bylo zmíněno, jednoznačně definované správné řešení a pro výpočet řady indexů je navíc nutné, aby byla celá komplexní úloha rozdělena do řady dílčích, samostatných úkolů (což dobře splní např. právě položky v testu, ale již ne třeba slohová práce). Tím je ovšem ze studia tohoto aspektu metakognice vyloučeno mnoho typů kognitivních problémů, které jsou přitom zjevně významné, a přítomnost metakognitivní aktivity se u nich dá opodstatněně předpokládat (např. různé problémy zaměřené na řešení vhladem, nestrukturovaná zadání vyžadující od probanda samotnou identifikaci problému apod.). Za druhé, a to částečně souvisí i s předchozím bodem, tento druh diagnostiky nepostihuje metakognici (a dokonce ani užší kategorii metakognitivních dovedností) v celé její šíři, neposkytuje například přímé a nezprostředkované informace o využívání metakognitivních strategií. Je proto poněkud zavádějící, pokud je v určitém výzkumu použita pouze nějaká metrika související s metakognitivním monitorováním, ale v závěrech studie se hovoří obecně o „metakognici“. S tímto terminologickým zmatením se lze v poslední době často setkat zejména v neurovědním výzkumu (viz např. Fleming & Frith, 2014).

Přestože příprava podnětových materiálů a jejich administrace může být poměrně jednoduchá, totéž již nelze říci o vyhodnocení a interpretaci získaných výsledků. Data vztahující se k metakognitivnímu monitorování mohou být kvantifikována řadou různých indexů – některé z těchto indexů lze ovšem dobře využít i jinde než v metakognitivním výzkumu, a proto by nemělo překvapovat, že byly navrženy a popsány dříve, než Flavell (1979) publikoval svoji stěžejní práci. Například koeficient γ , který se následně výrazně uplatnil v metakognitivně orientovaných studiích, navrhli původně Goodman a Kruskal již v polovině 50. let 20. století (Schraw, 2009). Velký pokrok v této oblasti pak nastal zejména v 70. a 80. letech minulého století a vzhledem k tomu, že odvození jednotlivých indexů se zakládá na obecně platných matematických vztazích, lze z většiny těchto starších prací čerpat i dnes. Přesto se ani v této oblasti vývoj nezastavil a zejména v průběhu uplynulé dekády bylo navrženo a formálně odvozeno – především s ohledem na potřeby neurovědní

orientovaného metakognitivního výzkumu – několik nových¹⁰ metrik, které částečně řeší dlouhodobě známé nedostatky starších a „tradičních“ indexů.



Registrovat a analyzovat projevy metakognitivního monitorování lze mnoha různými způsoby. Tento na první pohled poněkud zmatený obraz se pokusíme zpřehlednit uspořádáním jednotlivých možností do 4 základních dimenzí: a) úroveň, na které žádáme od probanda podání metakognitivního hodnocení (globální nebo lokální); b) časové uspořádání mezi podáním metakognitivního hodnocení ze strany probanda a vlastní kognitivní úlohou; c) konkrétní druh použité metriky; d) skutečnost, zda sledujeme absolutní či relativní přesnost metakognitivního hodnocení. Tyto dimenze si nyní představíme podrobněji.

I. Úroveň měření. Po posuzované osobě můžeme žádat, aby odhadla svůj výkon v testu jako celku – to označujeme jako *globální odhad* či *globální kalibraci* (Hacker et al., 2008). V nejjednodušším případě se například jednotlivých žáků po napsání písemné práce dotazujeme, jakou známku se domnívají, že z práce dostanou, a porovnáváme tyto odhady s následnými známkami skutečně udělenými učitelem. Alternativně zadáme znalostní test, sestávající třeba z deseti otázek, jejichž správnost lze zcela jednoznačně posoudit (čímž odpadá určitá subjektivita na straně učitele, která by se mohla projevit např. u písemné práce s volně tvořenými odpověďmi), za každou správně zodpovězenou otázku se přiděluje 1 bod. Žák tedy ví, že v testu jako celku mohl dosáhnout výsledku v rozmezí 0–10 bodů. Po dopsání testu a před tím, než jej dostane zpět opravený, necháme žáka napsat, kolik bodů si myslí, že dosáhl.

Pokud je metakognitivní odhad prováděn na *lokální* úrovni, vyjadřuje se proband zvláště ke každé jednotlivé úloze (či testové položce) – např. bezprostředně po jejím zodpovězení uvede, s jakou pravděpodobností se domnívá, že je jeho odpověď správná. I tyto lokální odhady je pak ovšem možné shrnout do nějakého celkového skóru za celý test, „lokalita“ je podstatná z pohledu toho, kde a kdy respondent odpovídá, nikoli nutně z hlediska následné analýzy. Výhodou globální kalibrace je pochopitelně především rychlost a snadnost měření, typicky žádáme od posuzované osoby jeden údaj na konci testu, což prakticky nijak nenavýší délku administrace a nemělo by to představovat ani žádnou dodatečnou zátěž pro probanda. Naproti tomu, pokud použijeme lokální odhady obtížnosti v testu, který sestává třeba z několika desítek položek, časová zátěž a náročnost na pozornost participantů mohou být značné.

Z tohoto hlediska je diskutabilní zejména použití lokálních odhadů obtížnosti u inteligenčních testů (nebo obecně jakýchkoli testů, které umožňují srovnání výsledků probanda s nějakou populační normou), které byly standardizovány pro přesně danou proceduru vyšetření, jejíž součástí takové mezikroky pochopitelně nejsou. Obecně lze takový postup doporučit pouze

¹⁰ I tyto metriky typu AUROC2, meta-d' apod., ovšem konceptuálně navazují na práce, které se datují až do 50. let 20. století.

experimentálně, s vědomím, že dosažená hodnota IQ může být zkreslena odchýlením se od standardní procedury. Výhledově si lze ale představit i standardizaci určitého inteligenčního testu speciálně pro toto použití – tj. testová procedura by obsahovala lokální odhady obtížnosti i při administraci standardizačnímu vzorku, na základě jehož výsledků jsou následně vytvořeny normy. V současnosti se ovšem tento postup běžně nepoužívá – autorovi této práce není znám příklad žádného rozšířeného a takto upraveného inteligenčního testu, a to ani u nás, ani v zahraničí.

Přes všechny nevýhody poskytuje nicméně lokální kalibrace zřejmě přesnější odhad metakognitivního monitorování. Vzhledem k tomu, že globální odhad obtížnosti provádí proband až s určitým odstupem po vyřešení konkrétních položek, lze jej chápat jako svého druhu off-line metodu, se všemi nedostatky, které vyplývají zejména z omezené paměťové kapacity a které byly v souvislosti s off-line metodami uvedeny dříve. Také některé empirické studie prováděné s nadanými studenty dospěly k závěrům, které svědčí pro větší spolehlivost registrování projevů metakognitivního monitorování na lokální úrovni (Snyder et al., 2011).

II. Časové uspořádání. Rozlišujeme odhady *prediktivní, souběžné (konkurentní) a postdiktivní* (Schraw, 2009). V případě prediktivních odhadů se ptáme před zahájením samotné kognitivní aktivity, např. žákům rozdáme test, necháme je prohlédnout si jednotlivé otázky, a na odhad celkového výsledku v testu se jich zeptáme ještě před tím, než na otázky začnou odpovídat. Jako souběžný zpravidla označujeme takový druh lokálního odhadu, kdy participant v průběhu testu podává odhad bezprostředně po zodpovězení každé položky¹¹. Postdiktivní odhad získáme tak, že proband hodnotí svůj výkon poté, co dokončil celý test. Jak vyplývá z již několikrát zmíněných teoretických východisek, obecně se dá předpokládat, že nejpřesnější údaje získáme za použití souběžných odhadů, a ty je proto vhodné používat přednostně všude, kde je to možné.

III. Konkrétní typ metakognitivní metriky. Všechny míry metakognitivního monitorování spojuje to, že jsou odvozeny od odhadu, kterým posuzovaná osoba hodnotí subjektivní obtížnost řešené úlohy. To je možné realizovat několika způsoby (např. procentuální odhad úspěšnosti, posouzení obtížnosti na nějaké ratingové škále atd.), s výslednými metrikami lze ale následně analyticky pracovat velmi podobně – i když je pravda, že některé indexy lze korektně spočítat například pouze na základě určitého typu dat (např. procentuální hodno-

¹¹ Zde jsou na místě dvě poznámky. Za prvé, striktně vzato je i takováto odpověď vlastně postdiktivní, protože testovaná osoba odpovídá až po vyřešení položky. Vzhledem k minimální časové prodlevě se ovšem na tento fakt nebere zřetel a „skutečně souběžná“ hodnocení, kdy by například participant verbalizoval svůj odhad prostřednictvím myšlení nahlas v průběhu řešení konkrétní položky, se pro tento účel běžně nepoužívají. Za druhé, samotné použití lokální kalibrace nemusí nutně vyžadovat souběžný způsob hodnocení. Lze si představit design, kdy proband nejprve vyplní test ve „standardním“ režimu a následně s ním examinátor položky znovu prochází a u každé žádá metakognitivní odhad. V tomto případě už by se nepochybně jednalo o postdiktivní hodnocení v pravém slova smyslu. I s takovýmto designem se ovšem setkáváme velmi zřídka.



cení, ale již ne bodový ratingový odhad a podobně). Pro ilustraci si zde přiblížíme tři důležité metriky, které se všechny používají při měření na lokální úrovni vzhledem k tomu, že práce s globálními odhady je zpravidla triviální (často je výstupem pouze stanovení rozdílu mezi globálním odhadem a reálným výsledkem v testu či sadě úloh).

- a) **Procentuální odhad správnosti vyřešení (kalibrace správnosti)**: tato metrika je svým způsobem nejpřímochařejší – proband po každé zodpovězené položce vyjadřuje subjektivní míru jistoty prostřednictvím odhadu pravděpodobnosti, že jeho řešení je správné. Zjevnou výhodou této metriky je právě její jednoznačnost a srozumitelnost, ať už v zadávání instrukcí probandům, či při následné interpretaci. Velkou nevýhodou představuje naopak skutečnost, že tento typ odhadu můžeme žádat pouze od dostatečně starých a dostatečně inteligentních probandů, kteří chápou pojem pravděpodobnosti a dokáží s ním bez potíží samostatně pracovat. Získání tohoto typu údajů je proto prakticky vyloučeno například u dětí na prvním stupni ZŠ. U starších žáků (nebo i u dospělých osob) zase může vést užití této metriky k nežádoucím výzkumným artefaktům vyplývajícím z toho, že část posuzovaných osob s pravděpodobnostními odhady dokáže korektně pracovat a část ne.
- b) **Odhad snadnosti řešení (Ease of solution judgement; EoSJ)**: v případech, kde nelze použít procentuální odhad obtížnosti, nebo by s sebou jeho použití neslo příliš velké riziko, nabízí se jako alternativa *odhad obtížnosti řešení* (dále jen EoSJ)¹². V tomto případě žádáme od posuzované osoby, aby po vyřešení samotné úlohy ohodnotila její subjektivně vnímanou obtížnost, zpravidla na určité formě Likertovy škály (např. 0 = nejjednodušší, 10 = nejtěžší). Kromě podobných diskrétních hodnot lze zejména při počítačové administraci využít i různé myši ovládané posuvníky, a získat tak veličinu, která je prakticky spojitá. Naopak je také možné žádat o odpovědi pouze v základních kategoriích, např. „Myslím si, že jsem úlohu správně vyřešil/a: „ANO vs. NE“, případně „Úloha se mi zdála: SNADNÁ vs. OBTÍŽNÁ“. Ve všech případech je pak, podobně jako u kalibrace správnosti, kritériem skutečnost, zda proband položku samostatně vyřešil správně či nikoli.

Na rozdíl od procentuálního odhadu je EoSJ celkově méně jednoznačnou metrikou, a to i pro probandy v průběhu testování. To vyplývá z faktu, že pokud již někdo rozumí pojmu pravděpodobnosti, tak tvrzení že dané řešení považuje správné s pravděpodobností 0 % či 20 % má srozumitelný a jasný význam. Naproti tomu různí probandi si mohou pod označením „maximálně jednoduchá položka“ představit pokaždé něco trochu jiného. Nutná je proto dostatečně podrobná instruktáž před začátkem experimentu či vyšetření a obecně je zřejmě vhodné upřednostnit procentuální kalibraci před EoSJ, pokud je to možné.

¹² V anglosaské literatuře se lze setkat i s jinými označeními a z nich odvozenými zkratkami, např. judgement of difficulty – JOD (Vangsness & Young, 2021).

- c) **Pocit znalosti (feeling of knowing, FOK):** pocit znalosti (dále jen FOK) vyjadřuje metakognitivní zkušenost, při níž má subjekt pocit, že určitou informaci se již v minulosti naučil, a ačkoli si ji nyní nedokáže vybavit, její fragmenty stále přetrvávají v jeho paměti a k vybavení by mu proto mohla pomoci vhodná nápověda či vodítko. Míru FOK lze vyjádřit několika způsoby, které se velmi podobají tomu, s čím jsme se již setkali v předcházejících případech. Možný je tedy jak procentuální odhad (zde subjekt udává míru jistoty, že se s danou informací již setkal a v minulosti ji znal), tak ohodnocení intenzity FOK na Likertově škále, stejně jako zařazení úlohy do jedné ze dvou či tří kategorií (např. „To jsem asi už někde četl“, „Nejsem si jistý“, „S tím jsem se zřejmě zatím nikdy neseťkal“). FOK je specifický především stanovením kritéria – zpravidla nepracujeme se správností/nesprávností samostatného řešení probanda, ale buď s jeho schopností vybavit si danou informaci až po poskytnutí nápovědy, nebo přímo s kontrolovanou evidencí toho, zda se proband s předmětnou informací v minulosti setkal či neseťkal.

Druhá možnost se prakticky uplatní pouze v experimentálních designech. Příkladem může být, když probandům prezentujeme při prvním sezení určitou sadu fotografií a po několika dnech jim předkládáme novou sadu, jejíž část ovšem tvoří i dříve prezentované fotografie. Instruueme je přitom, aby u každé fotografie uvedli, zda ji již dříve viděli či nikoli. V takovém případě může experimentátor předchozí expozici plně kontrolovat – eviduje, které fotografie předložil a které ne, a pokud použije obrázky cíleně vytvořené pro experiment, je prakticky vyloučeno, aby se s nimi testovaná osoba setkala někde jinde. Pokud studujeme FOK vázaný na informace, které lidé průběžně nabývají v běžném životě, je tento druh kontroly nereálný.

V případě poskytnutí vodítka může praktický postup vypadat zhruba takto: žák se musí předem naučit seznam slovíček v cizím jazyce. S odstupem času je testováno jejich paměťové vybavení – žákovi jsou předkládána česká slova a on má uvést jejich – kupříkladu anglický – překlad. Pokud odpověď zná, položka se s ohledem na FOK vůbec nehodnotí, pokud si nicméně dané anglické slovíčko nevybaví, ale zažívá při pokusu o jeho vybavení pocit „mám to na jazyku“¹³, uvede pro tuto položku vysokou hodnotu FOK (např. si je jistý na 80 %, že slovo se již dříve naučil), a naopak. Následně je žákovi poskytnuta nějaká forma nápovědy, např. počáteční hláska, počet a rytmus slabik apod., pokud si po ní cílové slovo vybaví, skóruje se kritérium jako správná odpověď (úspěch), v opačném případě jako odpověď chybná (neúspěšná). V některých designech lze skórování kritéria dále zjemnit tím, že nápověda může být odstupňovaná (v prvním kroku se poskytne první hláska, v druhém kroku poslední hláska atp.), kritérium je pak skórováno vícebodově a za každý krok navíc, který byl nutný pro konečné vybavení, se udělí méně bodů.

¹³ V anglosaské psycholingvistické literatuře se fráze „tip-of-the-tongue phenomenon“, tedy doslova „fenomén mám to na jazyku“, používá jako standardní odborný termín pro označení těchto jevů.

IV. Absolutní versus relativní přesnost. Zřejmě nejdůležitější dimenzí, která je ovšem zároveň poměrně neintuitivní, je to, zda nás při zjišťování metakognitivního odhadu zajímá jeho absolutní či relativní přesnost. Logika, která stojí za tímto rozdílem, se neomezuje pouze na zjišťování metakognitivních jevů a pro lepší názornost bude výhodnější přiblížit si ji nejprve na jednodušším modelu. Tím bude subjektivní posuzování fyzické charakteristiky – v tomto případě tedy hmotnosti – určitých předmětů¹⁴. Uvažujme dva posuzovatele, Janu a Tomáše. Úkolem každého z nich je postupně potěžit sadu pěti železných závaží o různé hmotnosti a pro každé konkrétní závaží se pokusit jeho hmotnost odhadnout. Experimentátor má samozřejmě k dispozici přesné údaje o skutečné hmotnosti jednotlivých závaží a může s nimi odhady pokusných osob porovnávat. Výsledky tohoto fiktivního experimentu shrnuje Tabulka 2.

Tabulka 2.

Odhady fyzikální charakteristiky (hmotnosti) sady závaží u dvou osob a skutečná hmotnost posuzovaných závaží

Závaží	Odhad Jany [kg]	Odhad Tomáše [kg]	Skutečná hmotnost [kg]
A	3	8	9
B	2	9	7
C	4	6	8
D	6	11	12
E	5	11	10

Na základě podobně strukturovaných dat lze spočítat několik různých indexů, které si nicméně představíme později (již přímo ve vztahu k metakognitivnímu hodnocení) a zde se jimi nebudeme zabývat. Nyní si vystačíme s prostým pohledem do tabulky, ze kterého je patrné – s případnou pomocí několika triviálních výpočtů – že způsob, jakým se odhady obou probandů vztahují ke skutečně naměřenými hodnotám, se poměrně výrazně liší. Porovnáme-li například aritmetický průměr odhadů Jany, Tomáše a reálných hmotností, vidíme, že průměr Tomášových odhadů (9 kg) se mnohem víc blíží průměru skutečných hodnot (9,2 kg), než průměr odhadů Jany (4 kg). Na úrovni jednotlivých položek, resp. v tomto případě jednotlivých závaží, se tato tendence manifestuje tak, že absolutní hodnota rozdílu mezi odhadem probanda a skutečnou hmotností závaží je obecně menší u Tomáše, který se v našem případě

¹⁴ Dnes může činnost, jakou si popíšeme v tomto příkladu, působit zcela absurdně a samoučelně, je ale faktem, že v počátcích psychologie jakožto akademické disciplíny na konci 19. století se v tehdejších psychologických laboratořích podobné experimenty prováděly relativně často.

odchyluje od skutečné hodnoty u jednoho závaží (v tomto případě jde o závaží B či C) maximálně o 2 kilogramy, než u Jany, která se odchyluje až o 6 kilogramů (a stojí za povšimnutí, že i *minimální* Janina odchylka – 4 kg v případě závaží C, je stále ještě větší, než Tomášovo *maximum*). Tento druh korespondence mezi odhadem a skutečnou hodnotou označujeme jako *absolutní přesnost*. Vyšší absolutní přesnosti dosahuje v našem případě Tomáš.

Zároveň je ale z bližšího pohledu do Tabulky 2 patrná i další důležitá charakteristika odpovědí: Janiny odhady hmotnosti jednotlivých závaží se sice od skutečných hodnot v absolutním smyslu výrazně liší, zároveň ale platí, že pokud je určité závaží reálně lehčí než jiné, je i Janin odhad hmotnosti tohoto závaží menší, než odhad hmotnosti toho druhého. V našem konkrétním příkladu, který je fiktivní a pro lepší názornost i poněkud extrémní, zhruba odpovídají i vzájemné poměry mezi skutečnými hmotnostmi a jejich odhady. Konkrétně třeba závaží C je o třetinu lehčí, než závaží D a odhad hmotnosti závaží C je rovněž o třetinu menší, než odhad hmotnosti závaží D. U ostatních dvojic sice není shoda poměrů zcela dokonalá, ale jejich hodnoty se k sobě vzájemně alespoň blíží. Dále si povšimněme, že nic z toho neplatí pro odhady Tomáše. U něj dochází i k situacím, kdy závaží, které vyhodnotil jako lehčí (např. ve dvojici A, B), je ve skutečnosti těžší. Také stojí za povšimnutí, že některá závaží, která se mu zdála stejně těžká, se reálně svou hmotností liší (D, E). Druh přesnosti, který postihuje vzájemné vztahy mezi položkami, pokud jde o hodnoty odhadu a kritéria označujeme jako *relativní přesnost*. V našem modelovém příkladu dosahuje větší relativní přesnosti Jana. Z výše uvedeného je zároveň zřejmé, že oba druhy přesnosti si nemusejí vzájemně odpovídat, tj. ve stejném úkolu a na stejných datech může jeden proband předčít druhého v úrovni své absolutní přesnosti, ale zároveň za ním zaostávat v přesnosti relativní¹⁵.

Tyto principy se analogickým způsobem uplatňují také v reálných výzkumných či diagnostických situacích, ve kterých již cíleně dochází k posuzování metakognitivního monitorování. Uvažujme například obsáhlý znalostní či didaktický test, jehož jednotlivé položky mají podobu poměrně komplexních úkolů, takže lze úspěšnost v řešení každé z nich hodnotit na vícebodové škále, např. tedy v rozsahu 0-10 bodů. Zároveň po zodpovězení každé položky studenty žádáme, aby se pokusili odhadnout počet bodů, který jim bude po opravení za položku přidělen. Po dokončení testu a jeho opravení pedagogem máme tedy pro každého studenta a pro každou položku, na niž student odpověděl, k dispozici jak jeho skutečnou úspěšnost, tak jeho subjektivní odhad vlastní úspěšnosti. Pokud se pokusíme shrnout všechny lokální odhady konkrétního studenta do jednoho celkového skóru a budeme porovnávat tuto hodnotu s reálně dosaženým výsledkem za celý test (ten se může skládat dejme tomu

¹⁵ Toto tvrzení vyjadřuje pouze fakt, že je *matematicky* možné, aby jedna a tatáž osoba dosáhla v porovnání s druhou vyšší úrovně absolutní přesnosti a zároveň nižší úrovně přesnosti relativní, a naopak. Zda v tomto směru existují nějaká *psychologická* omezení – tj. zda např. mají míry absolutní a relativní přesnosti vztahující se k metakognitivním odhadům tendenci spolu vzájemně korelovat, je již zcela jiná otázka, které se budeme ještě věnovat později.

z 20 položek a umožňuje tedy každému dosáhnout maximálně počtu 200 bodů), získáme míru absolutní přesnosti. Pokud nás naproti tomu bude zajímat, zda konkrétní položky, v nichž si student příliš nevěřil, také se zvýšenou četností vyřešil chybně (nebo jen s malým počtem bodů), musíme použít některou míru vhodnou pro vyjádření relativní přesnosti.

Každý z těchto údajů přitom vypovídá o něčem jiném a je nutné jej podle toho volit a interpretovat. Zajímá-li nás schopnost studentů odhadnout celkovou náročnost několika tematických bloků učiva, které musí všechny nastudovat, aby se připravili na nadcházející zkoušku (od takového odhadu se pak zákonitě bude odvíjet také množství času, které si ten který student na přípravu vyčlení), bude vhodným ukazatelem některý z indexů absolutní přesnosti (při jehož interpretaci budeme chápat veškeré učivo všech bloků jako jeden kompaktní celek). Naopak pokud nás zajímá, jak dokáže student mezi jednotlivými tematickými bloky identifikovat ty, které jsou pro něj celkově snadné a stačí mu v nich tedy materiály pouze zběžně prolistovat, a ty, které jsou pro něj značně náročné a musí mu v důsledku toho věnovat zvýšené množství času a pozornosti, sáhneme po některém indexu relativní přesnosti. Pochopitelně si lze alternativně představit i mnoho výzkumných či diagnostických situací, kdy buď potřebujeme zároveň oba tyto údaje, nebo nám naopak na rozdíl mezi absolutní a relativní přesností příliš nezáleží, třeba proto, že v daném kontextu předpokládáme, že obě míry by si měly být u týchž jedinců velmi podobné. Pro vyčíslení úrovně metakognitivního monitorování si pak můžeme v podstatě libovolně vybrat jeden konkrétní index, například takový, který je výpočetně nejméně náročný, či je s ním spojeno nejméně omezujících podmínek. V každém případě je nicméně vždy vhodné vědět, jaký druh informace ten který index vyjadřuje a jaký rozsah interpretace umožňuje.

V následujícím textu si postupně představíme několik důležitých indexů. Míry absolutní přesnosti a míry relativní přesnosti budou pojednány každá v samostatných kapitolách. Kromě uvedení příslušného vzorce a postupu výpočtu budou tam, kde je to nutné, doplněny další informace, např. týkající se speciálních podmínek, které musí být splněny, aby mohl být index korektně použit. Pro větší názornost si rovněž některé výpočetní postupy demonstrujeme na praktických příkladech¹⁶. Prezentovaný přehled vychází zejména z publikací Schrawa (2009), některé důležité či zajímavé detaily jsou doplněny z dalších zdrojů (Nelson, 1984; Schraw, 1995; Allwood et al., 2005; Hacker et al., 2008; a další). Vzhledem k významu a složitosti tématu se jedna podkapitola bude navíc cíleně zabývat porovnáním vlastností dvou konkrétních indexů – a sice Hamannova koeficientu (který představuje míru absolutní přesnosti) a koeficientu gamma (míra relativní přesnosti), a pokusí se poskytnout určitá obecná doporučení, kdy a proč upřednostnit jeden z nich.

¹⁶ Použité příklady jsou z didaktických důvodů sestaveny jako velmi jednoduché a krátké. V podobě, v jaké jsou zde prezentovány, by zpravidla nemohly být použity v reálném výzkumu. Rovněž při interpretaci nebudeme brát v úvahu záležitosti jako testy statistické významnosti apod. Text předpokládá, že čtenář je s těmito náležitostmi obeznámen z jiných zdrojů.

MÍRY ABSOLUTNÍ PŘESNOSTI

Pro větší názornost si některé aspekty indexů, které slouží k vyjádření absolutní přesnosti, demonstrujeme opět na fiktivních datech, která obsahuje Tabulka 3. Na rozdíl od prezentace, se kterou bychom se pravděpodobně setkali u dat z reálného výzkumu, obsahuje Tabulka 3 údaje od jediného žáka, které jsou ovšem v jednotlivých sloupcích vyjádřeny několika různými způsoby, s nimiž se můžeme ve studiích metakognitivního monitorování setkat – pro bližší vysvětlení viz legenda tabulky. Předpokládáme, že jsme zadali znalostní test o deseti otázkách (položkách), které jsou bodované pouze hodnotami 1 (správná odpověď), či 0 (nesprávná nebo chybějící odpověď).

Tabulka 3.

Údaje fiktivního posuzovaného žáka vztahující se k absolutní přesnosti jeho metakognitivního odhadu. Třetí sloupec vznikl umělou kategorizací pravděpodobnostních odhadů uvedených ve druhém sloupci (0–20 % → 1; 21–40 % → 2; 41–60 % → 3; 61–80 % → 4; 81–100 % → 5). Obdobně čtvrtý sloupec vznikl rozdělením údajů z druhého sloupce do dvou kategorií (0–50 % = nízká míra jistoty, kód 0; 51–100 % = vysoká míra jistoty, kód 1).

Kód položky	Odhad správnosti řešení v %	Odhad správnosti řešení na škále [1–5]	Kategorizované hodnoty odhadu správnosti řešení	Skutečný skór položky
A	85	5	1	1
B	12	1	0	0
C	65	4	1	0
D	44	3	0	1
E	18	1	0	1
F	75	4	1	1
G	90	5	1	1
H	7	1	0	0
I	100	5	1	1
J	35	2	0	0

Index absolutní přesnosti

Konceptuálně zřejmě nejjednodušší a nejsnáze uchopitelnou metrikou pro vyjádření absolutní přesnosti metakognitivního odhadu je index absolutní přesnosti (absolute accuracy index).¹⁷ Obecný vzorec pro jeho výpočet je:

$$AAI = \frac{1}{n} \sum_{i=1}^n (c_i - p_i)^2 \quad (1)$$

kde c_i označuje odhad obtížnosti i -té položky a p_i označuje skór vyjadřující správnost či nesprávnost probandova řešení této položky (v našem případě tedy 0 u chybné odpovědi či 1 u odpovědi správné). Číslo n značí počet položek v testu.

Pokud se přidržíme obvyklého úzu, kdy jsou procentuální odhady pravděpodobnosti vyjádřeny v desetinném formátu (např. 7 % → 0,07; 35 % → 0,35; 100 % → 1), pak platí následující: absolutní přesnost posuzované osoby je tím lepší, čím víc se její skór blíží k nule. Maximální hodnota indexu absolutní přesnosti může dosáhnout teoreticky hodnoty 1,0; takový výsledek je uváděn jako projev nejvyšší možné míry nepřesnosti (Schraw, 2009). Je ovšem důležité si uvědomit, že hodnota AAI=1 je ze své podstaty velmi zvláštní. I u člověka, který by hypoteticky vůbec nedisponoval schopností metakognitivního monitorování, nebo např. vůbec nepochopil instrukce pro udávání svého odhadu, lze totiž očekávat, že jeho AAI se bude pohybovat někde ve středu škály (tj. přibližně blízko hodnoty 0,5). Aby dosáhl hodnoty AAI rovné 1,0, musel by proband udávat svoje odhady „systematicky špatně“ – tj. u každé položky, kterou by zodpověděl správně, by musel udat subjektivní jistotu 0 % a u každé položky zodpovězené chybně by naopak musel signalizovat, že svoje řešení považuje za 100 % správné. Je jasné, že takovýto nenáhodný vzorec odpovědí signalizuje de facto dokonalou schopnost metakognitivního monitorování (alespoň vzhledem k absolutní přesnosti odhadu) a poukazuje spíše na nějaký motivační faktor na straně probanda (simulování či snaha z nějakého důvodu maskovat svoje schopnosti, žertování na účet experimentátora apod.).

Aplikujeme-li výpočet AAI na našeho fiktivního žáka, pak za hodnoty p_i dosazujeme vždy pouze 0 nebo 1, podle toho, zda položka byla zodpovězena správně, či chybně. Za c_i pak dosazujeme hodnotu odhadu správnosti řešení (uvedené ve sloupci 2 v Tabulce 3) v desetinném formátu. Například u položky A má výraz v závorce hodnotu $(0,85-1) = -0,15$. Po provedení celého výpočtu dostáváme pro žáka v našem modelovém případě hodnotu AAI = 0,17.

¹⁷ Konceptuálně totožný index se často používá i jinde než ve výzkumu metakognice, a dokonce jinde než v psychologii. S jeho pomocí lze např. posuzovat míru přesnosti, s jakou určitý expert či určitá metoda předvídají výskyt zemětřesení, tajfunů apod. V tomto kontextu se obvykle nepoužívá termínu index absolutní přesnosti, nýbrž se označuje jako Brierův index.

Index kalibrace s rozdělením posuzovací škály do kategorií

Někdy může nastat situace, že údaje o metakognitivním odhadu jsou sice registrovány na spojitě škále, výzkumník ovšem z nějakého důvodu považuje tuto škálu za zbytečně detailní a rozhodne se zjednodušit ji rozdělením do několika intervalů.¹⁸ Z takto strukturovaných dat lze pak vypočítat index, který jako první navrhl – zcela mimo kontext metakognitivního výzkumu – Murphy (1973) a dále jej popisují a pracují s ním např. Lichtensteinová a Fischhoff (1977) a Allwood s kolegy (2005). Zmínění autoři označují takto získanou mírou termínem kalibrace (calibration) – bez jakýchkoli přívlastků. Vzhledem k tomu, že termín kalibrace má v kontextu metakognitivního výzkumu poněkud širší význam, budeme zde pro tento konkrétní druh indexu používat termín „intervalová kalibrace“ (IK), přestože v původní zahraniční literatuře se s ním nesetkáme. Index intervalové kalibrace vypočteme podle vzorce:

$$IK = \frac{1}{n} \sum_{t=1}^T n_t (r_{tm} - c_t)^2 \quad (2)$$

kde n je celkový počet zodpovězených položek či úkolů, T označuje počet intervalů (ve smyslu „confidence class“), na které jsme původní škálu rozdělili, c_t je podíl správně zodpovězených položek v intervalu r_t , n_t označuje počet položek spadajících do intervalu r_t a r_{tm} vyjadřuje průměrnou míru metakognitivního odhadu v intervalu r_t .

Prakticky si můžeme celý postup opět ozřejmit na příkladu hodnot fiktivního žáka uvedených v Tabulce 3. Hodnoty metakognitivního odhadu, které byly původně vyjádřeny jako míra jistoty v procentech, tj. na škále 0–100 %, rozdělíme v prvním kroku např. do pěti kategorií, přičemž v první kategorii budou položky, kde žák udal míru jistoty v intervalu 0–20 %, ve druhé položky s mírou jistoty 21–40 % atd. Toto dělení je rovněž naznačeno ve třetím sloupci v Tabulce 3. Vidíme, že v první kategorii (tj. míra jistoty odhadu 0–20) jsou zařazeny tři položky – B, E a H. Pro první kategorii tedy platí $n_1 = 3$. Průměr hodnot metakognitivního odhadu těchto tří položek je po zaokrouhlení 0,12, čímž získáváme hodnotu r_{1m} . Z těchto tří položek byla v samotném výkonovém testu jedna zodpovězena správně a dvě chybně, poměr vyjádřený hodnotou c_1 je tedy přibližně 0,33. Obdobné výpočty je nutné provést pro všech $T = 5$ kategorií a získané hodnoty následně sečíst. Výslednou sumu v posledním kroku dělíme počtem položek n , v našem případě tedy 10. Pro posuzovaného žáka má tedy index intervalové kalibrace po zaokrouhlení hodnotu 0,07.

¹⁸ V literatuře zabývající se popisovaným indexem se pro tyto intervaly používá také pojem kategorie, což může být v některých kontextech poněkud matoucí. Alternativně se lze setkat také s obtížně přeložitelným označením „confidence class“ – česky tedy snad „konfidenční třída“ či „třída sebejistoty“.

Pokud srovnáme vzorec (2) s vzorcem (1), určeným pro výpočet indexu absolutní přesnosti, vidíme, že obě míry jsou si konceptuálně velmi podobné. Pokud máme k dispozici spojitá (nebo „téměř spojitá“) data, z kterých lze vypočítat AAI, je diskutabilní, nakolik dává smysl škálu metakognitivního odhadu následně kategorizovat za účelem výpočtu IK. V naprosté většině případů bude zřejmě přesnější a i jednodušší pracovat přímo s indexem absolutní přesnosti. IK má ovšem jednu výhodu, o které se zmiňuje Schraw (2009): lze jej – s určitou modifikací – použít i u odhadů, které nejsou původně vyjadřovány na škále 0–100, s čímž se můžeme setkat v situaci, kdy je například nutné použít místo kalibračního odhadu metriku typu EoSJ. Uvažujme výzkum, kdy pracujeme s dětmi na prvním stupni ZŠ, které ještě plně nechápu koncept pravděpodobnosti a metakognitivní odhad jsou proto schopné vyjádřit pouze v kategoriích typu („určitě to mám správně“, „asi to mám správně“, „nejsem si jistý/á“, „asi to mám špatně“, „určitě to mám špatně“). Tyto kategorie je možné skórovat na škále vyjadřující EoSJ – např. v rozsahu 0–4 či 1–5. Jinak řečeno, jsme v situaci, kdy máme v Tabulce 3 k dispozici pouze třetí sloupec, ale již nikoli údaje ve druhém sloupci. Pokud se nyní podíváme na vzorec (2), vidíme, že máme k dispozici všechny potřebné hodnoty, problematický je pouze člen r_{tm} (který byl v „klasickém“ výpočtu získán právě na základě údajů v druhém sloupci). Pokud je ovšem nahradíme nějakým kvalifikovaným odhadem (např. v logice výše nastíněného dělení na intervaly dosadíme u kategorie s nejnižší mírou jistoty přibližný střed intervalu 0–20 %, tedy 10 % – 0,1 – a podobně), index IK lze v zásadě spočítat. Je ovšem stále nutné mít na paměti, že odhadem členu r_{tm} jsme do výpočtu nevyhnutelně vnesli další nejistotu, takto získaný index by proto např. nebylo na místě použít v analýzách vyžadujících data na intervalové škále.

Bias index

Protože ve výpočtu indexu absolutní přesnosti jsou použity umocněné hodnoty rozdílů mezi výkonovým skórem a subjektivním odhadem správnosti, nevypovídá AAI nic o tom, zda posuzovaná osoba svoje skutečné výsledky systematicky přeceňuje, podceňuje či zda se v jejích nepřesnostech žádný systematický posun neprojevuje. Tendence k přeceňování, resp. podceňování vlastního výkonu je ovšem diagnosticky významná, zejména pokud je stabilní charakteristikou určité osoby. Výukové problémy studenta, který vyhledá školního či poradenského psychologa, mohou například pramenit primárně z toho, že tento student dlouhodobě přeceňuje svoje studijní schopnosti (a následně nevěnuje učení tolik času, kolik by bylo třeba). Psychologicky významný může být samozřejmě z řady různých důvodů i sklon k systematickému podceňování vlastních výkonů.

Tento druh informace zachycuje tzv. *bias index*¹⁹, který vypočítáme podle vzorce:

$$BI = \frac{1}{n} \sum_{i=1}^n (c_i - p_i) \quad (3)$$

kde c_i představuje odhad správnosti a p_i je skór vyjadřující správnost řešení položky v rámci samotného výkonového testu. Z porovnání vzorců (3) a (1) je patrné, že vzorec pro výpočet bias indexu se od vzorce pro výpočet indexu absolutní přesnosti liší právě jen tím, že hodnoty rozdílů mezi výkonovým skórem a jeho odhadem nejsou umocněny. Bias index tak může nabývat jak kladných, tak záporných hodnot, přičemž kladné hodnoty značí přílišnou sebevěru, nadhodnocování vlastního výkonu (overconfidence), zatímco záporné hodnoty signalizují sebedopodceňování, podhodnocování vlastního výkonu (underconfidence). Dosadíme-li do vzorce (3) hodnoty z Tabulky 3, vidíme, že náš fiktivní žák svůj výkon mírně podhodnocuje ($BI = -0,07$).

Indexy používané při práci s kategorickými druhy metakognitivních odhadů

Jak v případě indexu absolutní přesnosti, tak u bias indexu se pracuje s metakognitivními odhady na škále, která je spojitá, nebo alespoň „vícehodnotová“ (např. pokud probandy postavíme před nucenou volbu, kdy mohou na předložené škále vyjádřit svůj metakognitivní odhad pouze v celých desítkách procent – 0 %, 10 %, 20 % atd.). Taková škála obecně poskytuje oproti jiným alternativám nepochybně detailnější informace. V metakognitivním výzkumu se nicméně často setkáváme s nutností způsob odpovídání pro posuzované osoby co nejvíce zjednodušit, a mnohdy se proto používá design, v němž se proband vyjadřuje pouze prostřednictvím několika kategorií, v nejjednodušším případě pouze dvou (např. „jsem si jistý/á“ vs. „nejsem si jistý/á“, „moje odpověď je zřejmě správná“ vs. „moje odpověď je zřejmě chybná“ apod.). Na základě takovýchto kategorických dat nelze AAI korektně spočítat, a pro jejich analýzy se proto používají jiné indexy. Při ukázkách jejich výpočtu budeme pracovat s údaji, které jsou v Tabulce 3 uvedeny ve čtvrtém sloupci.

¹⁹ Pro anglický termín *bias* je notoricky obtížné najít vhodný český ekvivalent, mimo jiné proto, že i v originále má několik různých významů. V rámci sociologie či sociální psychologie vyjadřuje toto slovo zaujatost či předpojatost (např. vůči určité etnické skupině). Ve statistickém kontextu se někdy překládá jako zkreslení, což se asi nejvíce blíží významu tohoto termínu i ve vztahu k metakognitivnímu monitorování. Překládat termín „*bias index*“ jako „index zkreslení“ by ovšem bylo matoucí, protože by mohl vzniknout nesprávný dojem, že se jedná o zkreslení mající původ např. v procesu měření metakognitivní aktivity (a nikoli v duševní činnosti samotného probanda). Vzhledem k tomu, že v české odborné literatuře zabývající se metakognicí se objevují tendence ponechat tento termín v originální podobě (viz např. Chytrý et al., 2019), přidržíme se tohoto řešení i zde.

Vzhledem k tomu, že výkonový skóre u položky může nabývat pouze dvou hodnot (správně/nesprávně, resp. 0, 1) a metakognitivní odhad má v tomto případě také pouze dvě možné kategorie, je možné vzájemné vztahy mezi daty vyjádřit ve schématu, které vidíme na Obrázku 2. Označení jednotlivých kombinací v příslušných buňkách (a, b, c, d) budeme používat při dalším popisu.

		Správnost odpovědi ve výkonovém testu		Marginální hodnoty řádků
		ANO	NE	
Subjektivní odhad správnosti	ANO	a (zásah)	b (minutí)	a + b
	NE	c (minutí)	d (zásah)	c + d
Marginální hodnoty sloupců		a + c	b + d	

Obrázek 2.

Schéma vysvětlující jednotlivé členy, které vstupují do výpočtu prostého koeficientu shody a Hamannova koeficientu, podle Schrawa (2009). V zobrazené kontingenční tabulce typu 2x2 jsou sloupce uspořádány podle toho, zda proband odpovídal správně či chybně v kognitivní úloze. Řádky jsou uspořádány podle toho, zda svoje řešení subjektivně odhadoval jako správné, či nikoli. Buňka a tedy značí počet položek, které proband vyřešil správně a zároveň se sám domníval, že je vyřešil správně. Obdobně buňka b označuje počet položek, které reálně proband vyřešil chybně, ale domnívá se, že je má správně atd. Marginální hodnoty sloupců pak vyjadřují celkový počet správně (a+c), resp. chybně (b+d) zodpovězených položek v testu, bez přihlídnutí k subjektivnímu odhadu úspěšnosti. Celkové součty metakognitivních odhadů naopak zachycují marginální hodnoty řádků.²⁰

²⁰ Povšimněme si dále, že v tomto schématu – na rozdíl od podobného schématu spojeného s teorií detekce signálu (Obr. 4) – jsou za zásah (hit) považovány všechny případy, kdy je reálný výkon a jeho odhad v souladu, tj. jak buňka a, tak buňka d. Obdobně se situace vyjádřená buňkami b a c souhrnně označují jako minutí (miss). V tomto schématu se tedy nepracuje s pojmy falešný poplach a správné zamítnutí, se kterými se později setkáme právě v souvislosti s teorií detekce signálu.

Prostý koeficient shody (simple matching coefficient)

Dnes již poměrně málo užívaný index si zde uvedeme pouze pro úplnost. Jeho hodnota se spočítá podle vztahu:

$$SMC = \frac{a + d}{(a + d) + (b + c)} \quad (4)$$

Konceptuálně jde tedy o poměr správných metakognitivních odhadů (tj. případů kdy proband odpověděl na položku správně a svou správností si byl jistý, společně s případy, kdy odpověděl chybně a správností si přitom jistý nebyl) vůči celkovému počtu úkolů či položek v celé sadě či testu. Velkou nevýhodou tohoto indexu je skutečnost, že nedostatečně zohledňuje množství nesprávných odpovědí (tj. buňky b a c), a z toho důvodu je méně citlivý k náhodnému tipování (Schraw, 1995).

Hamannův koeficient

Další mírou absolutní přesnosti, jejíž výhodou je, že není zatížena některými nedostatky prostého koeficientu shody, je *Hamannův koeficient*.²¹ Vzorec pro jeho výpočet je následující:

$$HC = \frac{(a + d) - (b + c)}{(a + d) + (b + c)} \quad (5)$$

Jak je vidět z porovnání obou vzorců, na rozdíl od SMC vstupují do výpočtu Hamannova koeficientu i hodnoty v buňkách obsahujících minuty (neshody) – tedy situace, kdy si je proband jistý správností u chybné odpovědi a naopak; ve schématu na Obr. 2 se jedná o buňky b a c. Právě odečtení členu (b+c) v čitateli představuje jediný rozdíl mezi oběma koeficienty. Odtud je také patrné, v čem je Hamannův koeficient interpretačně výhodnější, pokud jde o vliv náhodného tipování. Uvažme situaci, kdy posuzovaná osoba není vůbec schopna metakognitivního monitorování a veškeré odhady obtížnosti proto – bez ohledu na správnost samotného kognitivního řešení – pouze náhodně tipuje. Výsledkem budou nenulové hodnoty v jedné, případně v obou z buněk a, d a zákonitě tedy i nenulová hodnota SMC při fakticky neexistujícím vlivu metakognitivního monitorování. Naproti tomu u Hamannova koeficientu

²¹ Někdy se lze také setkat s alternativním označením G-index (viz např. Schraw et al., 2014).

by měla být, při zcela náhodných odhadech obtížnosti, hodnota členu (a+d) zhruba rovna hodnotě členu (b+c), a následně by se měla i hodnota vypočteného koeficientu blížit nule. Z tohoto hlediska je Hamannův koeficient interpretačně přímočařejší a snáze uchopitelný. SMC i HC se na druhou stranu neliší v maximální dosažitelné hodnotě: pokud je proband zcela přesný ve svých odhadech, jsou hodnoty v buňkách b a c nulové, u obou koeficientů se následně vzorec fakticky zjednoduší na podíl $(a+d)/(a+d) = 1$.

V případě fiktivního žáka z Tabulky 3 použijeme pro výpočet Hamannova koeficientu „kategorizované“ hodnoty metakognitivního odhadu ve čtvrtém sloupci. To vede po dosazení k rovnici: $[(4+3)-(1+2)]/(4+1+2+3) = 0,4$.

S Hamannovým koeficientem jsou spojeny některé další specifické vlastnosti a omezení, kterými se ovšem budeme podrobněji zabývat až v samostatné kapitole věnované porovnání Hamannova koeficientu a analogické míry pro vyjádření relativní přesnosti – koeficientu gamma.

MÍRY RELATIVNÍ PŘESNOSTI

Pearsonův korelační koeficient

V případě, kdy je metakognitivní odhad vyjádřen na spojité, nebo „téměř spojité“ škále, lze jako míru relativní přesnosti snadno využít Pearsonův korelační koeficient (Schraw, 2009). Tento koeficient je zřejmě jedním z nejznámějších statistických konceptů a naprostá většina čtenářů je s ním pravděpodobně dobře obeznámena. Jeho využití pro vyjádření relativní přesnosti metakognitivního odhadu ovšem může být pro někoho na první pohled nezvyklé, a to proto, že Pearsonův korelační koeficient je zde použit jako míra metakognitivní schopnosti každého *jedince*, která může být teprve následně předmětem dalších analýz, v nichž se již srovnávají např. různé subjekty mezi sebou. Celý princip je opět možné demonstrovat na fiktivním, ale realistickém příkladu.

Uvažujme, že plánujeme porovnávat dvě školní třídy – v jedné proběhl systematický intervenční kurz zaměřený na rozvoj studijních a metakognitivních dovedností, druhá třída slouží jako kontrolní skupina, kde stále probíhala pouze standardní výuka. Po ukončení kurzu dostávají v průběhu několika dní či týdnů žáci za úkol vždy napsat krátké pojednání na určité nastudované téma, které následně učitel oboduje na škále 0-10 bodů. Každý žák zároveň po napsání uvede, kolik se domnívá, že za svoje zpracování dostane bodů. Žáci celkem zpracují a ohodnotí 8 textů. Klíčovou informací, která nás zajímá, je to, nakolik dokáží jednotliví žáci diferencovat mezi jednotlivými svými písemnými pracemi co do posouzení jejich přesnosti a kvality, jedná se tedy o typický příklad, kdy posuzujeme relativní přesnost metakognitivního monitorování.

Výsledky jednoho konkrétního žáka obsahuje Tabulka 4.

Tabulka 4.

Ukázková data pro ilustraci použití Pearsonova r jako míry relativní přesnosti metakognitivního monitorování

Kód zkoušky (eseje)	Žákem odhadovaný počet bodů	Skutečně přidělený počet bodů
ZK_1	10	9
ZK_2	8	10
ZK_3	9	10
ZK_4	10	6
ZK_5	10	10
ZK_6	10	10
ZK_7	0	9
ZK_8	10	10

Pearsonův korelační koeficient podle známého vzorce:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \quad (6)$$

pak počítáme pro každého jednotlivého žáka zvlášť tak, že korelujeme jeho odhady (sloupec II) a skutečný výkon (sloupec III). V tomto ilustračním příkladu vidíme, že hodnota korelačního koeficientu u posuzovaného žáka je téměř nulová ($r = 0,007$), tento žák tedy prakticky není schopen prostřednictvím metakognitivního monitorování diferencovat mezi situacemi, kdy podal dobrý či naopak nedostatečný výkon. Za zdůraznění také stojí fakt, že takto nízká úroveň metakognitivního monitorování se nijak nevyklučuje s relativně vysokou celkovou úspěšností v samotné sadě výkonových zkoušek, které žák dosáhl (v pěti úlohách získal maximální možný počet bodů a pouze v jedné úloze se mu vysloveně nedařilo).

Teprve když má každý žák určen svůj „osobní“ korelační koeficient, který vyjadřuje jeho relativní přesnost při provádění metakognitivního monitorování, můžeme tato data použít pro analýzu toho, co nás primárně zajímá – tedy zda je absolvování intervenčního kurzu spojeno s lepší schopností metakognitivního monitorování. V závislosti na splnění dalších podmínek provedeme tuto analýzu s pomocí některého běžného statistického testu určeného pro porovnávání dvou skupin, např. Mann-Whitneyova U testu apod.

V běžné výzkumné či pedagogické praxi najde zřejmě častěji uplatnění takový design, v němž je použit výkonový test sestávající s jednoduchých položek bodovaných pouze možnostmi „chybně“, „správně“, resp. 0, 1. Přestože to Schraw (2009) výslovně neuvádí, bylo by v těchto případech zřejmě na místě využít některou z variant Pearsonova r pro tento typ dat, například biseriální koeficient korelace.

Index rozlišení

Index rozlišení (resolution index) byl, podobně jako již dříve představený index kalibrace s rozdělením posuzovací škály do kategorií, poprvé navržen Murphym (1973). Obě metriky vycházejí z podobné logiky, index kalibrace ovšem vyjadřuje absolutní přesnost, zatímco index rozlišení je mírou přesnosti relativní.

Smysl indexu rozlišení si můžeme ukázat na následujícím příkladu. Uvažujme, že žák vyplnil znalostní test, který tvořilo 100 dílčích otázek (položek). Dále předpokládejme, že tento žák vyřešil 60 položek správně a 40 vyřešil chybně. Pokud bychom náhodně vybrali kteroukoli jednotlivou položku a měli určit, s jakou pravděpodobností v ní posuzovaný žák uspěje, pak – nemáme-li k dispozici jakékoli další informace – představuje nejlepší odhad hodnota 0,6 (tedy 60 %). Toto číslo představuje určitou *bazální hodnotu* pro celý test.

Proband má nízkou schopnost rozlišování, pokud se u všech položek testu blíží jeho subjektivní odhady úspěšnosti bazální hodnotě. Vysoká schopnost rozlišování se naopak projevuje tak, že proband dokáže rozdělit položky do kategorií (konfidenčních tříd) takovým způsobem, že poměr správně vyřešených položek v každé kategorii se výrazně odlišuje od bazální hodnoty pro celý test i od poměrů v ostatních kategoriích.

Index rozlišení (resolution index) vypočteme podle vzorce:

$$IR = \frac{1}{n} \sum_{t=1}^T n_t (c_t - c)^2 \quad (7)$$

kde n je celkový počet zodpovězených položek v testu, T označuje počet intervalů (kategorií, konfidenčních tříd), na něž je rozdělena škála subjektivních odhadů správnosti (celá škála má pochopitelně rozsah 0-100 %). Hodnota n_t označuje počet položek spadajících do intervalu r_t , c_t je podíl správně zodpovězených položek v intervalu r_t . Člen c pak označuje bazální hodnotu (tj. podíl správně zodpovězených položek v celém testu).

Koeficient gamma

Pokud potřebujeme vyjádřit relativní přesnost metakognitivního odhadu a máme k dispozici kategorická data, je častou volbou použití koeficientu gamma²². Právě s ohledem na to, že se jedná o míru relativní přesnosti, je koeficient gamma poněkud méně intuitivní (např. ve srovnání s obdobným indexem při vyjadřování absolutní přesnosti, tedy s Hamannovým koeficientem), a rovněž ilustrační příklad, na němž si výpočet a užití tohoto koeficientu ukážeme, proto bude nutně komplexnější. Příklad je volně inspirován výkladem, který podává Nelson (1984).

Uvažujme výzkumný projekt, jehož součástí je měření pocitu znalosti (FOK). Participantům jsou například předkládána česká slova a jejich úkolem je vybavit si a vyslovit anglický překlad. U slov, která si na první pokus nevybaví, participant následně posoudí, zda mají pocit, že příslušné anglické slovo se již učili (a zažívají pocit „mám to na jazyku“). Míru pocitu znalosti mohou udat pouze ve dvou kategoriích – ANO/NE. Následně je poskytnuto vodítko pro vybavení a experimentátor zaznamená, zda si s využitím tohoto vodítka ten který participant slovo vybavil, či nikoli. Tabulka 5 obsahuje výsledky jednoho fiktivního participanta.

Tabulka 5.

Data ukázkového participanta, na jejichž základě je ilustrován výpočet koeficientu gamma. Vzhledem k tomu, že v příkladu předpokládáme použití jednoduchého znalostního testu slovní zásoby, nejsou v tabulce jednotlivé položky číslovány, ale pro větší přehlednost je uveden přímo jejich obsah (tj. konkrétní slovo, které si má participant za úkol vybavit v cizím jazyce).

Položka	FOK (ANO/NE)	Paměťové vybavení s vodítkem (ANO/NE)
slon	ANO	ANO
hroch	ANO	ANO
velryba	ANO	ANO
gepard	ANO	NE
klokan	NE	ANO
opice	NE	ANO
velbloud	NE	NE
orel	NE	NE

Nyní je nutné definovat si pojmy *konkordance* a *diskordance*. Je důležité zdůraznit, že oba tyto pojmy se nevztahují k jednotlivým izolovaným položkám (např. ve smyslu souladu mezi

²² Někdy se můžeme setkat i s tím, že koeficient je označen přímo řeckým písmenem γ (viz např. Schraw, 1995).

FOK a následným paměťovým výkonem), ale ke všem možným *párům* položek v rámci celého úkolu. Dvě položky jsou vzájemně konkordantní, pokud má jedna vyšší hodnotu než druhá jak ve FOK, tak ve skóru paměťového vybavení. Jak je z Tabulky 5 patrné, pro posuzovaného participanta jsou konkordantní například páry položek „slon“ – „velbloud“, „slon“ – „orel“, „hroch“ – „velbloud“ atd. Dvojice položek je naopak diskordantní tehdy, když má jedna položka vyšší hodnotu než druhá např. ve FOK a zároveň nižší ve skóru paměťového vybavení, nebo naopak. V našem ilustračním případě jsou diskordantní dvojice položek „gepard“ – „klokan“, „gepard“ – „opice“.

Informace obsažené v Tabulce 5 lze alternativně vyjádřit i prostřednictvím schématu, které jsme již viděli na Obr. 2 v souvislosti s výpočtem Hamannova koeficientu. Obrázek 3 znázorňuje totéž schéma, za jednotlivé členy a, b, c, d jsou ovšem nyní navíc dosazeny číselné hodnoty pro tento konkrétní případ.²³

		Paměťové vybavení s nápovědou		Marginální hodnoty řádků
		ANO	NE	
FOK (pocit znalosti)	ANO	a (3)	b (1)	4
	NE	c (2)	d (2)	4
Marginální hodnoty sloupců		5	3	

Obrázek 3.

Schéma zobrazující kontingenční tabulku s hodnotami pro výpočet koeficientu gamma. Konkrétní hodnoty odpovídají datům v Tabulce 5.

²³ Tento příklad, stejně jako všechny další zde uvedené informace se vztahují ke klasické podobě koeficientu gamma, který je vypočten z kontingenční tabulky typu 2x2, pracuje se tedy pouze se dvěma kategoriemi výkonu a dvěma kategoriemi metakognitivního odhadu. Byly popsány i postupy, jak koeficient gamma zobecnit i na složitější případy (je tak možné pracovat např. se třemi úrovněmi odhadu – „jsem si jistý/á“, „nevím“, „nejsem si jistý/á“). Tato problematika je ovšem již natolik speciální, že se jí zde věnovat nebudeme – zájemce je možné odkázat např. na článek Nelsona a Narensa (1980), případně na studii Nelsona (1984), která obsahuje odkazy na další zdroje.

Jak dále uvádí Nelson (1984), míru přesnosti FOK lze obecně vyjádřit také pravděpodobnostně, v souladu se symbolikou, kterou používá, označíme tuto metriku symbolem V . Vztah:

$$V = P(A > B \text{ při vybavení} \mid A > B \text{ při FOK}) \quad (8)$$

pro všechna $A \neq B$,

vyjadřuje pravděpodobnost toho, že hodnota položky A bude vyšší než hodnota položky B při paměťovém vybavení za předpokladu, že hodnota FOK je v případě položky A vyšší než u položky B.

Přesný vzorec pro vyjádření metriky V je pak následující:

$$V = \frac{\text{počet konkordantních párů}}{\text{počet konkordantních párů} + \text{počet diskordantních párů}} \quad (9)$$

Počet konkordantních párů lze s odkazem na schéma na Obrázku 3 získat jakou součin hodnot v buňkách a, d, a počet diskordantních párů jako součin hodnot v buňkách b, c (Nelson, 1984). Pokud posuzovaná osoba není vůbec schopna metakognitivního monitorování a hodnota FOK (nebo obdobného metakognitivního odhadu) nemá žádnou souvislost s úspěšností paměťového vybavení, vede to k hodnotě $V = 0,5$. Dokonalé přesnosti odpovídá hodnota $V = 1,0$. Pokud je hodnota V nulová, jedná se o situaci, kdy proband udává svoje hodnocení nenáhodně a konzistentně chybně (jde o analogii situace popsané v podkapitole věnované indexu absolutní přesnosti).

Koeficient gamma je možné získat rozšířením koeficientu V o počet diskordantních párů v čitateli. Jde tedy do určité míry o analogický postup k tomu, jak je možné získat Hamannův koeficient rozšířením SMP. Úplný vzorec pro výpočet koeficientu gamma je následující:

$$\text{gamma} = \frac{ad - bc}{ad + bc} \quad (10)$$

Vztah mezi výše uvedenou mírou V a koeficientem gamma lze alternativně vyjádřit rovnicí:

$$V = 0,5 + 0,5 * \text{gamma} \quad (11)$$

Odtud lze dále odvodit, že hodnoty koeficientu gamma se mohou pohybovat v rozsahu od -1 do $+1$, přičemž hodnota 1 odpovídá maximální přesnosti metakognitivního odhadu, zatímco hodnoty okolo nuly naznačují naprostou neschopnost provádět v daném typu úlohy metakognitivní odhad (záporné hodnoty vyjadřují tendenci k „systematicky chybnému“ odpovídání v podobném duchu, jako jsme už viděli např. v kapitole věnované indexu absolutní přesnosti). Hodnota koeficientu gamma odpovídající datům v Tabulce 5 je rovna $0,5$.

Odvození koeficientu gamma od metriky V je výhodné v tom, že umožňuje hodnoty gamma interpretovat pravděpodobnostním a vcelku intuitivním způsobem, jaký zmiňují např. Nelson a Dunlosky (1991). Vztáhneme-li jej na náš příklad, můžeme se u konkrétních dvojic položek ptát: jestliže žák uvede u jednoho slova vyšší FOK než u druhého, a jedno z těchto slov si následně s pomocí nápovědy vybaví, jaká je pravděpodobnost, že vybaveno bude to slovo, u kterého předtím uvedl vyšší FOK? Vzhledem k tomu, že při zcela náhodném tipování při uvádění hodnot FOK je tato pravděpodobnost $0,50$ a při zcela dokonalé schopnosti metakognitivního monitorování $1,00$, lze pak pravděpodobnost, kterou je možné určit na základě hodnoty gamma (viz vzorec (11)), porovnávat s těmito krajnostmi. To dává poměrně intuitivní představu, protože je tak názorně vidět, zda má metakognitivní výkon probanda blíže spíše k chaotickému, náhodnému charakteru (např. pravděpodobnost $0,60$), nebo naopak k dokonalému monitorování (např. pravděpodobnost $0,90$).

Porovnání Hamannova koeficientu a koeficientu gamma

Zřejmou výhodou výzkumného designu, při němž se metakognitivní odhad registruje pouze na úrovni dvou kategorií, je jeho jednoduchost, nejen z hlediska následných analýz, ale většinou i z pohledu samotných probandů v čase, kdy plní zadaný úkol. Takto koncipovaných výzkumů proto bylo realizováno velké množství, a často se proto můžeme setkat s použitím Hamannova koeficientu (pokud je třeba zjišťovat absolutní přesnosti odhadu), či koeficientu gamma (pokud má být zjišťována přesnost relativní). Oba koeficienty pracují s podobným schématem vstupních dat, tedy s kontingenční tabulkou typu 2×2 , porovnávající metakognitivní odhad a správnost řešení zadané kognitivní úlohy. Často proto nemusí být zřejmé, jaký koeficient v jaké situaci použít. V minulosti probíhaly okolo tohoto tématu poměrně ostré odborné diskuze (např. Schraw, 1995; Nelson, 1996) a i pozdější studie (např. Maki et al., 2005; Nietfeld et al., 2006; Schraw et al., 2014; Lingel et al., 2019) přinesly v této oblasti některá důležitá, i když často nejednoznačná zjištění. S každým z těchto indexů jsou totiž spojeny určité nevýhody i omezení, a může být proto užitečné porovnat obě míry na jednom místě, např. za účelem informovaného rozhodnutí pro to, kterou z nich použít v plánovaném či připravovaném výzkumu.

Pokud je ze samotného designu výzkumu (nebo z účelu praktické diagnostiky) zřejmé, zda nás zajímá absolutní či relativní přesnost metakognitivního odhadu, je rozhodnutí vcelku jednoduché. Dobrá znalost některých aspektů zvoleného koeficientu může být v takové si-

tuaci užitečná zejména pro včasnou úpravu výzkumného designu tak, aby následné analýzy nebyly zkresleny omezeními, která jsou s koeficientem spojena (např. úprava položek tak, aby se minimalizoval počet nulových buněk v kontingenčních tabulkách při použití koeficientu gamma – viz dále). V některých případech ale může výzkumník či diagnostik vyžadovat v podstatě jakýkoli indikátor metakognitivního monitorování, respektive předpokládat, že jak relativní, tak absolutní přesnost budou u jednotlivých probandů velmi podobné. Tento předpoklad je sám o sobě značně diskutabilní, jak ještě zmíníme dále, pokud jej ovšem přijmeme, dává větší smysl zvolit si konkrétní indikátor především s ohledem na jeho praktickou použitelnost.

Pro ilustraci některých problémů spojených s koeficientem gamma si nejprve uvedeme Tabulku 6, která je inspirována podobným schématem, jaké ve své studii uvádí Schraw (1995).

Tabulka 6.

Ilustrační data pro vysvětlení některých rozdílů mezi koeficientem gamma a Hamannovým koeficientem a pro ukázkou některých problematických vlastností obou koeficientů. Hodnoty buněk, které se při práci s těmito koeficienty standardně zobrazují v kontingenčních tabulkách – viz např. schémata na Obr. 2, 3 – jsou zde uvedeny vždy na jednom řádku, společně s odpovídajícími hodnotami obou koeficientů.

Kód žáka	a	b	c	d	Koeficient gamma	Hamannův koeficient
žák_1	15	5	0	0	nedef.	0,50
žák_2	0	2	0	18	nedef.	0,80
žák_3	2	9	0	9	1,00	0,10
žák_4	16	2	2	0	-1,00	0,60
žák_5	5	5	5	5	0,00	0,00
žák_6	5	9	1	5	0,47	0,00

Jako velmi podstatné se pak jeví zejména tyto dva problémy:

1. Pokud se v kontingenční tabulce vyskytne prázdná buňka (přesněji řečeno, buňka s hodnotou 0), má to výrazný dopad na hodnotu koeficientu gamma. Tento jev může mít několik různých podob. Jestliže mají nulové hodnoty obě buňky ve stejném sloupci, či ve stejném řádku kontingenční tabulky (tj. situace, kdy $a=b=0$; $a=c=0$ atd.), hodnota jmenovatele v rovnici (10) je rovněž nula, koeficient gamma v takovém případě nelze vypočítat a jeho hodnota je nedefinována. K takové situaci dojde vždy, pokud proband udává u všech úkolů totožnou kategorii metakognitivního odhadu (např. kompletně u všech položek udá, že si je jistý správností svého řešení). Druhá situace, která nevyhnutelně vede k nedefinované

hodnotě koeficientu gamma, spočívá v naprosté úspěšnosti, resp. naprosté neúspěšnosti probanda v samotné výkonostní zkoušce (tj. pokud proband např. všechny položky zodpoví chybně, je hodnota gamma nedefinovaná bez ohledu na rozložení metakognitivních odhadů). Koeficient gamma se bude rovnat jedné, kdykoli bude mít nulovou hodnotu kterákoli z buněk b, c, a analogicky bude roven mínus jedné, pokud bude mít nulovou hodnotu kterákoli z buněk a, d.

Z hlediska výzkumu metakognice v souvislosti s mimořádným intelektovým nadáním představuje zejména varianta „prázdného sloupce“ (tj. ve výkonovém testu má posuzovaná osoba buď všechny položky správně, nebo všechny položky chybně) významný intervenující faktor, který může zároveň nastat velmi snadno a je nutné s ním realisticky počítat. To se týká především výzkumných designů, v nichž jsou nadané děti srovnávány s kontrolní skupinou běžných žáků, a pro obě skupiny se přitom využívá tatáž sada experimentálních úkolů. Snadno pak může dojít k tomu, že úlohy, které jsou pro průměrné žáky zhruba středně obtížné a v této skupině je tak patrná určitá variabilita, jsou pro nadané celkově snadné a řada dětí v nich pak skutečně téměř nechybje. Pochopitelně si lze představit i zrcadlově opačnou situaci, tj. administraci úloh navržených pro intelektově nadané žáky průměrným dětem, což povede k vysoké, eventuálně až stoprocentní neúspěšnosti. V obou případech se dá s opodstatněním předpokládat výrazné zkreslení údajů o metakognitivních schopnostech vyjádřených prostřednictvím koeficientu gamma v té skupině, v níž jsou výsledky výkonových zkoušek podobným způsobem nevyrovnané.

Tyto obavy z velké části potvrzují výsledky nedávné studie Lingela a kolegů (2019), která se zaměřovala cíleně právě na posouzení vhodnosti několika různých metakognitivních indexů. Badatelé sledovali vzorek 109 žáků na druhém stupni základní školy (7. ročník, věk cca 12 let), jako podnětové úkoly pro vyvolání metakognitivního monitorování použili sadu 20 matematických příkladů. Autoři mimo jiné uvádějí, že až čtvrtina participantů v jejich souboru měla v některých analýzách nedefinované hodnoty koeficientu gamma v důsledku prázdných buněk a až u poloviny nabývala hodnota koeficientu gamma některé z hraničních hodnot (tj. -1 , 0 , nebo 1), přičemž k této situaci docházelo významně častěji u úspěšnějších (z hlediska správnosti řešení zadaných matematických úloh) participantů. Autory vedly tyto závěry k pochybnostem, zda je vůbec vhodné jakékoli použití koeficientu gamma v pedagogickém výzkumu. Schraw (1995) zmiňuje liberálnější doporučení – koeficient gamma lze využít tehdy, pokud žádná buňka v kontingenční tabulce neobsahuje nulovou hodnotu a zároveň 80 % buněk obsahuje alespoň 5 prvků (tato hodnota odkazuje na rozšířenou variantu koeficientu gamma, který lze vypočítat i při více kategoriích, než ve schématu 2×2). Zároveň je vhodné podotknout, že Hamannův koeficient obecně není citlivý na přítomnost buněk s nulovou hodnotou, i když – jak ještě uvidíme – i v případě této míry je problematická přítomnost celého nulového řádku či sloupce.

2. Nerovnoměrné rozložení hodnot metakognitivního odhadu či výsledků výkonových úloh může vést k zavádějícím výsledkům analýz i v případech, které nejsou tak extrémní jako ty

uváděné v předchozím bodě. Jde třeba o situace, kdy součet hodnot v jednom sloupci (např. $b+c$) je výrazně, často i mnohonásobně menší než součet hodnot v druhém sloupci, není ovšem nulový. V takových případech je hodnota koeficientu γ citlivá na rozložení hodnot podél diagonál a/d a b/c . To je patrné například na srovnání žáků 5 a 6 v Tabulce 6. Obě posuzované osoby mají stejný počet odpovědí, které byly správné a subjektivní odhad je také jako správné ohodnotil (buňka a), i hodnot, které byly nesprávné a subjekt při nich udal nízkou jistotu správnosti řešení (buňka d). Vzhledem k velkým nerovnoměrnostem na diagonálách $b-c$ se ovšem jejich hodnoty koeficientu γ poměrně výrazně liší.

Jak ovšem zdůraznil Nelson (1996), i prostý koeficient shody, a následně i Hamannův koeficient, který je od něj odvozen, jsou náchylné ke zkreslení při určitém charakteru nerovnoměrně rozložených dat. Pokud se podíváme na žáka 2 v Tabulce 6, vidíme, že tento člověk měl jednak celkově velmi nízký počet správných odpovědí ve výkonovém testu. Nulové hodnoty v buňkách a, b zároveň naznačují, že při udávání subjektivního odhadu mohl zvolit strategii, kdy bez ohledu na cokoli vždy podal negativní odpověď. Přestože jeho subjektivní odhady tedy nemusí mít jakoukoli spojitost se schopností metakognitivního monitorování, dosáhne celkově vysoké hodnoty Hamannova koeficientu. Nelson (1996) tak problematizuje vhodnost HC i jako míry absolutní přesnosti.

Lingel a kolegové (2019) v nedávném článku poukazují na přetrvávající nedostatek empirických studií zaměřených na metodologické otázky spojené s výzkumem metakognitivního monitorování, o to překvapivějším, že tato schopnost je čím dál častěji předmětem pedagogického výzkumu. Přesto ovšem byly v uplynulých zhruba dvaceti letech publikovány některé další práce zabývající se použitím obou koeficientů. Jejich nejdůležitější zjištění se pokusíme stručně shrnout.

Za prvé je zde otázka, naznačená již v úvodu této kapitoly, nakolik mohou být míry absolutní a relativní přesnosti v různých výzkumných designech vzájemně zaměnitelné. Rozhodnutí zvolit libovolně pouze jeden z porovnávaných koeficientů jako obecný indikátor metakognitivního monitorování se implicitně opírá o předpoklad, že oba koeficienty vyjadřují v podstatě stejný konstrukt a že u týčů jedinců je jejich hodnota víceméně podobná (jinými slovy, předpokládá se celkově vysoká korelace obou měř). Práce, které se cíleně zaměřily na tuto otázku, dochází k nejednoznačným závěrům. Například Schraw a kolegové (2014) pozorovali celkově silné korelace mezi mírami absolutní a relativní přesnosti (v této studii byly porovnávány i další míry, nejen Hamannův koeficient a γ). Podobný vztah zaznamenali i Lingel a kolegové (2019). I když zjištěné korelace mezi ukazateli absolutní a relativní přesnosti byly v jejich studii podstatně slabší než v předchozí studii, podle autorů nejsou v rozporu s modelem, v němž absolutní i relativní přesnost metakognitivního odhadu reflektuje různé projevy téhož konstrukt, který stojí v jejich pozadí. Zároveň je ale vhodné zmínit např. studii, kterou provedli Maki et al. (2005), kteří naopak uvádějí prakticky nulové korelace mezi mírami absolutní a relativní přesnosti (u absolutní přesnosti se nicméně nejednalo o Hamannův koeficient, nýbrž o variantu AAI).

Předmětem zmíněné studie Makiové a kolegů byl také další důležitý problém, a sice otázka reliability koeficientu gamma. Autoři se pokoušeli ověřit závěry některých předcházejících výzkumů, které vyznívaly v tom smyslu, že výsledky metakognitivních úloh vyjádřené prostřednictvím koeficientu gamma jsou z velké míry nereliabilní (jak z hlediska test-retest reliability, tak z hlediska reliability split-half). K problému přistoupili tak, že po participantech žádali ke stejným úlohám jak prediktivní, tak postdiktivní odhad, pro oba odhady spočetli hodnoty gamma a reliabilitu posuzovali na základě vzájemných korelací těchto hodnot. Oba odhady byli konzistentní, což autoři interpretovali tak, že míry absolutní přesnosti mohou lépe vyjadřovat určitý stabilní, trvalý sklon jedince, zatímco míry relativní přesnosti spíše odrážejí krátkodobé, situační faktory (např. míru stresu spojeného se zkouškou atp.).

Z výše uvedeného tedy vyplývá, že míry absolutní a relativní přesnosti nemusí nutně vyjadřovat tutéž schopnost, a je tedy problematické volit např. mezi Hamannovým koeficientem a koeficientem gamma zcela libovolně, případně pouze s ohledem na omezení související s jeho praktickým užitím. Nejasné výsledky citovaných studií mohou – vedle zmíněných situačních faktorů – souviset i s vývojovou dynamikou (např. Lingel a kolegové, 2019, zkoumali žáky 2. stupně ZŠ, Schraw a kolegové, 2014, pracovali s vysokoškoláky...), s odlišným typem experimentálních úkolů a podobně. Problém si nepochybně zaslouží i do budoucna další zkoumání.

Vzhledem k tomu, že v typickém výzkumném designu – s možnou výjimkou čistě exploračních studií – jsou oba diskutované koeficienty využívány především jako podklad pro vnitro-subjektové či mezisubjektové statistické porovnávání, mohou být pro badatele přínosné i výsledky simulační studie, kterou realizovali Nietfeld a kolegové (2006). Autoři se prostřednictvím metody Monte Carlo pokusili stanovit některé statistické vlastnosti Hamannova koeficientu a koeficientu gamma, a to jednak ve scénáři obsahujícím tisíc položek, jednak ve scénáři s padesáti položkami. V případě tisícipoložkového scénáře si byly distribuční vlastnosti obou koeficientů dost podobné, především se v obou případech jednalo o normální rozdělení hodnot. Je ale zřejmé, že z pohledu praktické diagnostiky či výzkumu je takové uspořádání značně nerealistické – těžko si představit studii, v níž bude možné zadat participantům test s tisíci dílčími úkoly. V případě mnohem realističtějšího padesátipoložkového scénáře pak již byly mezi oběma koeficienty patrné významné rozdíly: zatímco Hamannův koeficient vykazuje i v tomto případě normální rozdělení, rozdělení koeficientů gamma je zešíkmené a leptokurtické. Z pohledu plánování výzkumu může být proto užitečné vědět, že při využití koeficientu gamma nebude pravděpodobně možné provádět následně testy vyžadující normalitu rozdělení.

Poslední dvě zjištění, která zde zmíníme, se týkají podoby podnětových úkolů a jejich možného vlivu na hodnoty koeficientů metakognitivní přesnosti. I zde by byly samozřejmě žádoucí robustnější důkazy, na základě studií, které jsou k dispozici se nicméně zdá, že formát podnětových úloh, který je výhodnější při měření absolutní přesnosti, může být problematický při měření přesnosti relativní a naopak. Nedávno publikovaná studie, kterou provedli

Vuorre a Metcalfová (2021), která také využívala simulace prostřednictvím metody Monte Carlo, došla k závěru, že při měření *relativní přesnosti* (vyjádřené, mimo jiné, koeficientem gamma), představuje určité riziko formát testu s volbou správné odpovědi z několika nabízených alternativ (multiple-choice). Možnost skórovat v části položek pouze na základě náhody zkresluje závislost mezi výkonem a hodnotou gamma odvozenou od metakognitivního odhadu. Autoři z toho dovozují, že mnoho dříve publikovaných zjištění v tom smyslu, že lepší úroveň metakognitivního monitorování (vyjádřeného koeficientem gamma) vede k lepšímu studijnímu výkonu, tak mohlo být ve skutečnosti pouze výzkumným artefaktem. V případě zjišťování *absolutní přesnosti* vyjádřené prostřednictvím Hamannova koeficientu se na základě výsledků, ke kterým došli Lingel a kolegové (2019), naopak jeví jako výhodnější formát úkolů s nucenou volbou odpovědi, který v jejich studii vedl k celkově větší přesnosti metakognitivního odhadu.

I s přihlédnutím k těmto novějším studiím se stále jeví jako smysluplná a relevantní doporučení, která před více než čtvrt stoletím formuloval Schraw (1995) a která můžeme parafrázovat takto: 1) pokud se pracuje s kategorickými daty, je vhodné všude, kde je to možné, počítat, analyzovat a reportovat jak koeficient gamma, tak Hamannův koeficient současně. Jestliže se obě míry výrazně liší, je nutné při interpretaci těchto rozdílů zohlednit specifické vlastnosti a omezení obou indexů (tj. rozdíl nemusí mít vždy psychologický význam, může být pouze artefaktem dané metriky); 2) jestliže existuje zvýšené riziko, že jeden či oba z těchto koeficientů mohou být zkresleny z důvodu výše uvedených omezení, je lepší upřednostnit takové míry absolutní či relativní přesnosti, které nepracují s kategorickými daty (např. index absolutní přesnosti, Personův korelační koeficient); 3) pokud už je nutné pracovat pouze s kategorizovanými odhady a je nutné zvolit pouze jeden indikátor, měl by se přednostně volit Hamannův koeficient, pokud je velikost výzkumného souboru relativně malá, nebo je pravděpodobné, že typ použitých úloh povede k prázdným buňkám v kontingenční tabulce. V ostatních případech je možné volit koeficient gamma.

METRIKY ZALOŽENÉ NA TEORII DETEKCE SIGNÁLU

Poslední dvě metriky, které si představíme – *AUROC2* a *meta-d'*, jsou odvozeny od tzv. *teorie detekce signálu* (signal detection theory, SDT). Ve své aktuální podobě jsou plodem poměrně nedávného vývoje v metodologii zjišťování metakognice, přestože i v tomto případě lze nalézt kořeny a inspirační zdroje již v některých průkopnických pracích datujících se až do 50. let minulého století (např. Clarke et al., 1959). Problematika využití SDT jako východiska pro odvození nových metakognitivních indikátorů byla ovšem významně rozpracována zejména během uplynulé dekády, zásadní význam měla především studie Maniscalca a Laua z roku 2012, na niž pak navázaly některé další práce (Fleming & Lau, 2014; Maniscalco & Lau, 2014; Fleming, 2017).

Prakticky všechny metriky, dosud uvedené v této publikaci, se dají vyjádřit poměrně jednoduchým vzorcem a čtenář si je ve své praxi může v případě potřeby vypočítat ručně, případně s pomocí běžného tabulkového kalkulátoru. Výpočet metrik odvozených od SDT se v tomto ohledu liší, je celkově komplikovanější a úplné pochopení logiky výpočtu vyžaduje alespoň středně pokročilé matematické znalosti – např. odvození *meta-d'* (Maniscalco & Lau, 2014) se neobejde bez pochopení základních principů diferenciálního a integrálního počtu. Detailní a formálně úplné odvození zde tedy provádět nebudeme, čtenáře se zájmem o hlubší vhled do této problematiky je možné odkázat na právě zmíněné původní zdroje. Výpočet *AUROC2* je dále podrobně popsán např. v článku, jehož autory jsou Hauser a kolegové (2017). Zde se zaměříme pouze na stručný nástin základních principů SDT a jejich rozvinutí nezbytného pro měření metakognice. Výhodou obou představených metrik je na druhou stranu skutečnost, že pro jejich praktické použití není detailní znalost výpočetního postupu nutná: v obou případech existují pro tento účel volně dostupné softwarové nástroje v podobě skriptů spouštěných v prostředí MATLAB.²⁴

²⁴ Skript pro výpočet *meta-d'* je (stav k 1. 7. 2021) dostupný na adrese: <http://www.columbia.edu/~bsm2105/type2sdt/archive/index.html>. Skript pro výpočet *AUROC2* je k dispozici v doplňujících materiálech ke studii: Fleming & Lau, 2014.

Hlavní přínos využití SDT jako východiska pro měření metakognitivního monitorování spočívá v tom, že tento přístup nabízí řešení závažného metodologického problému, zatěžujícího zejména „klasické“ míry relativní přesnosti. Podstatou zmíněného problému je fakt, že tyto míry mohou být zkresleny celkovým biasem respondenta. Pro ilustraci se ještě jednou vrátíme k využití Pearsonova r jako míry relativní přesnosti. Uvažujme dva žáky, kteří mají zcela totožný výkon v kognitivním testu (i na úrovni odpovědí na jednotlivé položky), a liší se přitom pouze ve svých metakognitivních odhadech. První žák – pojmenujme jej například Roman – má přitom celkově větší záporný bias, jinak řečeno má větší sklon k podceňování svého výkonu než druhý žák (kterého pojmenujeme třeba Marek). Na první pohled by se mohlo zdát, že celkový bias by nemusel příliš ovlivnit míru *relativní* přesnosti, pokud by bylo sebepodcenění rovnoměrně distribuováno mezi jednotlivé položky. To ovšem není z principu možné vzhledem k položkám, u nichž dosahuje výkon respondentů krajních (extrémně nízkých či extrémně vysokých) hodnot.

Jestliže je možné kvalitu odpovědi na každou z položek v testu ohodnotit na škále 0–10 bodů, u položek, v nichž oba žáci dosáhli zhruba „středního“ výkonu (tj. jejich odpovědím bylo následně hodnotitelem přiděleno např. 5 či 6 bodů), se Romanův bias systematicky projeví víceméně konstantně nižšími odhady ve srovnání s Markem. Například u položek, ve kterých oba chlapci reálně dosáhli 6 bodů, může Roman průměrně odhadovat svůj výkon třemi body, zatímco odhady Marka (u něhož se může naopak projevit mírný pozitivní bias, tedy tendence svůj výkon nadhodnocovat) budou u týchž položek odhady průměrně například sedmibodové. Pokud by ale tento rozdíl měl být zachován konzistentně, tedy včetně hodnocení těžších položek, v nichž se oběma žákům příliš nedařilo, tak např. u položky, kde Marek udává 3 body, by Roman musel udat –1 bod, což je pochopitelně nereálné a nesmyslné.

Tím, že možné hodnoty odhadů jsou zdola i shora omezeny (proband nemůže z principu odhadnout nižší hodnotu než nula a vyšší hodnotu, než činí maximální bodová hodnota položky či obdobného základního prvku posouzení), mohou být při určitém profilu odpovědí metriky vyjadřující relativní přesnost poměrně výrazně zkresleny. Budeme-li např. posuzovat dvě osoby, které se liší svým celkovým biasem, a zadáme jim test sestávající z přibližně průměrně obtížných položek, mohou mít oba prakticky stejné hodnoty relativní přesnosti, vyjádřené některým z klasických indexů. Zadáme-li týmž osobám jiný test, obsahující větší množství extrémních (velmi snadných nebo velmi obtížných položek), mohou se jejich hodnoty relativní přesnosti poměrně výrazně lišit. Pro úplnost je vhodné dodat, že tento problém se neomezuje pouze na klasické Pearsonovo r , ale týká se i ostatních měr relativní přesnosti, jako je např. gamma, speciální formy Pearsonovy korelace atd. (Fleming & Lau, 2014). Základní výhodou metrik odvozených od SDT je skutečnost, že dokáží vyjádřit relativní přesnost metakognitivního monitorování jedince nezatíženou jeho případným biasem.

Samotná problematika teorie detekce signálu je velmi komplexní a na tomto místě není možné poskytnout více, než pouze rámcový a nezbytně značně povrchní pohled. Zájemce o hlubší pochopení SDT lze odkázat např. na obsáhlou publikaci, jejímž autorem je Wickens (2002). Přístup SDT byl vyvinut zhruba v polovině 20. století, a to původně ze zcela praktických důvodů, konkrétně pro posuzování schopnosti operátorů leteckých radarů správně vyhodnotit přijímané informace a identifikovat podstatné události (např. přelet neznámého letadla). Analytické postupy, které byly za tímto účelem vyvinuty, se nicméně následně široce uplatnily i v základním výzkumu v mnoha vědních oborech, včetně psychologie – zde např. ve výzkumu vnímání, ale i paměti a dalších složitějších kognitivních jevů.

Základní principy SDT si demonstrujeme na následující modelové situaci: uvažujme výzkumné sezení, které sestává z velkého množství jednotlivých dílčích pokusů (angl. trials). V každém pokusu je po určitý čas prezentován náhodný šum (předpokládejme, že pracujeme s podněty ve sluchové modalitě, šum tedy bude mít podobu směsi náhodných zvuků o různých frekvencích). V některých pokusech je šum tím jediným, co je pokusné osobě prezentováno, v některých pokusech ovšem zazní na pozadí šumu také signál – předem známý zvuk o konkrétní frekvenci a intenzitě. Úkolem posuzované osoby je v každém pokusu určit, zda byl signál přítomen či nikoli.

Výstup z takto pojatého experimentálního sezení zobrazuje schéma na Obrázku 4. Schéma je velice podobné tomu, se kterým jsme se již setkali např. v souvislosti s Hamannovým koeficientem (viz Obr. 2). Základní rozdíl je konceptuální, v SDT se explicitně rozlišují všechny typy kombinací mezi reálnou přítomností signálu a reportováním přítomnosti signálu

		<i>Typ pokusu</i>	
		Signál	Šum
<i>Signál detekován</i>	<i>ANO</i>	zásah (hit)	falešný poplach (false alarm)
	<i>NE</i>	minutí (miss)	správné odmítnutí (correct rejection)

Obrázek 4.

Schéma zachycující možné výsledky jednotlivých pokusů či položek v modelu teorie detekce signálu.

v odpovědi probanda. Z logiky věci může v jednom pokusu nastat některá ze čtyř možností: a) signál byl reálně prezentován a proband jeho přítomnost zaznamenal. Tyto případy označujeme jako *zásah* (hit); b) signál byl přítomen, ale proband jeho přítomnost nezaznamenal. Takové situace označujeme jako *minutí* (miss); c) signál nebyl přítomen, ale proband se chybně domnívá, že jeho přítomnost zaznamenal, v tomto případě hovoříme o *falešném poplachu* (false alarm); d) konečně poslední situace spočívá v tom, že signál nebyl přítomen, což proband správně určil. Tyto případy označujeme jako *správné odmítnutí* (correct rejection).

Metriku, kterou získáme tak, že počet zásahů (tedy číslo v levé horní buňce) vydělíme celkovým počtem pokusů, v nichž byl signál reálně přítomen (tedy číslo odpovídající marginální četnosti prvního sloupce), označujeme jako *podíl zásahů* (hit rate, HR). Formálně bude vztah pro výpočet vypadat následovně:

$$HR = p(\text{odpověď probanda} = \text{„signál přítomen“} / \text{skutečnost} \\ = \text{„signál přítomen“}) = \frac{\text{celkový počet zásahů}}{\text{celkový počet zásahů} + \text{celkový počet minutí}} \quad (12)$$

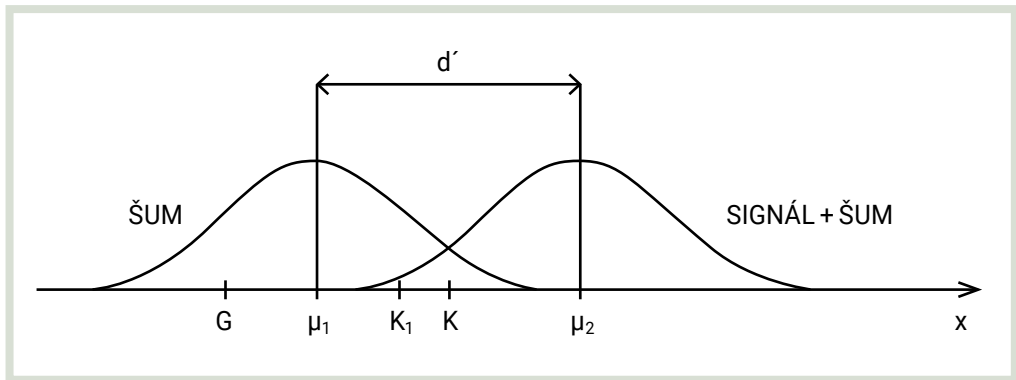
Obdobně získáme metriku označovanou jako *podíl falešných poplachů* (false alarm rate, FAR). V tomto případě ovšem bude v čitateli hodnota v pravé horní buňce a ve jmenovateli celkový počet pokusů, v nichž signál reálně nebyl přítomen. Formálně jde tedy o vzorec:

$$FAR = p(\text{odpověď probanda} = \text{„signál přítomen“} / \text{skutečnost} \\ = \text{„signál nepřítomen“}) = \frac{\text{celkový počet falešných poplachů}}{\text{celkový počet falešných poplachů} + \text{celkový počet správných odmítnutí}} \quad (13)$$

Stejným způsobem by se daly získat i poměry pro zbývající dvě kategorie, tedy pro minutí a správná odmítnutí. Vzhledem k tomu, že veškeré informace nutné pro analýzy v rámci SDT jsou pro konkrétní osobu a konkrétní výzkumné sezení plně obsaženy v hodnotách HR a FAR, se tyto doplňkové metriky většinou nepočítají a neuvádějí (Wickens, 2002).

SDT se primárně zabývá tím, jak spolu oba indikátory, tedy HR a FAR, vzájemně souvisí, na čem závisí jejich hodnoty a jak lze těmito hodnotami případně systematicky manipulovat. Se značnou mírou zjednodušení lze říci, že hodnoty HR a FAR závisí na třech základních faktorech. Prvním faktorem je intenzita signálu, přesněji řečeno její poměr k intenzitě šumu na pozadí. Je celkem samozřejmé, že snadněji se bude identifikovat silný signál na pozadí slabého šumu než slabý signál na pozadí silného šumu. Za druhé budou hodnoty HR a FAR dále záviset na percepčních (či jiných, v závislosti na povaze úkolu) schopnostech daného jedince. V terminologii SDT se schopnost člověka správně rozlišovat situace s přítomností

a nepřítomností signálu označuje termínem *senzitivita* a vychází se z předpokladu, že v míře senzitivity se jednotliví lidé vzájemně liší, analýzy vycházející z SDT tedy fungují jako nástroj pro diagnostiku této schopnosti. Za třetí pak hodnoty HR a FAR, a jejich vzájemný poměr, závisí na tendenci posuzovatele odpovídat spíše liberálně (tj. když si není jistý, raději reportuje přítomnost, než nepřítomnost signálu), či konzervativně (v situaci nejistoty má sklon udávat spíše nepřítomnost signálu, než jeho přítomnost). Premisy, ze kterých vychází SDT, umožní lépe pochopit schéma zobrazené na Obrázku 5.



Obrázek 5.

Schematické znázornění rozdělení vnitřních stavů, odpovídajících přítomnosti samotného šumu, nebo kombinaci signál + šum v nejjednodušším SDT modelu. Hodnota d' , která je rovna rozdílu středních hodnot obou rozdělení, udává senzitivitu posuzovaného jedince.

Základní princip spočívá v tom, že fyzikální podněty z vnějšího prostředí (v našem případě tedy zvukové), jsou zpracovány příslušnými funkčními mechanismy nervové soustavy a následně se subjektivně odrážejí v mysli jedince jako určitý „vnitřní stav“²⁵. Jedinec se pak rozhoduje, zda v daném pokusu zaznamenal či nezaznamenal přítomnost signálu právě na základě tohoto vnitřního stavu, nikoli přímo na základě objektivní hodnoty fyzikálního podnětu. Fakt, že objektivní povahu vnějšího podnětu a jeho mentální reprezentaci nelze ztotožňovat je důsledkem toho, že lidé se navzájem liší ve svých schopnostech. Kdyby tomu tak nebylo a rozhodnutí by záviselo čistě na fyzikálních charakteristikách podnětu, pak by při stejných podnětových materiálech dosahovaly všechny pokusné osoby prakticky totožných

²⁵ V anglicky psané literatuře se někdy můžeme setkat i s termínem „distribution of internal signals“ pro toto rozdělení (např. Fleming & Lau, 2014). Tento termín je nicméně poněkud matoucí, protože se fakticky vztahuje i na rozdělení v pokusech, ve kterých je přítomen pouze šum.

hodnot HR a FAR (k čemuž ve skutečnosti nedochází). Intenzita vnějších podnětů se nicméně do rozhodování promítne nepřímo, protože je zřejmé, že silnější podnět obecně vyvolá silnější vnitřní odpověď, tato závislost je ovšem spíše volná a také není lineární.

Osa x na Obr. 5 vyjadřuje subjektivně vnímanou intenzitu vnitřních stavů, vyvolaných prezentací externích podnětů. SDT předpokládá, že rozdělení vnitřních stavů v pokusech, v nichž je přítomen pouze šum, je přibližně normální, na zmíněném obrázku tomu odpovídá křivka normálního rozdělení umístěná více vlevo s průměrem μ_1 . Hodnota směrodatné odchylky tohoto rozdělení se zpravidla uvažuje jednotková. Pro názornost je na ose x umístěn bod G, ve vzdálenosti -1 SD od průměru rozložení pro pokusy pouze se šumem. S přihlédnutím ke známým vlastnostem normálního rozdělení můžeme vyslovit následující interpretaci: ve zhruba 15 % pokusů, v nichž je prezentován pouze šum, bude pokusná osoba vnímat intenzitu zvukového podnětu jako nižší, než odpovídá subjektivnímu vnímání této intenzity v bodě G. Dále je nutné si uvědomit, že v pokusech, v nichž je přítomen signál, není nikdy prezentován *pouze* signál, vždy se jedná o signál na pozadí šumu. V nejjednodušším modelu se dále předpokládá, že variabilita samotného signálu je vzhledem k variabilitě šumu zanedbatelná²⁶, signál tedy můžeme chápat jako konstantu, která je vždy přičtena k hodnotě šumu. Důsledkem toho je, že rozdělení subjektivních stavů odpovídajících pokusům s kombinací signál+šum, je také normální a má stejnou směrodatnou odchylku (jednotkovou) jako rozdělení pokusů se samotným šumem, střed tohoto rozdělení je nicméně na ose x posunut k vyšším hodnotám. Na Obr. 5 je rozdělení vnitřních stavů odpovídajících pokusům signál + šum znázorněno křivkou umístěnou více napravo, s průměrem v bodě μ_2 .

Zdůrazněme pro jistotu ještě jednou, že hodnoty na ose x odpovídají subjektivnímu vyhodnocení vjemů (případně složitějších podnětů). U jedinců, kteří jsou vystaveni zcela identickým podnětům, pak mohou být tyto křivky normálního rozdělení od sebe různě vzdálené. Rozdíl hodnot μ_1 , μ_2 , označovaný symbolem d' , pak odpovídá senzitivitě, tedy schopnosti daného jedince od sebe oba druhy situací (tedy pouze šum, vs. šum + signál) správně rozlišovat. Čím více se křivky překrývají, resp. čím menší je hodnota $\mu_1 - \mu_2$, tím hůře je posuzovaná osoba schopna detekovat signál na pozadí šumu. Extrémním případem je situace, kdy $\mu_1 = \mu_2$, tehdy je senzitivita posuzované osoby nulová a její odpovědi napříč jednotlivými pokusy jsou zákonitě zcela náhodné.

Věc je ovšem dále komplikována tím, že ve skutečnosti se probandi nerozhodují přímo na základě hodnot μ_1 , μ_2 , ani na základě jejich rozdílu d' . SDT předpokládá, že každý jedinec má stanovené určité vnitřní kritérium, tedy úroveň subjektivně vnímané intenzity podnětu, na obr. 5 je toto kritérium označené bodem K. Pokud je subjektivně vyhodnocená intenzita

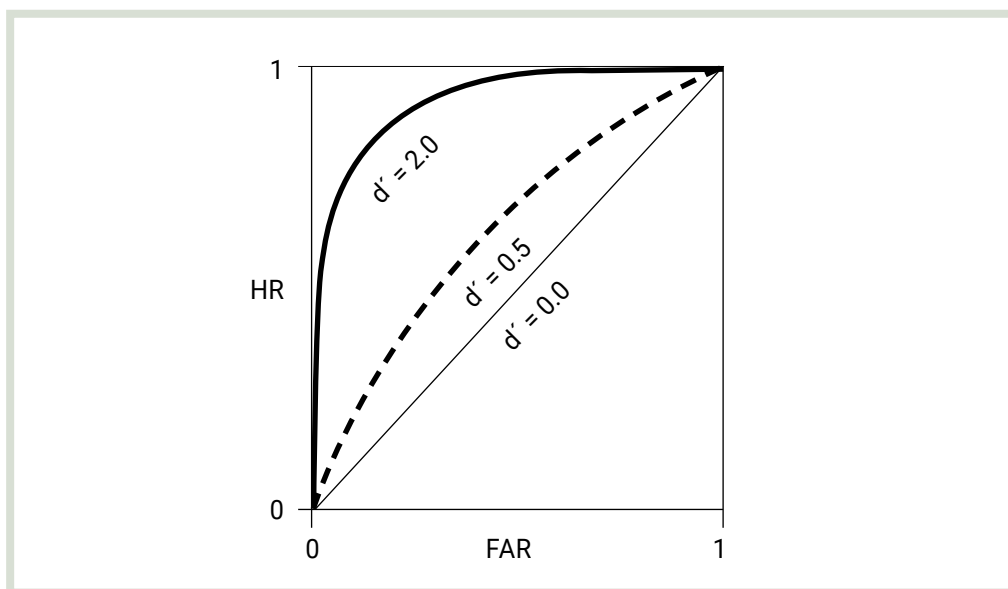
²⁶ Je nicméně vhodné zmínit, že tento předpoklad je v mnoha situacích značně diskutabilní. Byly proto navrženy komplexnější modely, které s tímto předpokladem nepracují. V této publikaci se jim dále věnovat nebudeme, systematický výklad nalezneme čtenář např. v práci Wickense (2002).

podnětu menší než kritérium K , považuje to posuzovaná osoba za absenci signálu, pokud je naopak vyšší, vyhodnocuje ji jako příznak přítomnosti signálu. V „neutrálním“ případě, kdy pokusná osoba není ani osobnostně disponována k určitému typu odpovědi (tj. liberálnějšímu nebo konzervativnějším), ani není v tomto ohledu nějak experimentálně podnícena, se obvykle předpokládá umístění kritéria v polovině vzdálenosti $\mu_1 - \mu_2$ (právě tuto situaci vidíme na Obr. 5). V některých případech ovšem kritérium v této střední poloze být umístěné nemusí. Pokud je například proband osobnostně spíše impulzivní, bude mít zřejmě tendenci detekovat přítomnost signálu celkově častěji, hodnota jeho kritéria bude na ose x posunuta doleva (bod $K1$), u celkově opatrnějšího či úzkostnějšího člověka tomu bude zřejmě naopak. Hodnotou kritéria lze navíc i cíleně experimentálně manipulovat. Informujeme-li třeba pokusnou osobu, že je celkově důležitější detekovat všechny situace, kdy je signál přítomen, a že příliš nevádí, pokud dojde někdy k falešnému poplachu, povede to s velkou pravděpodobností k posunu kritéria směrem k liberálnějším hodnotám. I velikost tohoto posunu je možné do určité míry ovládat – např. výškou odměny, kterou proband obdrží za každý zásah (za předpokladu, že se naopak nepokutují falešné poplachy) atp. Tyto posuny kritéria pak logicky ovlivňují hodnoty HR a FAR.

Povšimněme si, že tendence člověka odpovídat liberálněji či konzervativněji – ať již vychází z jeho osobnostního nastavení, nebo je navozena uměle – se na konceptuální rovině velmi podobá biasu, jak jsme o něm dosud pojednávali v souvislosti s metakognicí. V případě SDT mohou dvě osoby se stejnou mírou schopnosti (senzitivitou) dosahovat jiných hodnot HR a FAR díky rozdílné tendenci odpovídat liberálně či konzervativně. V kontextu metakognitivního monitorování mohou analogicky dvě osoby, jejichž míra relativní přesnosti (tedy také určitá schopnost) je totožná, podávat odlišná hodnocení svého vlastního výkonu v závislosti na tendenci se přeceňovat či podceňovat. Fakt, že v rámci SDT byl nalezen způsob, jak vyčíslit hodnotu senzitivity nezávisle na sklonu člověka odpovídat liberálnějším či konzervativnějším způsobem, pak byl pobídkou k využití týchž analytických postupů i při zjišťování metakognice.



Uvažujme, že s konkrétním jedincem bylo provedeno několik výzkumných sezení, která se lišila experimentálně navozenou tendencí odpovídat liberálněji či konzervativněji (např. v jednom sezení získal za každý zásah odměnu 1 Kč, v dalším sezení odměnu 5 Kč, v jiném byly naopak pokutovány falešné poplachy, ještě v jiném nebyl ani odměňován, ani pokutován atp.). Pro téhož člověka, u kterého předpokládáme konstantní úroveň posuzované schopnosti (senzitivitu) napříč všemi sezeními, tak získáme obecně různé hodnoty HR a FAR. Jednotlivé dvojice HR, FAR za každé sezení lze graficky vyjádřit způsobem, který je znázorněn na Obr. 6, kde na osu x vynášíme hodnoty FAR a na osu y hodnoty HR.



Obrázek 6.

Ukázka křivek ROC, které odpovídají odlišné míře senzitivity. Při nulové senzitivě splyne křivka ROC s diagonálou spojující bod 0,0 s bodem 1,1. Obě zbývající křivky odpovídají osobám s nenulovou mírou senzitivity, senzitivita osoby, jejíž ROC je zobrazena plnou čarou, je nicméně výrazně vyšší než senzitivita osoby, jejíž křivka ROC je znázorněna čarou čerchovanou.

Jak vidíme na Obr. 6, jednotlivými body (tj. dvojicemi hodnot HR, FAR z jednoho sezení) lze pak analyticky proložit křivku, která se v terminologii teorie detekce signálu označuje jako ROC křivka²⁷. Dále lze spočítat, jaká část plochy celého grafu se nachází pod touto křivkou (poznamenejme, že graf má rozměry 1x1 a tedy i jednotkovou plochu). Tento údaj, tedy plocha pod křivkou, se standardně označuje anglickou zkratkou AUROC (area under ROC). Hodnota plochy pod křivkou vyjadřuje senzitivitu jedince, tedy jeho schopnost rozlišovat mezi přítomností a nepřítomností signálu. Pokud je křivka ROC totožná s diagonálou, spojující levý dolní a pravý horní roh grafu, je plocha pod křivkou rovna hodnotě 0,5; to značí nulovou senzitivitu jedince, všechny jeho odpovědi ve všech sezeních a ve všech pokusech

²⁷ Jedná se o zkratku z anglického termínu „receiver operating characteristic“, tedy – velmi volně přeloženo – „charakteristika osoby obsluhující přijímač“. Termín upomíná na počátky SDT, kdy byl tento přístup rozpracován právě proto, aby s jeho pomocí mohly být vybírány osoby obsluhující letecké radary, které pro tuto činnost měly nejlepší předpoklady. Zkratka ROC se pak natolik vžila, že se standardně používá prakticky ve všech aplikacích SDT, přestože většina z nich nemá již s jejím původním významem nic společného.

v rámci těchto sezení jsou čistě náhodné. Čím víc se křivka ROC odchyluje od této diagonály směrem k levému hornímu rohu, tím je senzitivita osoby větší. Např. senzitivita probanda, jemuž na Obr. 6 odpovídá ROC křivka zobrazená plnou čarou, ze zjevně vyšší než senzitivita probanda, jehož ROC křivka je tvořena čarou čerchovanou. Podstatné je, že hodnota AUROC vyjadřuje senzitivitu probanda nezávisle na jeho sklonu hodnotit podněty liberálněji či konzervativněji. Tato tendence ovlivní pouze polohu konkrétního bodu, tj. konkrétní dvojici hodnot FAR, HR na křivce odpovídající schopnostem daného probanda, ale nikoli velikost plochy pod křivkou.



Model, na kterém jsme si demonstrovali základní principy teorie detekce signálu, se velice blížil původnímu zaměření SDT. Jak už ale bylo jednou zmíněno, metodologii původně vyvinutou pro výzkum psychologie vnímání lze použít i na mnoho jiných, často velmi komplexních kognitivních procesů. Pokud například zkoumáme dlouhodobou paměť, můžeme experiment navrhnout tak, že v prvním sezení prezentujeme probandům určitou sadu fotografií. Při druhé části experimentu, která může proběhnout i s odstupem několika dní, prezentujeme jinou sadu fotografií, přičemž část z nich byla již obsažena v první sadě a část je zcela nová. V takovém případě je analogií přítomnosti signálu fakt, že fotografie byla zahrnuta již v první sadě, pokud nebyla, představuje to analogii nepřítomnosti signálu. Při vyhodnocení je přítom ale možné použít tytéž analytické postupy, vytvořené původně pro studium vnímání (včetně stanovení křivky ROC, hodnoty AUROC aj.). Výzkumný design, který se tematicky více blíží studiu nadaných dětí může také vypadat tak, že žáci posuzují předložené matematické příklady, ke kterým zároveň experimentátor poskytuje i řešení. Část těchto řešení může být správná (analogie přítomnosti signálu), část může být záměrně chybná (odpovídá nepřítomnosti signálu).

V těchto a podobných případech, kterých je možné v psychologii nalézt skutečně mnoho, je ovšem SDT použita jako nástroj pro analýzu určité *kognitivní* schopnosti. Sám o sobě tento typ analýzy nevypovídá nic o *metakognici* jedince. Proto, aby mohl být přístup teorie detekce signálu využit k měření metakognitivního monitorování, musel být dále upraven a rozvinut. Některé základní principy, které se uplatnily později, nastínili ji v 50. letech minulého století Clarke a kolegové (1959), na ně poté navázali Galvinová a kolegové (2003), Evans a Azzopardi (2007) a především Maniscalco a Lau (2012). Tito autoři označují „klasický“ SDT přístup, jak jsme si jej dosud přiblížili, jako tzv. *model teorie detekce signálu prvního typu* (type 1 signal detection theory, dále též SDT1). Jeho modifikace, která umožní měřit a analyzovat i metakognitivní monitorování, spočívá v tom, že k základní (kognitivní) úloze – např. rozpoznat přítomnost či nepřítomnost, správnost či nesprávnost atd., určitého podnětu – je následně přidružen dodatečný úkol: provést metakognitivní odhad, tj. určit míru jistoty, jakou pociťuje proband ohledně svého řešení v kognitivní úloze, případně posoudit míru obtížnosti kognitivní úlohy (při využití metriky typu EoSJ u osob, které neovládají koncept pravděpodobnosti) a podobně.

S těmito metakognitivními odhady lze pak provádět obdobné analýzy a počítat obdobné metriky, jako v případě „klasické“ SDT1. Celý tento přístup, zaměřený na metakognici místo na kognici, pak výše uvedení autoři označují jako *model teorie detekce signálu druhého typu* (type 2 signal detection theory). Analogické indikátory, metriky apod. se pak označují stejnými symboly či zkratkami, k nimž je však pro rozlišení připojena číslovka 2 (např. HR2, AUROC2). V případě SDT1 jsme analyzovali vztah mezi kritériem, kterým byla (obvykle) reálná přítomnost či nepřítomnost určitého podnětu a skutečností, zda proband přítomnost či nepřítomnost podnětu správně detekoval. U SDT2 je kritériem správnost řešení primární SDT1 úlohy – tedy to, zda odpověď probanda v samotné kognitivní úloze spadá do kategorie zásah nebo správné odmítnutí, a nikoli do kategorie falešný poplach či minutí. S tímto kritériem pak porovnááme údaj o tom, nakolik si byl správností své odpovědi v kognitivní úloze proband jistý.

Analogií poměru zásahů (HR), se kterým pracuje teorie detekce signálu prvního typu, je v případě SDT2 indikátor $p(\text{míra jistoty probanda} \mid \text{správná řešení})^{28}$, tedy proporce případů, kdy proband zodpověděl kognitivní úkol správně a zároveň si byl jistý správností svého řešení, ze všech případů, kdy zodpověděl kognitivní úkol správně. Ve shodě s Maniscalcem a Lauem (2014) budeme dále tuto hodnotu označovat zkratkou HR2 (type 2 hit rate). Obdobně lze definovat analogii k poměru falešných poplachů, a to jako proporcii případů, kdy proband vyřešil kognitivní úlohu chybně a byl si přitom jistý správností své odpovědi, oproti celkovému počtu všech chybně vyřešených položek v základní kognitivní úloze, tedy $p(\text{míra jistoty probanda} \mid \text{nesprávná řešení})^{29}$. Tuto hodnotu budeme dále označovat jako FAR2 (type 2 false alarm rate).

Tabulka 7.

Ukázková data pro ilustraci vztahů mezi SDT1 a SDT2

Kód položky	Signál reálně přítomen [ANO/NE]	Proband detekoval přítomnost signálu [ANO/NE]	Proband si je jistý správností své detekce [ANO/NE]
1	ANO	ANO	ANO
2	NE	ANO	NE
3	NE	NE	NE
4	ANO	ANO	NE

²⁸ V angl. originále $p(\text{confidence} \mid \text{correct})$

²⁹ V angl. originále $p(\text{confidence} \mid \text{incorrect})$

Celý princip si opět pro názornost ukážeme na krátkém praktickém příkladu, který je zachycený v Tabulce 7. Poměrně neproblematická je položka 1 – zde byl reálně signál přítomen, a subjekt také jeho přítomnost správně detekoval. Z pohledu SDT1 se tedy jedná o zásah. Při následném metakognitivním odhadu si byl proband jistý správností své odpovědi, vzhledem k tomu, že odpověď skutečně byla správná, jedná se o zásah i z pohledu SDT2. Zajímavější je situace u položky č. 2. Zde signál nebyl přítomen, subjekt ale jeho přítomnost chybně detekoval, z pohledu SDT1 jde tedy o nesprávné řešení – falešný poplach. Zároveň si ale – jak vidíme z hodnoty ve třetím sloupci, nebyl správností svého řešení jistý. V určitém smyslu tedy položka zachycuje adekvátní kvalitu metakognitivního monitorování – subjekt sice odpovídá nějakým způsobem, na určité úrovni ovšem „tuší“, že odpověď je problematická. Protože řešení položky je chybné a subjekt zároveň udává nejistotu, jedná se z pohledu SDT2 o zásah. Položky 3 a 4 pak představují opačný případ, tj. řešení je z pohledu SDT1 správné, ale následně metakognitivní hodnocení udává nízký stupeň jistoty, z pohledu SDT2 jde tedy o minutí.

AUROC2

Pokud by byla míra jistoty vyjádřená pouze ve dvou kategoriích, tak jako v uvedeném příkladu (takže subjekt by mohl například volit pouze mezi možnostmi „jsem si jistý/á“ vs. „nejsem si jistý/á“), umožňovalo by to určit pouze jedinou dvojici hodnot HR2, FAR2, což by nepostačovalo k jednoznačnému určení křivky ROC druhého typu, a tedy ani k určení plochy pod touto křivkou. Existuje ovšem poměrně snadný postup, jak toto omezení obejít (Fleming & Lau, 2014). Podmínkou pro jeho uplatnění je to, aby zkoumaná osoba vyjadřovala svoje metakognitivní hodnocení nikoli pouze ve dvou kategoriích, nýbrž na širší posuzovací likertovské škále. Škála přitom nemusí vyjadřovat míru jistoty v procentech, tento druh analýzy lze proto provádět i s ratingy obtížnosti typu EoSJ a podobně.

Pro ilustraci postupu předpokládáme, že proband vyjadřuje u jednotlivých úloh míru jistoty vztaženou ke zvolenému rozhodnutí na škále 0-9, kde 0 značí nejmenší míru jistoty a 9 naopak míru největší. Výzkumník pak postupuje následujícím způsobem: v prvním kroku rozdělí odpovědi probanda do dvou kategorií (nízká úroveň jistoty, vysoká úroveň jistoty) tak, že za položky s nízkou mírou jistoty jsou považovány pouze ty, u nichž proband uvedl hodnocení 0, všechna ostatní hodnocení v rozsahu 1-9 jsou považovány za vysokou míru jistoty. Na základě takto kategorizovaných položek je pak možné jednoduše spočítat první dvojici hodnot HR2, FAR2. Ve druhém kroku jsou za nízkou míru jistoty považována hodnocení 0, 1; za vysokou míru jistoty pak hodnocení v rozsahu 2-9. Tato nová kategorizace umožní spočítat další dvojici hodnot HR2, FAR2, která se od první bude obecně lišit, a to kvůli odlišnému zařazení položek, u nichž subjekt ohodnotil svoji jistotu ratingem 1. V dalším kroku bude následovat dělení 0-2 vs. 3-9 a tak dále, až k poslednímu kroku, v němž budou za vysokou míru jistoty považována pouze hodnocení, u nichž proband vyjádřil míru své jistoty hodnotou 9. V tomto konkrétním případě tedy můžeme získat 10 různých dvojic HR2, FAR2. Následně již lze využít běžné postupy rozvinuté v rámci SDT1, určenými body proložit křivku ROC,

a spočítat hodnotu plochy pod touto křivkou, tzv. AUROC2. Tato hodnota je vyjádřením *metakognitivní senzitivity* posuzované osoby, a nezávisí na jejím případném globálním biasu (Fleming & Lau, 2014).

AUROC2 má ovšem bohužel jiné podstatné omezení. Aby bylo možné jednotlivé osoby s použitím této metriky korektně porovnávat, vyžaduje se, aby jednotlivé úkoly, použité pro testování určité kognitivní schopnosti v primární SDT1 analýze, byly z pohledu všech probandů zhruba stejně obtížné. Není-li tento požadavek dodržen, může být hodnota AUROC2 u různých osob různým způsobem zkreslena, a neexistuje žádný jednoduchý a obecně použitelný postup pro korekci tohoto zkreslení. Pokud je metakognitivní monitorování studováno v kontextu psychologie vnímání či v neurovědním výzkumu, nepředstavuje toto omezení zásadní problém, protože v těchto případech lze podněty celkem snadno upravit tak, aby zmíněný požadavek splňovaly. Fleming a kolegové (2010) například ve studii zaměřené na lokalizaci funkčních systémů zodpovědných za metakognitivní monitorování v mozku pracovali se zrakovými podněty, které se vzájemně lišily svým kontrastem. Intenzita tohoto kontrastu byla pro každého probanda nastavena tak, aby byl v průměru úspěšný (z pohledu SDT1) zhruba v 71 % případů. Tím bylo zajištěno, že všichni probandi řešili úkoly, které pro ně byly relativně stejně obtížné.

Splnit ovšem podobnou podmínku v pedagogickém či v psychologickém výzkumu orientovaném na metakognici ve vztahu ke komplexnějším kognitivním funkcím (jako je logické myšlení, vybavování informací z dlouhodobé paměti na základě určitého sémantického kritéria a podobně) se jeví jako daleko obtížnější. Teoreticky by bylo možné – s použitím Raschova modelu nebo jiného IRT přístupu – vytvořit databanku položek o známé obtížnosti a zároveň provést dopředu u každého participanta screening úrovně jeho schopností ve zkoumané oblasti. Jednotlivým participantům by pak také mohly být například matematické příklady vybírány „na míru“, aby bylo dopředu jasné, že každý proband má každou zadanou položku šanci vyřešit se zhruba stejnou, předem stanovenou pravděpodobností. Takový postup by byl ale časově i organizačně extrémně náročný a autorovi této publikace není znám žádný případ, kdy by se někdo o něco podobného pokusil, ať již u nás, či v zahraničí. Naštěstí to ani není nezbytné, protože pro posouzení senzitivity metakognitivního monitorování při měnící se obtížnosti podnětových úkolů existuje nyní jiná vhodná metrika, rovněž konceptuálně vycházející z teorie detekce signálu – tzv. meta-d', již bude věnována následující kapitola.

Meta-d'

Jak jsme již viděli dříve, u analýz v rámci SDT1 se předpokládá, že vnitřní stavy subjektu, které jsou výsledkem jeho zpracování určitého externího podnětu a na jejichž základě se subjekt rozhoduje mezi přítomností a nepřítomností signálu, mají normální rozdělení o stejném rozptylu. Vzdálenost mezi středními hodnotami pro rozdělení odpovídající pouze šumu a pro rozdělení odpovídající šumu + signálu – označovanou symbolem d' (senzitivitu), lze

chápat jako vyjádření schopnosti vnímání, případně jiné, komplexnější, kognitivní schopnosti. Pro analýzy vycházející z SDT2 by se tedy zdánlivě nabízel stejný postup, který ovšem naráží na závažný metodologický problém. Jak zmiňují Fleming a Lau (2014), v případě SDT1 je předpoklad výše zmíněných normálních rozdělání s určitými výhradami akceptovatelný, u SDT2 pro něj ovšem chybí přesvědčivé teoretické zdůvodnění. Galvinová a kolegové (2003) navíc ukázali, že tyto distribuce se ve skutečnosti výrazně odchyľují od normality a zpravidla se výrazně liší i svým rozptylem. Aplikace standardních výpočetních postupů vyvinutých v rámci SDT1 za účelem stanovení hodnoty d' na metakognitivní data by pak vedla k výrazně zkresleným hodnotám, což naznačila například i simulační studie, kterou provedli Evans a Azzopardi (2007). Tento problém lze ale naštěstí obejít postupem, který rozpracovali Maniscalco a Lau (2012). Ten využívá skutečnosti, že pokud známe parametry rozdělání „vnitřních stavů subjektu“ získané v rámci SDT1 analýz, lze z nich odvodit dostatečné omezující podmínky pro rozdělání analyzovaná v rámci SDT2. Při znalosti těchto omezení stačí pak údaje o reálném výkonu probanda při metakognitivním hodnocení (tj. jediná dvojice hodnot HR2, FAR2), aby bylo možné určit jeho metakognitivní senzitivitu.

Jak už bylo jednou zmíněno, přesný a úplný popis jednotlivých výpočtů je poměrně komplikovaný, případně zájemce lze v tomto směru odkázat např. na vyčerpávající výklad v publikaci Maniscalco a Lau (2014). Logiku, která za tímto výpočtem stojí, si nicméně můžeme přiblížit na extrémním, zároveň ale velmi názorném a intuitivním příkladu. Uvažujme člověka, který má tak nízkou úroveň kognitivní schopnosti nutně pro splnění primárního úkolu (analyzovaného prostřednictvím SDT1), že jeho odpovědi jsou prakticky náhodné (což se mimo jiné projeví na hodnotě plochy pod křivkou ROC prvního typu, která bude zhruba 0,5). V takovém případě musí být zákonitě i jeho metakognitivní senzitivita vztažená k tomuto úkolu nulová, protože pokud v primární úloze „určil“ některé úlohy správně pouze díky náhodě, nemůže v této souvislosti ani monitorovat a reflektovat průběh své kognitivní schopnosti (která přitom do procesu „řešení“ prakticky nevstoupila). Potenciální úroveň metakognitivní senzitivity je tedy zdola omezená úrovní senzitivity v primární SDT1 úloze.

Obecně lze na základě takových úvah definovat celou rodinu křivek ROC druhého typu, z nichž každá odpovídá jiné předpokládané hodnotě SDT1 senzitivity za předpokladu, že posuzovaný jedinec je tzv. metakognitivně ideální (to znamená, že disponuje maximální mírou SDT2 senzitivity, kterou mu „dovoluje“ jeho úroveň SDT1 senzitivity). Zároveň můžeme z poskytnutých odhadů probanda spočítat jednu dvojici hodnot HR2, FAR2. Díky tomu můžeme určit, která z rodiny hypotetických křivek ROC2 skutečně vyjadřuje metakognitivní senzitivitu tohoto probanda. Znalost konkrétní křivky pak umožňuje určit metriku označovanou jako $meta-d'$. Její přesný význam je následující: jedná se o senzitivitu, jakou lze u jedinice předpokládat na škále vztahující se ke kognitivní schopnosti posuzované v rámci primární SDT1 analýzy, pokud známe empiricky zjištěné hodnoty HR2, FAR2, za předpokladu, že tento jedinec je metakognitivně ideální. Přestože se tato metrika vyjadřuje na stejné škále, jako je ta použitá v rámci SDT1, nevyjadřuje $meta-d'$ míru kognitivní, nýbrž *metakognitivní* senzitivity, tedy zjednodušeně řečeno schopnosti provádět metakognitivní monitorování.

Při nastíněném postupu je ovšem možné provést i veškeré standardní SDT1 analýzy a získat tak i hodnotu d' – tedy skutečnou míru senzitivity týkající se primární kognitivní úlohy. Obě hodnoty (tedy d' a meta d') je pak možné vzájemně porovnat. Podíl obou hodnot – meta- d'/d' , se označuje jako *metakognitivní efektivita* (metacognitive efficiency). Tato metrika může nabývat hodnot v rozsahu 0-1, přičemž hodnota 1 značí metakognitivně ideálního posuzovatele, ve smyslu uvedeném výše. Jak meta d' , tak i odvozený podíl meta d'/d' nejsou zkresleny ani rozdíly v obtížnosti podnětových úloh, ani celkovým biasem posuzované osoby. Fleming a Lau (2014) je z tohoto důvodu považují za v současnosti nejvhodnější ukazatele při měření metakognitivního monitorování.

Meta d' a odvozené metriky³⁰ mají několik velkých výhod. Za prvé představují velmi dobrý podklad pro korektní porovnávání metakognitivního monitorování ve vztahu k rozdílným kognitivním schopnostem, v různých doménách a v odlišných typech úkolů. Jako příklad můžeme uvést nedávnou studii Mazancieuxové a kolegů (2020), která srovnávala metakognitivní monitorování u týchž osob v úlohách zaměřených na epizodickou paměť, sémantickou paměť, vizuální percepci a pozornost. Tento typ výzkumů má velký potenciál do budoucnosti, což už naznačila samotná zmíněná studie. Její závěry lze s určitou opatrností interpretovat tak, že možná existuje jeden společný mechanismus, který se uplatňuje při metakognitivním monitorování zacíleném na široké spektrum různých kognitivních schopností.

Druhý velký přínos představuje skutečnost, že s využitím meta- d' lze snadno a metodologicky korektně porovnávat metakognitivní schopnosti mezi skupinami, které se výrazně liší v úrovni své kognice. Takto byly například porovnávány osoby lišící se věkem, nebo pacienti s kontrolní skupinou zdravých osob (Fleming & Lau, 2014). V podobném duchu můžeme předpokládat, že meta d' bude sloužit v budoucnu jako vhodná metrika pro analýzy srovnávající nadané děti s běžnou populací. Zatím ale zůstává tato možnost spíše příslibem do budoucna, aktuálně v tomto typu výzkumů meta d' běžně používáno není.

I přes uvedené přednosti je nutné závěrem zmínit, že i meta d' má přinejmenším jednu podstatnou nevýhodu: tuto metriku lze použít pouze v návaznosti na podnětové úkoly, které mohou být standardně analyzovány v rámci SDT1 (Vuorre & Metcalfe, 2021). Musí se tedy jednat o úlohy, při nichž se proband i v prvním kroku (tj. v samotné kognitivní úloze) *rozhoduje*, zpravidla mezi dvěma alternativami. Meta d' tedy nelze odvodit např. z metakognitivních odhadů vztahujících se k otázkám s volně tvořenou odpovědí. Vzhledem k tomu, že úlohy s nucenou volbou nemohou ekologicky validním způsobem postihnout řadu jevů důležitých pro pedagogickou praxi, nepředstavuje ani meta d' nějaké ideální či univerzální řešení, které by mohlo bez dalšího nahradit všechny ostatní užívané metriky.

³⁰ Kromě podílu meta d'/d' se někdy používá také rozdíl obou hodnot, tedy meta $d' - d'$ (Fleming & Lau, 2014).

ZÁVĚREM

Cílem této monografie bylo představit a kriticky zhodnotit základní metody používané pro měření metakognice, zejména s perspektivou využití u intelektově nadaných dětí. Nejedná se přitom o vyčerpávající přehled všech existujících metod. V knize byla záměrně pominuta řada velmi specializovaných technik či postupů, jako je např. paradigma detekce chyb, opt-out paradigma či úlohy typu CLOZE. Tyto postupy jsou využívány téměř výhradně experimentálně, a přestože umožnily a stále umožňují dobrat se k důležitým a zajímavým poznatkům, je jejich zaměření natolik úzké, že i v pozici odborníka v pedagogice či psychologii se lze většinou bez jejich znalosti obejít. Domnívám se, že zaměření na nejpodstatnější a nejrozšířenější metody a postupy učinilo studii přehlednější a srozumitelnější.

Relativně největší část publikace byla věnována specifickým metrikám a indexům, využívaným při měření metakognitivního monitorování. Pro zvláštní důraz na toto téma existují přinejmenším dva dobré důvody. Jednak je tato problematika složitá a v našem prostředí poměrně málo známá. Pro odborníky zabývající se pedagogickou či psychologickou diagnostikou by proto prezentované informace mohly být v mnoha ohledech obohacující. Především je ale výhodné, že měření metakognitivního monitorování obtížnosti není vázané na žádný konkrétní test či diagnostický nástroj. Tím, že lze za tímto účelem snadno upravit většinu běžně používaných didaktických, znalostních či jiných výkonových testů, představuje využití některé z prezentovaných metrik ve spojení s vybraným testem zřejmě nejrychlejší a nejsnazší cestu, jak v našich podmínkách prakticky diagnostikovat či zkoumat metakognici (resp. pouze jeden, nicméně velmi důležitý aspekt metakognice). Nutnou podmínkou pro úspěch takového snažení je pochopitelně volba kvalitního „podkladového“ nástroje – tedy výkonového testu, na jehož položkách budou zároveň prováděny metakognitivní odhady. Dostupnost takových nástrojů je v českém prostředí zřejmě nicméně přece jen lepší, než dostupnost metod cíleně vyvinutých za účelem měření metakognice.

Přidržíme-li se konkrétních nástrojů a metod, které byly v knize prezentovány, můžeme se na závěr pokusit o jejich celkové zhodnocení ze dvou hledisek: za prvé z pohledu potenciálního uživatele – pedagoga či psychologa, který uvažuje o využití nějakého nástroje pro zjišťování metakognice ve své praxi; za druhé z pohledu odborníků, kteří se zabývají pedagogickou či psychologickou diagnostikou a mohli by uvažovat o adaptaci některého nástroje v českém prostředí, případně o vývoji obdobné, zcela nové metody.

Pedagog či psycholog, který by rád prováděl diagnostiku metakognitivních schopností u svých žáků či klientů, má v českém prostředí aktuálně velmi omezené možnosti. Pokud jde o sebeposuzovací metody, tedy dotazníky a strukturované rozhovory, nelze reálně uvažovat o používání zahraničních metod bez předcházející standardizace na naši žákovskou populaci. Tato zásada platí obecně pro většinu psychodiagnostických nástrojů, v případě metakognitivních dotazníků je ovšem ještě naléhavější. Vzhledem k tomu, nakolik je proces porozumění jednotlivým položkám náročný jak na pozornost, tak na úsudek probandů, může i drobný významový posun v překladu vést k diametrálně odlišným výsledkům v porovnání s originální verzí. Rozhodně není na místě používat pouze přeložené a dále neověřované verze zahraničních metod. Dotazníky, které byly vyvinuty v anglosaském prostředí a které jsou tam často používány – např. zmiňované MAI či JMAI, českou verzi aktuálně nemají. Výjimečným příkladem zahraniční metody, která byla standardizována na českou populaci, je dotazník MAESTRA. Metoda nebyla v této knize blíže pojednána vzhledem k tomu, že mnohem obsáhlejší a detailnější popis je k dispozici přímo od autorů české verze (Chytrý et al., 2014).

Pokud jde o on-line metody, ani zde není situace ideální. Různé softwarové nástroje, které registrují projevy metakognitivní aktivity, nejsou v ČR volně ani komerčně dostupné v takové podobě, která by umožnila jejich praktické použití bez nutnosti dalších úprav. Eye-tracking zůstává metodou spíše experimentální, a navíc stále finančně náročnou s ohledem na potřebné vybavení. Jako jediná alespoň trochu reálná možnost se tak jeví využití protokolové analýzy, případně metody stimulovaného vybavení. Zahraniční zkušenosti ukazují, že analýza TAP nemusí být využita pouze výzkumně, nýbrž bylo popsáno i její úspěšné praktické uplatnění. Čtenáře se zájmem o bližší informace lze odkázat např. na práce Randallové a kolegyň (1986) či Cowana (2019).

Pokud jde o perspektivy adaptace či vývoje nové metody, je na prvním místě nutné zvážit, nakolik se vyplatí investovat čas a finance do adaptace dotazníků, s ohledem na zmíněné problémy s jejich validitou. Ty jsou na základě zahraničních zkušeností nepochybné a závažné. Proto lze soudit, že adaptovat některé z často užívaných dotazníků „tak jak jsou“, tedy s podobnou skladbou položek, se současným formátem odpovídání atd., by mohlo být riskantní z pohledu vynaložených financí i práce. Na druhou stranu se ale nedomnívám, že by na základě dosavadních zahraničních zkušeností bylo na místě sebeposuzovací metody jako nástroj na měření metakognice zcela zavrhnout.

Za prvé, jak vyplývá z analýzy Craigové a spolupracovníků (2020), posuzované dotazníky sice mají obecně problémy s validitou, nikoliv však nutně s reliabilitou. Řada publikovaných dotazníků tedy stabilně a spolehlivě měří „něco“, ono „něco“ je ovšem spíše subjektivní představa jedince o vlastních metakognitivních strategiích a nikoli to, jak tyto strategie reálně využívá. Jestliže bychom měli pro téhož jedince k dispozici výsledek dotazníku a zároveň výstup z nějaké on-line metody, mohl by mít právě rozpor (nebo naopak soulad) mezi oběma údaji sám o sobě diagnostickou hodnotu. Pokud je například student vnitřně přesvědčen, že využívá optimálně strategie, které usnadňují učení, ale ve skutečnosti tomu tak není, může právě zde ležet příčina jeho studijní neúspěšnosti, kvůli níž vyhledal poradenskou pomoc.

Za druhé, i při tvorbě metakognitivních dotazníků lze využít zatím méně obvyklé, inovativní postupy. Například výše zmíněný dotazník MAESTRA (Chytrý et al., 2014) je konstruován tak, že žák nejprve řeší nějaký problémový úkol (v tomto případě v podobě matematického příkladu) a následně sebeuposuzující otázky jsou vztaženy k tomuto konkrétnímu úkolu. To je velký rozdíl od mnoha tradičních metakognitivních dotazníků, které, jak jsme viděli na uvedených příkladech, sestávají z obecných a pro žáky tak mnohdy obtížně zodpověditelných otázek. Pro nějaké silnější závěry bude třeba více zkušeností s podobnými inovativními postupy, je ale možné, že právě tudy vede cesta k validnějším a tím i v praxi použitelnějším off-line metodám pro diagnostiku metakognice.

Domnívám se ovšem, že ze všech představených metod vycházejí nakonec – po zvážení všech výhod i nevýhod – jako nejlepší alternativa počítačové metody registrující činnost probandů při řešení zadaných úloh. Pokud by vznikl nějaký nástroj tohoto druhu, který by byl dostupný v českém jazyce, byly by k němu k dispozici normy pro místní populaci a případně by i zohledňoval obsah kurikula na školách v ČR, nepochybně by zde našel široké uplatnění jak ve výzkumu, tak v pedagogicko-psychologické praxi.



SUMMARY

The main purpose of this book is to present and critically appraise the most important and most frequently used methods and procedures designed to assess metacognitive abilities, with special emphasis on measuring these abilities in intellectually gifted children. The monograph is supposed to provide a clear and comprehensive overview of these methods for the use of Czech psychologists, teachers and educational specialists, who are in need of planning their own research or of assessing the metacognitive abilities of a concrete child within their counselling practice.

At the beginning, a short review of recent publications bearing on the topic of metacognition and published in the Czech language is presented and it is argued why the assessment of metacognition is essential in the field of developmental psychology and educational science. At the same time, the construct of metacognition itself, along with the widely used taxonomical system for classification of metacognitive phenomena is briefly explained. A separate chapter is also devoted to the issue of intellectual giftedness and its relation to metacognition.

Particular methods and diagnostic tools are presented in keeping with the classification of these methods, which is based on whether the data are registered during the execution of the stimulus task (on-line methods), or whether they are garnered with a certain time lag (off-line methods).

As to the on-line methods, the following approaches are presented: the analysis of think-aloud protocols (TAP), the analysis of the participant's activity as registered by means of a computer and the eye-tracking. In view of its theoretical importance, special emphasis is put on the analysis of think-aloud protocols. The origin and development of the technique is described, the caveats and potential limitations are discussed and the respective chapter contains practical illustrations of TAP, both from cognitive and metacognitive research. It also deals with the problem of whether and to what extent the very process of taking the TAP may alter the surveyed metacognitive processes themselves. The off-line methods presented in the book comprise metacognitive questionnaires, the metacognitive interview and the stimulated recall. The metacognitive questionnaires are reflected on most thoroughly, especially with the regard to the serious and long known validity issues. The book also briefly mentions several possibilities of how these limitations could be mitigated or overcome.



The most extensive part of the book deals with the measurement of metacognitive monitoring in a sense narrowed down to the subjective confidence or perceived difficulty ratings. By way of practical examples, the most important and most frequently used measures designed to quantify the ability of metacognitive monitoring are introduced and their advantages and drawbacks are discussed. Both measures of absolute and relative accuracy are treated. The first group comprises absolute accuracy, calibration and bias indices, as well as simple matching coefficient and Hamann coefficient. The latter category encompasses Pearson's r , resolution index and gamma coefficient. The final section of this part is devoted to the recent progress in the domain of measuring metacognitive monitoring, in which the emphasis is on developing and refining measures grounded in the signal detection theory. The logic of the measures, such as AUROC2 and meta- d' is explained, together with the considerations regarding their potential and the prospects of their future use.

At the end, the monograph attempts to draw some general conclusions regarding the suitability of the presented methods and approaches, both for the research and for the psychological and educational practice.

LITERATURA

- Afflerbach, P. (2000). Verbal reports and protocol analysis. In M. L. M. Kamil, P. B. Mosenthal, D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (sv. III, s. 163–179). Erlbaum.
- Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, 125, 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>
- Alexander, J. M., Carr, M., & Schwanenflugel, P. J. (1995). Development of metacognition in gifted children: Directions for future research. *Developmental review*, 15(1), 1–37. <https://doi.org/10.1006/drev.1995.1001>
- Alexander, J. M., Johnson, K. E., Albano, J., Freygang, T., & Scott, B. (2006). Relations between intelligence and the development of metaconceptual knowledge. *Metacognition and Learning*, 1(1), 51–67. <https://doi.org/10.1007/s11409-006-6586-8>
- Allon, M., Gutkin, T. B., & Bruning, R. (1994). The relationship between metacognition and intelligence in normal adolescents: Some tentative but surprising findings. *Psychology in the Schools*, 31(2), 93–97. [https://doi.org/10.1002/1520-6807\(199404\)31:2<93::AID-PITS2310310202>3.0.CO;2-X](https://doi.org/10.1002/1520-6807(199404)31:2<93::AID-PITS2310310202>3.0.CO;2-X)
- Allwood, C. M., Jonsson, A. C., & Granhag, P. A. (2005). The effects of source and type of feedback on child witnesses' metamemory accuracy. *Applied Cognitive Psychology*, 19(3), 331–344. <https://doi.org/10.1002/acp.1071>
- Azevedo, R., & Alevin, V. (Eds.). (2013). *International handbook of metacognition and learning technologies* (sv. 26). Springer. <https://doi.org/10.1007/978-1-4419-5546-3>
- Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29, 344–370. <https://doi.org/10.1016/j.cedpsych.2003.09.002>
- Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally facilitated regulated learning more effective than self-regulated learning with hypermedia?. *Educational Technology Research and Development*, 56(1), 45–72. <https://doi.org/10.1007/s11423-007-9067-0>
- Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning?. *Metacognition and Learning*, 3(1), 39–58. <https://doi.org/10.1007/s11409-007-9009-6>

- Bowles, M. A. (2010). *The think-aloud controversy in second language research*. Routledge.
<https://doi.org/10.4324/9780203856338>
- Bråten, I., & Strømsø, H. I. (2003). A longitudinal think-aloud study of spontaneous strategic processing during the reading of multiple expository texts. *Reading and Writing, 16*(3), 195–218.
<https://doi.org/10.1023/A:1022895207490>
- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F.E. Weinert, R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (s. 65–116). Erlbaum.
- Carr, M., & Taasobshirazi, G. (2008). Metacognition in the gifted: Connections to expertise. In M. F. Shaughnessy, M. V. J. Veenman, & C. Kleyn-Kennedy (Eds.), *Meta-Cognition: A recent review of research, theory, and perspectives* (s. 109–125). Nova Science Publishers.
- Clarke, F. R., Birdsall, T. G., & Tanner, W. P., Jr. (1959). Two types of ROC curves and definitions of parameters. *The Journal of the Acoustical Society of America, 31*(5), 629–630.
<https://doi.org/10.1121/1.1907764>
- Cooper, M. M., Sandi-Urena, S., & Stevens, R. (2008). Reliable multi method assessment of metacognition use in chemistry problem solving. *Chemistry Education Research and Practice, 9*(1), 18–24. <https://doi.org/10.1039/B801287N>
- Cowan, J. (2019). The potential of cognitive think-aloud protocols for educational action-research. *Active Learning in Higher Education, 20*(3), 219–232. <https://doi.org/10.1177/1469787417735614>
- Craig, K., Hale, D., Grainger, C., & Stewart, M. E. (2020). Evaluating metacognitive self-reports: systematic reviews of the value of self-report in metacognitive research. *Metacognition and Learning, 15*, 155–213. <https://doi.org/10.1007/s11409-020-09222-y>
- Cromley, J. G., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring?. *Metacognition and Learning, 1*(3), 229–247. <https://doi.org/10.1007/s11409-006-9002-5>
- Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., ... & Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly, 50*(4), 393–416.
<https://doi.org/10.1002/rrq.105>
- Dori, Y. J., Mevarech, Z. R., & Baker, D. R. (Eds.). (2018). *Cognition, Metacognition, and Culture in STEM Education*. Springer. <https://doi.org/10.1007/978-3-319-66659-4>
- Edwards-Leis, C. (2006). *Variations to stimulated recall protocols to enhance student reflection: I did, I saw, I remembered* [Tištěná verze konferenčního příspěvku]. AARE Annual Conference, Adelaide, Austrálie.
- Elshout, J. J., & Veenman, M. V. (1992). Relation between intellectual ability and working method as predictors of learning. *The Journal of Educational Research, 85*(3), 134–143.
<https://doi.org/10.1080/00220671.1992.9944429>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis. Verbal Reports as Data*. MIT Press.
<https://doi.org/10.7551/mitpress/5657.001.0001>



- Evans, S., & Azzopardi, P. (2007). Evaluation of a 'bias-free' measure of awareness. *Spatial vision*, 20(1), 61–78. <https://doi.org/10.1163/156856807779369742>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, 34(10), 906. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleming, S. M. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of consciousness*, 2017(1), nix007. <https://doi.org/10.1093/nc/nix007>
- Fleming, S. M., & Frith, C. D. (Eds.). (2014). *The cognitive neuroscience of metacognition*. Springer. <https://doi.org/10.1007/978-3-642-45190-4>
- Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in human neuroscience*, 8, 443. <https://doi.org/10.3389/fnhum.2014.00443>
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, 329(5998), 1541–1543. <https://doi.org/10.1126/science.1191883>
- Foltýnová, D. (2009). Vliv metakognitivních strategií na rozvoj dovedností žáků autoregulovat své učení. *Pedagogická orientace*, 19(2), 72–88.
- Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic bulletin & review*, 10(4), 843–876. <https://doi.org/10.3758/BF03196546>
- Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, 5(1), 3–57. <https://doi.org/10.1002/rev3.3077>
- Greene, J. A., Moos, D. C., Azevedo, R., & Winters, F. I. (2008). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education*, 50(3), 1069–1083. <https://doi.org/10.1016/j.compedu.2006.10.004>
- Gridley, B. E., Norman, K. A., Rizza, M. G., & Decker, S. L. (2003). Assessment of gifted children with the Woodcock-Johnson III. In F. A. Schrank & D. P. Flanagan (Eds.), *WJ III Clinical Use and Interpretation* (s. 285–317). Academic Press. <https://doi.org/10.1016/B978-012628982-4/50010-6>
- Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory*, (s. 429–455). Psychology Press.
- Harrison, G. M., & Vallin, L. M. (2018). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and Learning*, 13(1), 15–38. <https://doi.org/10.1007/s11409-017-9176-z>
- Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *Elife*, 6, e24901. <https://doi.org/10.7554/eLife.24901.009>
- Havigerová, J. M. (2011). *Pět pohledů na nadání*. Grada.

- Helms-Lorenz, M., & Jacobse, A. E. (2008). Metacognitive skills of the gifted from a cross-cultural perspective. In M. F. Shaughnessy, M. V. J. Veenman, & C. Kleyn-Kennedy (Eds.), *Meta-Cognition: A recent review of research, theory, and perspectives* (s. 3–43). Nova Science Publishers.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Hrbáčková, K. (2004). Vliv metakognitivních strategií na rozvoj autoregulace učení. *Pedagogická orientace*, 14(1), 81–88.
- Hrbáčková, K. (2011). Vliv metakognitivní intervence na rozvoj myšlení dětí předškolního věku. *E-Pedagogium*, 3, 49–63. <https://doi.org/10.5507/epd.2011.033>
- Hu, J., & Gao, X. A. (2017). Using think-aloud protocol in self-regulated reading research. *Educational Research Review*, 22, 181–193. <https://doi.org/10.1016/j.edurev.2017.09.004>
- Hunt, M. (2015). *Dějiny psychologie*. Portál.
- Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*, 94(1), 44. <https://doi.org/10.1037/0022-0663.94.1.44>
- Chytrý, V., Pešout, O., & Říčan, J. (2014). *Preference metakognitivních strategií na pozadí úkolových situací v matematice u žáků druhého stupně ZŠ*. Univerzita Jana Evangelisty Purkyně v Ústí nad Labem.
- Chytrý, V., Říčan, J., & Živná, D. (2019). Matematická výkonnost a metakognice žáků základních škol běžných, základních škol Montessori a žáků vyučovaných podle Hejného metody. *Studia paedagogica*, 24(1), 107–133. <https://doi.org/10.5817/SP2019-1-5>
- Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high-and low-achieving primary school children's reading. *Learning and Instruction*, 5(2), 143–165. [https://doi.org/10.1016/0959-4752\(95\)00009-R](https://doi.org/10.1016/0959-4752(95)00009-R)
- Kinnunen, R., & Vauras, M. (2010). Tracking on-line metacognition: monitoring and regulating comprehension in reading. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (s. 209–229). Springer. https://doi.org/10.1007/978-1-4419-6546-2_10
- Krykorková, H., & Chvál, M. (2001). Rozvoj metakognice—cesta k hodnotnějšímu poznání. *Pedagogika*, 51(2), 185–196.
- Lajoie, S. P., Poitras, E. G., Doleck, T., & Jarrell, A. (2015). Modeling metacognitive activities in medical problem-solving with BioWorld. In A. Peña-Ayala (Ed.), *Metacognition: Fundamentals, applications, and trends* (s. 323–343). Springer. https://doi.org/10.1007/978-3-319-11062-2_13
- Lam, W. Y. (2008). Metacognitive strategy use: Accessing ESL learners' inner voices via stimulated recall. *International Journal of Innovation in Language Learning and Teaching*, 2(3), 207–223. <https://doi.org/10.1080/17501220802158917>
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know?. *Organizational behavior and human performance*, 20(2), 159–183. [https://doi.org/10.1016/0030-5073\(77\)90001-0](https://doi.org/10.1016/0030-5073(77)90001-0)

- Lingel, K., Lenhart, J., & Schneider, W. (2019). Metacognition in mathematics: do different metacognitive monitoring measures make a difference?. *ZDM*, 51(4), 587–600.
<https://doi.org/10.1007/s11858-019-01062-8>
- Lokajíčková, V. (2014). Metakognice – vymezení pojmu a jeho uchopení v kontextu výuky. *Pedagogika*, 64(3), 287–306.
- Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual Differences in Absolute and Relative Metacomprehension Accuracy. *Journal of Educational Psychology*, 97(4), 723.
<https://doi.org/10.1037/0022-0663.97.4.723>
- Málková, D., & Jirotková, D. (2019). Metakognitivní projevy u žáků s nadáním v průběhu řešení nestandardních úloh. *Online Journal of Primary and Preschool Education*, 3(1), 16–26.
<https://doi.org/10.21062/ujep/328.2019/a/2533-7106/OJPPE/2019/3/16>
- Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and cognition*, 21(1), 422–430.
<https://doi.org/10.1016/j.concog.2011.09.021>
- Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: meta-d', response-specific meta-d', and the unequal variance SDT model. In S. M. Fleming, & C. D. Frith (Eds.), *The cognitive neuroscience of metacognition* (s. 25–66). Springer.
https://doi.org/10.1007/978-3-642-45190-4_3
- Mareš, J. (2013a). *Pedagogická psychologie*. Portál.
- Mareš, J. (2013b). Přehledové studie: jejich typologie, funkce a způsob vytváření. *Pedagogická orientace*, 23(4), 427–454. <https://doi.org/10.5817/PedOr2013-4-427>
- Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, 60(1), 95–109. <https://doi.org/10.1016/j.compedu.2012.07.011>
- Mazancieux, A., Fleming, S. M., Souchay, C., & Moulin, C. J. (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *Journal of Experimental Psychology: General*, 149(9), 1788. <https://doi.org/10.1037/xge0000746>
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1–10.
<https://doi.org/10.1037/xge0000746>
- McGuire, S. Y. (2015). *Teach students how to learn: Strategies you can incorporate into any course to improve student metacognition, study skills, and motivation*. Stylus Publishing, LLC.
- Meijer, J., Veenman, M. V., & van Hout-Wolters, B. H. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation*, 12(3), 209–237. <https://doi.org/10.1080/13803610500479991>
- Meijer, J., Veenman, M. V., & van Hout-Wolters, B. (2012). Multi-domain, multi-method measures of metacognitive activity: what is all the fuss about metacognition... indeed?. *Research papers in education*, 27(5), 597–627. <https://doi.org/10.1080/02671522.2010.550011>



- Metcalfe, J., & Dunlosky, J. (2009). *Metacognition: A textbook for cognitive, educational, life span & applied psychology*. SAGE Publications.
- Missett, T., & McCormick, K. (2014). Conceptions of giftedness. In J. A. Plucker, & C. M. Callahan (Eds.), *Critical issues and practices in gifted education* (s. 143–157). Prufrock Press.
- Moon, S. M. (2009). Myth 15: High-ability students don't face problems and challenges. *Gifted Child Quarterly*, 53(4), 274–276. <https://doi.org/10.1177/0016986209346943>
- Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology and Climatology*, 12(4), 595–600. [https://doi.org/10.1175/1520-0450\(1973\)012<0595:ANVPOT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2)
- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological bulletin*, 95(1), 109. <https://doi.org/10.1037/0033-2909.95.1.109>
- Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item comments on Schraw (1995). *Applied Cognitive Psychology*, 10(3), 257–260. [https://doi.org/10.1002/\(SICI\)1099-0720\(199606\)10:3<257::AID-ACP400>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1099-0720(199606)10:3<257::AID-ACP400>3.0.CO;2-9)
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science*, 2(4), 267–271. <https://doi.org/10.1111/j.1467-9280.1991.tb00147.x>
- Nelson, T. O., & Narens, L. (1980). A new technique for investigating the feeling of knowing. *Acta Psychologica*, 46(1), 69–80. [https://doi.org/10.1016/0001-6918\(80\)90060-8](https://doi.org/10.1016/0001-6918(80)90060-8)
- Nietfeld, J. L., Enders, C. K., & Schraw, G. (2006). A Monte Carlo comparison of measures of relative and absolute monitoring accuracy. *Educational and Psychological Measurement*, 66(2), 258–271. <https://doi.org/10.1177/0013164404273945>
- Ning, H. K. (2016). Examining heterogeneity in student metacognition: A factor mixture analysis. *Learning and Individual Differences*, 49, 373–377. <https://doi.org/10.1016/j.lindif.2016.06.004>
- Nováková, A., Chytrý, V., & Říčan, J. (2018). Vědecké myšlení a metakognitivní monitorování studentů učitelství pro 1. stupeň základní školy. *Scientia in educatione*, 9(1), 66–80. <https://doi.org/10.14712/18047106.1041>
- O'Brien, J. (1993). Action research through stimulated recall. *Research in Science Education*, 23(1), 214–221. <https://doi.org/10.1007/BF02357063>
- Peña-Ayala, A. (Ed.). (2015). *Metacognition: Fundamentals, Applications, and Trends*. Springer. <https://doi.org/10.1007/978-3-319-11062-2>
- Pfeiffer, S. I. (2013). *Serving the gifted: Evidence-based clinical and psychoeducational practice*. Routledge.
- Poitras, E. G. (2015). The MetaHistoReasoning Tool: Studying Domain-Specific Metacognitive Activities in an Intelligent Tutoring System for History. In A. Peña-Ayala (Ed.), *Metacognition: Fundamentals, Applications, and Trends* (s. 345–365). Springer. https://doi.org/10.1007/978-3-319-11062-2_14

- Popelka, S. (2018). *Eye-tracking (nejen) v kognitivní kartografii: praktický průvodce tvorbou a vyhodnocením experimentu*. Univerzita Palackého v Olomouci. <https://doi.org/10.5507/prf.18.24453132>
- Portešová, Š. (2011). *Rozumově nadané děti s dyslexií*. Portál.
- Portešová, Š., Straka, O., Veenman, M. V., Lacko, D., Šašinka, Č., & Jabůrek, M. (2021). Learning strategies in gifted students with low and high levels of metacognition: An eye-tracking study [publikace v recenzním řízení].
- Randall, A., Fairbanks, M. M., & Kennedy, M. L. (1986). Using think-aloud protocols diagnostically with college readers. *Literacy Research and Instruction*, 25(4), 240–253. <https://doi.org/10.1080/19388078609557883>
- Renzulli, J. S. (2005). The three-ring conception of giftedness: A developmental model for creative productivity. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conception of giftedness* (s. 53–92). Cambridge University Press. <https://doi.org/10.1017/CBO9780511610455.015>
- Richer, J., Mongeau, P., Lafortune, L., Deaudelin, C., Doudin, P. A., & Martin, D. (2004). Outil d'évaluation de la métacognition: processus de validation et utilisation à des fins pédagogiques. In R. Pallascio, D. Marie-France, & L. Lafortune (Eds.), *Pensée et réflexivité. Théories et pratiques* (s. 73–106). Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv18ph0t1.7>
- Shaughnessy, M. F., Veenman, M., & Kennedy, C. K. (Eds.). (2008). *Meta-cognition: A recent review of research, theory, and perspectives*. Nova Science Publishers
- Schellings, G. L., van Hout-Wolters, B. H., Veenman, M. V., & Meijer, J. (2013). Assessing metacognitive activities: the in-depth comparison of a task-specific questionnaire with think-aloud protocols. *European journal of psychology of education*, 28(3), 963–990. <https://doi.org/10.1007/s10212-012-0149-y>
- Schmitt, M. C., & Sha, S. (2009). The developmental nature of meta-cognition and the relationship between knowledge and control over time. *Journal of Research in Reading*, 32(2), 254–271. <https://doi.org/10.1111/j.1467-9817.2008.01388.x>
- Schneider, W., & Pressley, M. (2013). *Memory development between two and twenty*. Psychology Press. <https://doi.org/10.4324/9780203774496>
- Schraw, G. (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology*, 9(4), 321–332. <https://doi.org/10.1002/acp.2350090405>
- Schraw, G. (2001). Metacognition in basic skills instruction. In H. J. Hartman (Ed.), *Metacognition in learning and instruction: theory, research and practice* (s. 3–16). Kluwer Academic Publishers.
- Schraw, G. (2009). Measuring Metacognitive Judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (s. 415–429). Routledge.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary educational psychology*, 19(4), 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Schraw, G., Kuch, F., Gutierrez, A. P., & Richmond, A. S. (2014). Exploring a three-level model of calibration accuracy. *Journal of Educational Psychology*, 106(4), 1192. <https://doi.org/10.1037/a0036653>



- Schwartz, D. L., Chase, C., Chin, D. B., Oppezzo, M., Kwong, H., Okita, S., ... & Wagster, J. D. (2009). Interactive metacognition: Monitoring and regulating a teachable agent. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (s. 340–358). Routledge.
- Snyder, K. E., Nietfeld, J. L., & Linnenbrink-Garcia, L. (2011). Giftedness and metacognition: A short-term longitudinal investigation of metacognitive monitoring in the classroom. *Gifted Child Quarterly*, 55(3), 181–193. <https://doi.org/10.1177/0016986211412769>
- Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary educational psychology*, 27(1), 51–79. <https://doi.org/10.1006/ceps.2001.1091>
- Stevens, R., Johnson, D. F., & Soller, A. (2005). Probabilities and predictions: modeling the development of scientific problem-solving skills. *Cell biology education*, 4(1), 42–57. <https://doi.org/10.1187/cbe.04-03-0036>
- Sternberg, R. (2005). The Triarchic Theory of Successful Intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (s. 103–119). Guilford Press.
- Straka, O., & Portešová, Š. (2019). Metakognitivní charakteristika intelektově mimořádně nadaných dětí. *Československá Psychologie*, 63(5), 543–561.
- Terman, L. M. (1926). *Genetic studies of genius*. Stanford University Press.
- van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In R. Azevedo, & V. Alevan (Eds.), *International handbook of metacognition and learning technologies* (s. 143–156). Springer. https://doi.org/10.1007/978-1-4419-5546-3_10
- Vangsness, L., & Young, M. E. (2021). More isn't always better: when metacognitive prompts are misleading. *Metacognition and Learning*, 16(1), 135–156. <https://doi.org/10.1007/s11409-020-09241-9>
- Veenman, M. V. J. (2008). Giftedness: Predicting the speed of expertise acquisition by intellectual ability and metacognitive skillfulness of novices. In M. F. Shaughnessy, M. V. J. Veenman, & C. Kley-Kennedy (Eds.), *Meta-Cognition: A recent review of research, theory, and perspectives* (s. 207–220). Nova Science Publishers.
- Veenman, M. V. J. (2011). Learning to self-monitor and self-regulate. In R. Mayer, & P. Alexander (Eds.), *Handbook of research on learning and instruction*, (s. 233–257). Routledge.
- Veenman, M. V. J. (2012). Metacognition in Science Education: Definitions, Constituents, and Their Intricate Relation with Cognition. In A. Zohar, Y. J. Dori (Eds.), *Metacognition in Science Education* (s. 21–36). New York: Springer. https://doi.org/10.1007/978-94-007-2132-6_2
- Veenman, M. V. J. (2015). Metacognition: 'Know thyself'. Use that knowledge especially to regulate your own behavior. *De Psycholoog*, 50 (special. ed.), 8–18.
- Veenman, M. V. J. (2016). *Experience with using computer logfiles*. [osobní sdělení]. Leiden, Nizozemí.



- Veenman, M. V., Bavelaar, L., De Wolf, L., & van Haaren, M. G. (2014a). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, 29, 123–130. <https://doi.org/10.1016/j.lindif.2013.01.003>
- Veenman, M. V., Hesselink, R. D., Sleeuwaegen, S., Liem, S. I., & van Haaren, M. G. (2014b). Assessing developmental differences in metacognitive skills with computer logfiles: Gender by age interactions. *Psychological Topics*, 23(1), 99–113.
- Veenman, M. V., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science*, 33(3), 193–211. <https://doi.org/10.1007/s11251-004-2274-8>
- Veenman, M. V., & van Cleef, D. (2019). Measuring metacognitive skills for mathematics: students' self-reports versus on-line assessment methods. *ZDM*, 51(4), 691–701. <https://doi.org/10.1007/s11858-018-1006-5>
- Veenman, M. V., van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and learning*, 1(1), 3–14. <https://doi.org/10.1007/s11409-006-6893-0>
- Vuorre, M., & Metcalfe, J. (2021). Measures of relative metacognitive accuracy are confounded with task performance in tasks that permit guessing. *Metacognition and Learning*, 1–23. <https://doi.org/10.1007/s11409-020-09257-1>
- Waters, H. S. E., & Schneider, W. E. (Eds.) (2010). *Metacognition, strategy use, and instruction*. Guilford Press.
- Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195092509.001.0001>
- Wilhelm, P., Beishuizen, J. J., & van Rijn, H. (2005). Studying inquiry learning with FILE. *Computers in human behavior*, 21(6), 933–943. <https://doi.org/10.1016/j.chb.2004.03.013>
- Young, A. E., & Worrell, F. C. (2018). Comparing metacognition assessments of mathematics in academically talented students. *Gifted Child Quarterly*, 62(3), 259–275. <https://doi.org/10.1177/0016986218755915>
- Zohar, A., & Dori, Y. J. (Eds.). (2011). *Metacognition in science education: Trends in current research*. Springer. <https://doi.org/10.1007/978-94-007-2132-6>

JAK MĚŘIT METAKOGNICI (NEJEN) U NADANÝCH DĚTÍ

Ondřej Straka

Obálka a grafická úprava: Reprocentrum, a. s., Bezručova 29, 678 01 Blansko

Jazyková redakce: Veronika Muselíková

Vydala Masarykova univerzita, Žerotínovo nám. 617/9, 601 77 Brno

1., elektronické vydání, 2021

ISBN 978-80-210-9905

<https://doi.org/10.5817/CZ.MUNI.M210-9905-2021>

www.fss.muni.cz

www.press.muni.cz



MUNI
PRESS

MUNI
FSS