

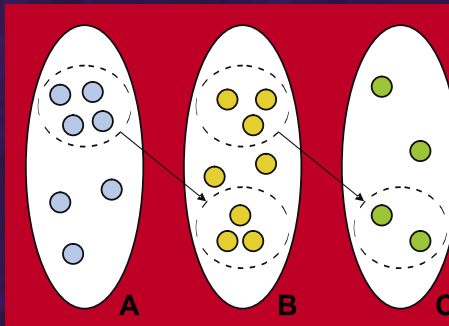
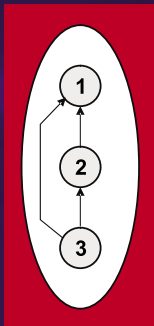
PSYCHOMETRIKA

Měření v psychologii

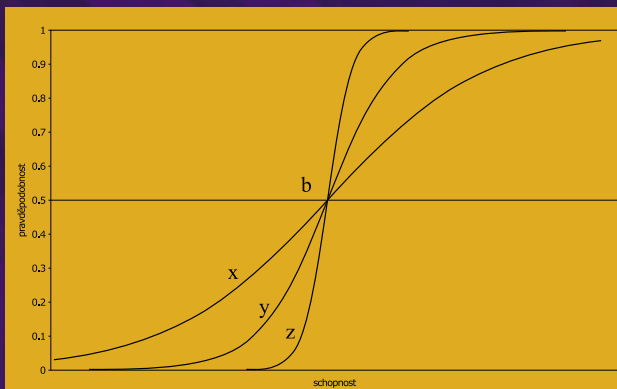
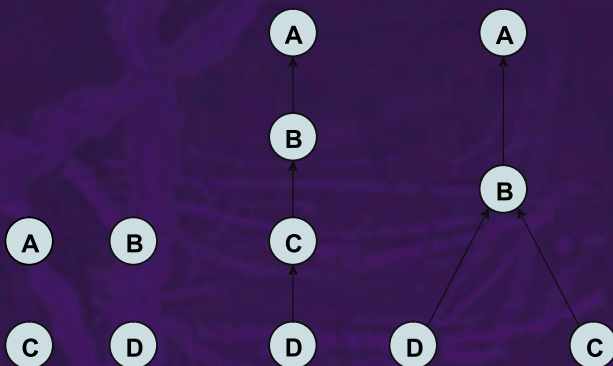
Tomáš URBÁNEK

Denisa DENGLEROVÁ

Jan ŠIRŮČEK



	A	B	C	D
A	1	0	0	0
B	0	1	0	0
C	0	0	1	0
D	0	0	0	1



PSYCHOMETRIKA

Měření v psychologii

Tomáš Urbánek
Denisa Denglerová
Jan Širůček

Urbánek, Tomáš

Psychometrika : měření v psychologii / Tomáš Urbánek, Denisa Denglerová, Jan Širůček. – Vyd. 1. – Praha : Portál, 2011. – 320 s.
ISBN 978-80-7367-836-4 (váz.)

159.938.3

- psychometrie
- kolektivní monografie

159.9 – Psychologie [17]

UPOZORNĚNÍ PRO ČTENÁŘE A UŽIVATELE TÉTO KNIHY

Všechna práva vyhrazena. Žádná část této elektronické knihy nesmí být reprodukována a šířena v papírové, elektronické či jiné podobě bez předchozího písemného souhlasu nakladatele.

Neoprávněné užití této knihy bude trestně stíháno.

Tomáš Urbánek se psaní knihy zúčastnil v rámci plnění výzkumného záměru Psychologického ústavu AV ČR, v.v.i., s číslem AV0Z70250504 (Člověk v kontextech celoživotního vývoje).

Denisa Denglerová se psaní knihy zúčastnila v rámci plnění grantu MUNI/A/1024/2009 (Výzkum dětí ze sociálně znevýhodněných skupin).

Jan Širůček se psaní knihy zúčastnil v rámci plnění projektu MSM0021622406 (Psychologické a sociální charakteristiky dětí, mládeže a rodiny, vývoj osobnosti v době proměn moderní společnosti).

Knihu lektorsky posoudili:

prof. PaedDr. Iva Stuchlíková, CSc.

doc. PhDr. Peter Halama, Ph.D.

© Tomáš Urbánek, Denisa Denglerová, Jan Širůček, 2011

© Portál, s.r.o., Praha 2011

ISBN

978-80-7367-836-4 (tištěná kniha)

978-80-262-0143-4 (pdf)

978-80-262-0430-5 (ePub)

978-80-262-0431-2 (mobi)

Summary	8
Předmluva.....	9
1 Psychometrika v souvislostech	13
Úvod.....	13
Co je psychometrika.....	13
Psychometrika v kontextu dalších disciplín	13
Měření v psychologii z hlediska historie.....	14
Institucionální kontext psychometrie	47
Shrnutí	48
2 Aplikace měření v psychologii	49
Úvod.....	49
Typy použití psychodiagnostických metod	49
Typy psychodiagnostických metod	52
Klinický a statistický přístup.....	56
Shrnutí	58
3 Měření a „měření“ v psychologii	59
Stevensovy „úrovně měření“	60
Škálování jako pravidla korespondence	66
Spojené měření.....	71
Povaha měřených atributů	78
Shrnutí	80
4 Charakteristiky psychodiagnostických metod	81
Úvod.....	81
Praktické problémy měření v psychologii	81
Klasifikace chyb měření.....	83
Charakteristiky psychodiagnostických metod.....	88
Shrnutí	93

5	Různá pojetí reliability a způsoby jejího zjišťování	94
	Náhodné chyby a pravé skóry	94
	Teoretická definice reliability.....	95
	Odhad reliability	98
	Teorie zobecnitelnosti	112
	Faktory ovlivňující odhad reliability	120
	Minimální žádoucí reliabilita metody.....	129
	Shrnutí	131
6	Zdroje důkazů o validitě.....	132
	Třídění zdrojů důkazů o validitě.....	132
	Obsahové zdroje důkazů o validitě.....	133
	Empirické důkazy o validitě	137
	Konstruktové důkazy o validitě	143
	Kritika současného pojetí validity.....	149
	Shrnutí	151
7	Teorie odpovědi na položku.....	152
	Historie IRT	152
	Základní předpoklady a vlastnosti IRT	154
	Modely IRT	156
	Odhad parametrů položek.....	163
	Informační funkce položky.....	165
	Možnosti aplikace IRT	168
	Konkrétní aplikace IRT modelů.....	173
	Shrnutí	179
8	Teorie vědomostního prostoru.....	180
	Základní pojmy a definice teorie vědomostního prostoru.....	181
	Příklady ilustrující definované pojmy.....	183
	Prerekvizitní relace	185
	Prostor kompetencí a prostor výkonu	197
	Vybrané aplikace teorie vědomostního prostoru v psychologii	198
	Využití KST a shrnutí	202

9	Konstrukce testových metod a analýza položek.....	204
	Konstrukce testu nebo dotazníku	204
	Psaní položek.....	208
	Hlavní formáty odpovídání	210
	Skór položky a skór testu.....	217
	Analýza položek	217
	Poznámky k zásadám adaptace zahraničních metod.....	230
	Shrnutí	235
10	Normalizace a normy	236
	Základní pojmy	236
	Normální rozložení	237
	Možné výjimky.....	238
	Soubory pro tvorbu norem	241
	Pravděpodobnostní výběr	241
	Nepravděpodobnostní výběr.....	243
	Poznámky k metodám výběru z různých populací.....	244
	Typy odvozených skórů	245
	Alternativní metody interpretace testových skórů	249
	Existují situace, kdy normy nejsou důležité?	254
	Shrnutí	254
11	Porovnání současných teorií měření v psychologii	255
	Pojmenování nejdůležitějších témat jednotlivých teorií.....	255
	Srovnání klíčových témat jednotlivých teorií.....	263
	Shrnutí	265
	Slovníček pojmů	266
	Literatura.....	302
	Rejstřík jmenný	311
	Rejstřík věcný	315

Summary

This book is intended mainly for psychologists, students who have to study psychometrics during their studies, and experienced practitioners who want to refresh some of the concepts and approaches. Generally, it is intended for anyone who is interested in measurement in psychology or in related disciplines. In our experience, we have found that one cannot ignore the fact that psychometrics is truly not a domain that usually makes students study psychology, and unfortunately even those who practice do not have enough knowledge of it. This often leads to the incorrect application of tests and questionnaires, their unjustified refusal or, on the contrary, undue confidence in their results, and erroneous interpretation and communication of the results to clients. It can also lead to the mistaken notion that projective techniques do not need to fulfill the same standards as tests and questionnaires, etc. In addition, it leads to an even more serious misunderstanding of what so-called measurement in psychology is, and what its specifics (and deficiencies) compared to measurement in other scientific disciplines are.

For practitioners in psychology, psychometrics is a very important field, because both European and global trends concerning the use of psychodiagnostic methods lead to greater standardization and control by their users. The goal is, on one hand, to protect clients in psychological services against unqualified use or misuse of these methods, and, on the other hand, to protect the reputation of psychology itself. Of course, this issue is not related only to psychologists, as some studies show that a large number of psychodiagnostic methods are administered and scored by non-psychologists in Europe. This practice in itself does not necessarily lead to damage to clients or the reputation of the discipline; however, this holds only provided that all people that test clients perform each phase of testing in accordance with high professional and ethical standards. In the case of professional standards, this requires the demonstration of theoretical knowledge and practical skills. This book should convey at least a part of this knowledge to its readers.

In addition, the book can be used by experts in other disciplines such as education, sociology, medicine, marketing, and management; that is by everyone who, in their jobs, uses tools that are adapted, or created on the base of their own knowledge and experience – questionnaires, surveys, satisfaction inventories, standardized interviews, etc. All these types of structured research and diagnostic tools can be analyzed, critically evaluated or improved by means of psychometric methods.

Předmluva

Před zhruba 15 lety se první autor této knihy začal na Psychologickém ústavu AV ČR profesionálně zabývat psychometrikou, kterou posléze i začal vyučovat na Filosofické fakultě Masarykovy univerzity v Brně. Zpočátku se spolehlal na starší učební texty, ale záhy vznikla potřeba vytvořit vlastní učební text, a ten pak vyšel v roce 2002 v Nakladatelství MU. Byl sice přijat vcelku příznivě, ale vyšel v nízkém nákladu, a proto jsme se před asi dvěma lety rozhodli s bývalými studenty, současnými přáteli a kolegy Miroslavem Filipem a Denisou Denglerovou, že ho přepracujeme a také rozšíříme o přístupy, které se začaly používat teprve v nedávné době, nebo které se sice leckde již používají, ale v České republice dosud všeobecně známé a používané nejsou. Současně jsme cítili potřebu některé základní myšlenky psychometrie uvést do současného kontextu. To, co se nám ještě před pár lety jevilo jako monolitní bezrozporná disciplína, se v kontextu některých nových knih, článků i vlastních úvah jeví spíše jako neuralgický bod celé psychologické vědy nebo přinejmenším její kvantitativně orientované části.

V původní trojici nakonec Miroslava Filipa, který se psaní knihy nemohl věnovat, nahradil Jan Širůček, další metodolog. Většinu kapitol tvoří podstatným způsobem přepracovaný text původních skript doplněný o některé starší, dříve nepoužité prameny a také o nejnovější zdroje. Kapitoly 5 psali ve spolupráci T. Urbánek a J. Širůček, kapitolu 11 T. Urbánek a D. Denglerová, kapitoly 7 a 8 psala D. Denglerová, závěrečný Slovníček pojmů je kompilací různých zdrojů s hesly, která dodali jednotliví spoluautoři, a podstatným způsobem ho redigoval J. Širůček. Jednotlivé kapitoly jsme si však navzájem četli a připomínkovali, takže se nejedná o editovaný sborník, ale skutečné spoluautorství.

Tato kniha je určena především psychologům, od studentů, kteří mají výuku psychometrie povinnou v rámci studia, až po dlouholeté praktiky, kteří mají potřebu si některé pojmy a postupy připomenout. Především je ale určena každému, kdo se zajímá o měření v psychologii a příbuzných disciplínách. Podle našich zkušeností nelze zavírat oči před tím, že psychometrie nepředstavuje pro většinu studentů disciplínu, kvůli níž se začali psychologii věnovat, a ani mezi praktiky bohužel nebývají její znalosti dostatečné. To pak

vede k chybnému používání testových a dotazníkových metod, k jejich neodůvodněnému odmítání, nebo naopak k nadměrné důvěře v jejich výsledky, k chybným sdělením a interpretacím výsledků klientům, k mylné představě, že na projektivní metody není nutné klást stejné požadavky jako na testy a dotazníky apod. Kromě toho to však vede k mnohem závažnějšímu nepochopení, co tzv. měření v psychologii znamená a jaké jsou jeho zvláštnosti (a nedostatky) ve srovnání s jinými vědními disciplínami.

Pro praktické psychology je psychometrika značně důležitým oborem, protože evropské, ale i celosvětové trendy v používání psychodiagnostických metod směřují k větší standardizaci a kontrole uživatelů. Cílem je na jedné straně chránit klienty psychologických služeb před nekvalifikovaným používáním nebo zneužíváním těchto metod a na druhé straně chránit pověst oboru samotného. To se pochopitelně netýká jen psychologů. Existují výzkumy, které dokazují, že v Evropě velké množství psychodiagnostických metod administrují a vyhodnocují lidé, kteří nestudovali psychologii (viz např. Bartram, 2001). Tato praxe sama o sobě v případě některých testů nemusí znamenat poškození ani klientů, ani pověsti oboru. To vše ale pouze za předpokladu, že každou fázi testování budou provádět pouze lidé, kteří budou dodržovat veškeré odborné a etické standardy (viz např. AERA, APA, NCME, 2001). V případě odborných standardů to znamená, že tito lidé prokáží teoretické znalosti a praktické dovednosti. Tato kniha má čtenářům zprostředkovat alespoň část z nich.

Kromě toho mohou tuto knihu využít i odborníci z dalších oborů – např. pedagogiky, sociologie, medicíny, marketingu, managementu a dalších. Tedy všichni, kdo při své práci používají již hotové, adaptované nebo na základě vlastních znalostí a zkušeností vytvořené testy, dotazníky, ankety, průzkumy spokojenosti, standardizované rozhovory apod. Všechny typy těchto strukturovaných výzkumných a diagnostických nástrojů lze pomocí psychometrických metod analyzovat, kriticky hodnotit a případně vylepšovat.

Rádi bychom poskytli stručný návod, jak by měl tuto knihu číst čtenář bez znalostí nebo jen s malými znalostmi z psychometrie (obeznámený čtenář může pochopitelně začít kdekoli). Kapitoly 1–3 lze číst relativně samostatně a představují důležitý úvod a širší kontext oboru. Kapitoly 4–6 obsahují relativně ucelený výklad tzv. klasické teorie testů, který je z hlediska pochopení základních pojmů a postupů dobré číst popořadě. Doplnují ho kapitoly 9

a 10, které se věnují postupům konstrukce a revize psychometrických metod a postupům analýzy položek. Kapitoly 7 a 8 jsou vůči zbytku knihy víceméně paralelní – to znamená, že se odkazují k dalším částem knihy, ale lze je pochopit samostatně. Věnují se novějším, ve světě již používaným přístupům, které se v České republice začnou prosazovat v blízké budoucnosti. Kapitola 11 je pokusem o závěrečné porovnání předností a nedostatků jednotlivých přístupů. Na konci knihy je uveden Slovníček pojmů. Ten je možné konzultovat průběžně a hledat v něm definice klíčových pojmů psychometrie nebo specifické (možná idiosynkratické) definice běžných pojmů používaných v kontextu psychometrie.

Nezbytnou součástí jakéhokoli textu o psychometrii jsou matematické vzorce. V současné době, kdy jakákoli analýza dat je prováděna za pomoci počítačů, by se mohlo zdát, že tyto vzorce není nutné ovládat. Protože však pro matematicky gramotného čtenáře obsahují velmi zhuštěnou informaci o postupu výpočtu konkrétního indexu, skutečný zájemce o psychometrii by se s nimi měl seznámit, protože se jedná o jediný způsob, jak skutečně porozumět tomu, co se s daty děje.

Na závěr bychom se chtěli stručně vyjádřit k určitému podráždění, kterého jsme si všimli u některých zasloužilejších kolegů v souvislosti s námi preferovaným tvarem „psychometrie“ oproti jimi preferovanému názvu „psychometrie“. Uznáváme, že existují analogické názvy jako „planimetrie“, „biometrie“, „sociometrie“ a další, ale vycházíme z anglosaské tradice, kde je předmět našeho zájmu tradičně označován jako „psychometrics“, a existuje dokonce vlivný časopis *Psychometrika*. Současně víme o tom, že tvar „psychometry“ odpovídající onomu českému (ale i německému a francouzskému) „psychometrie“ je ve dvou vlivných encyklopediích (anglická verze *Wikipedie* a *Encyclopaedia Britannica*) vymezen jako součást parapsychologie. Z těchto důvodů nám připadá lepší používat zatím nezatížený název „psychometrie“. Je to však spíše výraz našeho osobního vkusu, a nikoli autoritativní názor, na kterém bychom chtěli rigidně ulpívat. Vždyť jak říká Shakespearova Julie: „Co je po jméně? Co růží zvou i zváno jinak, vonělo by stejně.“

Rádi bychom, aby tato kniha pomáhala také v komunikaci odborníků zabývajících se psychometrií ať už v oblasti výzkumu, nebo praktické

psychodiagnostiky. Z toho důvodu jsme založili webové diskusní fórum na adrese <www.psychometrika.cz>, kam všechny zájemce o psychometriku a příbuzná témata zveme.

Rovněž velice děkujeme oběma lektorům, kteří pomohli ke zlepšení této knihy.

V Brně dne 22. 2. 2011

Tomáš Urbánek
Denisa Denglerová
Jan Širůček

1 Psychometrika v souvislostech

Úvod

Žádná vědecká disciplína neexistuje ve vzduchoprázdnu, ale v mnoha souvislostech a kontextech. K pochopení toho, proč je psychometrika pro psychologii důležitá, se pokusíme zohlednit alespoň tři z nich – (1) kontext dalších psychologických disciplín i ostatních vědních oborů, (2) historický kontext a (3) institucionální kontext. Nejprve se však pokusíme psychometriku definovat.

Co je psychometrika

Psychometrika je nejčastěji chápána jako jakási pomocná psychologická disciplína zaměřená na tvorbu, adaptaci, úpravy a hodnocení psychodiagnostických metod. Je však něčím mnohem důležitějším. Jedná se o obor zabývající se teoretickými otázkami měření v psychologii, a teprve druhotně aplikací těchto teoretických principů v praxi. K tomu účelu používá množství matematických a statistických postupů, které byly v průběhu vývoje této disciplíny vytvářeny pro konkrétní potřeby psychologické praxe.

Psychometrika má tedy k praktické psychodiagnostice podobný vztah, jako mají metodologie a statistická analýza dat k výzkumu v psychologii – představuje její nedílnou součást. Přestože se při konstrukci testu používají i jiné postupy než psychometrické, konstrukce a hodnocení testových a dotazníkových metod se bez ní neobejdou, podobně jako se psychologický výzkum v současnosti neobejde bez kvalitativního nebo kvantitativního, převážně statistického zpracování dat. Znalosti psychometrických postupů jsou tedy důležité pro každého psychologa.

Psychometrika v kontextu dalších disciplín

Psychometrika je nutně propojena se všemi teoretickými i aplikovanými psychologickými disciplínami. Neoddělitelná je především od **psychodiagnostiky**, pro kterou její aplikace mohou poskytovat servis v podobě tvorby, adaptace, ověřování

a modifikace testových a dotazníkových metod. Formuluje však také obecné principy, které platí i u metod, ke kterým lze z psychometrického hlediska přistupovat o něco obtížněji (např. u projektivních metod). Z **obecné a kognitivní psychologie** přebírá poznatky o zákonitostech fungování psychických jevů a z **vývojové psychologie** poznatky o psychickém vývoji. Protože aplikovaná psychometrika je empirická metoda, čerpá z **metodologie** principy výzkumu, metody a techniky, které je nutné použít pro získání zobecnitelných výsledků. **Matematika a matematická statistika** poskytují principy a postupy zpracování empirických dat: různé metody analýzy, statistického testování hypotéz, různé indexy a metody jejich odhadu atd.

Prostřednictvím psychodiagnostiky je psychometrika klíčová pro oblast **diferenciální psychologie**. Dá se současně říci, že jeden ze zdrojů zájmu o psychodiagnostiku a také o psychometriku leží právě ve vědeckém zájmu o studium individuálních rozdílů mezi lidmi, jehož průkopníkem byli např. James McKeen Cattell nebo William Stern, protože ke studiu těchto rozdílů poskytuje nástroje. Autoři nových diagnostických nástrojů ale k jejich tvorbě přistupují na základě poznatků diferenciální psychologie, takže ovlivňování obou oborů je vzájemné. Zcela nezbytná je psychodiagnostika, a tedy i psychometrika, pro **klinickou a poradenskou psychologii** a **další aplikované disciplíny** (psychologii práce, personalistiku atd.), protože bez psychodiagnostiky není v těchto oborech možné kvalifikovaně rozhodovat. Současně však tyto obory obohacují psychodiagnostiku o další poznatky a vytvářejí tlak na tvorbu a modifikaci metod. Zcela nezbytná je pak psychometrika pro jakékoli **výzkumné aktivity** využívající kvantitativní přístupy, protože ty se bez nějaké formy měření neobejdou.

To neznamená, že agendou psychometrie by mělo být poskytování služby ostatním oborům psychologie, pro které by vytvářela nástroje měření. Spíše je možné říci, že psychometrika by v současné době měla být souborem hledisek, jejichž prostřednictvím v psychologii reflektujeme kvantitativní postupy, a to v kterékoli z psychologických disciplín. Toto tvrzení bude jasnější z následujících kapitol.

Měření v psychologii z hlediska historie

Pro pochopení současného stavu nějaké disciplíny je velmi důležité pokusit se pochopit, jakým způsobem se vyvíjela. Pro dějiny psychometrie to platí dvojnásob, protože dlouho nebyly adekvátně zpracovány. V mnoha knihách jako např.

v Helmstadterově (1964), Aikenově (1999) nebo Huntově (2000) se objevují převážně jenom úspěchy psychometrie, nebo spíše to, co je za úspěchy vydáváno. Helmstadter a Hunt zdůrazňují hlavně tři jména – Francise Galtona, Jamese McKeena Cattella a Alfreda Bineta (viz níže), popřípadě ještě zdůrazňují přínos autorů zabývajících se faktorovou analýzou – např. Charlese Spearmana. Gouldova (1998) kniha se spíše než historii psychometrie věnuje historii testování, tzn. historii aplikací psychometricky vytvořených psychodiagnostických metod. Jiné knihy, jako např. Klinova (1993) nebo McDonaldova (1999), se historii nevěnují vůbec. Přitom z historie psychologie je známo, že byla mnoha autory koncipována zpočátku jako kvantitativní věda, takže měření muselo být nedílnou součástí praxe psychologického výzkumu. Pouze Michellova (1999) kniha představuje skutečně zásadní počin obsahující informace, které v ostatních knihách nejsou obsaženy, spolu s psychologicky fundovaným kritickým postojem k vývoji oboru. Následující stránky obsahují pokus o syntézu různých informací a komentářů, přičemž je nutné přiznat především silný vliv právě Michella (1999).

Předvědecké období

Známý bonmot, že psychologie má dlouhou minulost, ale krátkou historii (McDonald, 1999), platí ve stejné míře také o psychodiagnostice a psychometrice. Povědomí o individuálních rozdílech v psychických vlastnostech lze vysledovat už v nejstarších literárních památkách, např. v *Eposu o Gilgamešovi* nebo *Bibli* (Kniha Soudců, 7,3–7). Počátky praxe, kterou bychom dnes nazvali psychodiagnostikou, lze popsat už kolem roku 2200 př. n. l. v Číně, kde se používaly ústní zkoušky pro výběr vládních úředníků a probíhalo také jejich pravidelné přezkušování (Aiken, 1999).

Obrátíme-li se k antické filosofii, pak zjistíme, že už filosofové Platon a Aristoteles se zabývali otázkami individuálních rozdílů. Řecký lékař Hippokrates v 5. století př. n. l. tvrdil, že choroby mají přirozené příčiny, jejichž vznik podle jeho názoru souvisel s rovnováhou nebo nerovnováhou čtyř tělních tekutin – krve, hlenu, žluči a černé žluči. Na jeho představy navázal o pár století později další Řek Galenos, který v souvislosti s převládající tělní tekutinou popsal čtyři typy temperamentu.

Aristotelův pokračovatel v Lykeionu Teofrastos popsal v díle *Charaktery* 30 typů osobnosti (např. lichotník, mluvka, lhář, lakomec, mrzout), ale bez

empirických dokladů o jejich existenci v dnešním slova smyslu. Také o mnoho století později se v dílech spisovatelů objevují psychologicky propracované popisy osobnostních typů – např. Cervantesův Don Quijote, postavy v dílech Williama Shakespeara či Geoffreye Chaucera a později i literární typy v knihách takových klasiků světové literatury, jakými byli Honoré de Balzac, Charles Dickens, Stendhal, Victor Hugo, Thomas Mann, Fjodor Michailovič Dostojevskij.

Ani jeden z obou uvedených přístupů, lékařský a literární, však nelze považovat za psychologický a vědecký v tom slova smyslu, jak mu rozumíme dnes. Ani další obory, které se odedávna snažily o poznání lidské psychiky (třebaže měly kromě toho i další cíle), jako astrologii, chiromantii, frenologii, fyziognomii, numerologii apod., nelze pokládat za psychologické a vědecké v dnešním smyslu. Přestože ne všechny poznatky nebo tvrzení těchto oborů můžeme bez výjimky odmítnout jako nepravdivé (ve smyslu nevalidní), je to téma na širší a hlubší diskusi, a v tomto textu se jimi zabývat nebudeme.

Zájem o studium psychiky nepochybně vyvolávali lidé duševně nemocní. Ve středověku byli často považováni za posedlé ďáblem nebo zlými duchy a jejich „lčba“ spadala do kompetence tzv. exorcistů, kteří měli pomoci náboženských obřadů tyto bytosti z nemocného vyhánět. V literatuře se často uvádí Velká francouzská revoluce jako klíčová historická událost, po níž začalo docházet ke změnám v postojích k duševně nemocným. Známostou postavou z této doby je Philippe Pinel, o kterém se poněkud pateticky píše, že „zbavil šílenice okovů“ (viz Černoušek, 1994). Faktem je, že začátkem 19. století se začal klást větší důraz na organické základy duševních chorob. Ještě později se začalo uvažovat o existenci funkčních příčin, proto se také začaly prosazovat psychologické formy terapie namísto pouze medicínských. Průkopníky tohoto chápání psychických poruch byli ve Francii např. Jean-Martin Charcot a Pierre Janet, kteří také začali s vývojem prvních metod pro jejich diagnostiku (Hunt, 2000).

Vědecké období

I při zpracování historických aspektů vývoje oboru je možné zohledňovat různé aspekty tohoto vývoje. Většina pasáží o historii v knihách o psychometrice nebo měření v psychologii je spíše než historii měření věnována historii testování. Protože i toto hledisko může být pro nás zajímavé, začneme nejprve jím. Později se pokusíme o kritičtější hodnocení vývoje.

Helmstadter (1964) uvádí, že kolem roku 1830 se začaly projevovat tři relativně nezávislé trendy, které později zásadním způsobem ovlivnily vznikající psychometriku. **Ve Francii** byl tradičně velký zájem o studium osob vybočujících z šablon normálního chování. Pracovali zde takoví lidé jako lékaři Esquirol, Itard, Seguin, psycholog Ribot a již zmínění Charcot (neurolog) a Pinel (dnes bychom jej označili jako psychiatra). Oproti tomu **v Německu** se prováděl spíše experimentální výzkum chování normálních dospělých lidí (jeho autory byli např. původně lékaři Weber a Wundt nebo původně fyzik Fechner). Třetí trend není určen místně, ale oborem **matematiky**. Už v roce 1733 byla Abrahamem de Moivreem matematicky vyjádřena funkce normálního rozložení pravděpodobnosti a matematici Gauss a Laplace rozpoznali její užitečnost jako matematického nástroje. Při měření výšky postavy byla využita belgickým astronomem, sociologem a matematikem Queteletem kolem roku 1835. Pro zajímavost, tento autor navrhl dodnes velmi používaný Body Mass Index (BMI), který se proto také někdy nazývá Queteletův index.

Kdyby byla vyhlášena anketa o titul „praotce mentálního testování“, podle Helmstadtera (1964, viz také Hunt, 2000) by se nejčastěji objevovala následující tři jména: Francis Galton, James McKeen Cattell a Alfred Binet. Každý z nich k němu přispěl svým podstatným dílem.

Francis Galton (1822–1911) byl britský polyhistor, nadšený soukromník-amatér bez vazby na tehdejší akademické instituce, což mu umožňovala jeho finanční situace. Byl silně ovlivněn pracemi svého bratrance Charlese Darwina (*On the Origin of Species*, 1859, česky např. 2007) a zajímal se o zkoumání dědičnosti mentálních rysů. V roce 1869 vyšla jeho kniha *Hereditary Genius*, ve které se zabýval otázkami dědičnosti inteligence a vyvinul k tomuto účelu metodu dnes nazývanou historiometrie. V této práci pokračoval i v knize *English Men of Science* z roku 1874, v jejímž podtitulu poprvé použil slovní spojení *nature or nurture*, dodnes používané jako nálepka pro spory mezi vlivy dědičnosti a prostředí (Hunt, 2000). V dědičnost inteligence zřejmě silně věřil a domníval se, že je vhodné motivovat inteligentní lidi, aby měli potomky, takže je považován za jednoho ze zakladatelů vysoce problematického oboru eugeniky.

Jak však poznat inteligentní lidi? V tom se obrátil k dalšímu vzoru tehdejší doby, kvantitativní vědě. Byl prý silně ovlivněn požadavkem, který

zformuloval italský učenec Galileo Galilei (1564–1642): „Měř to, co je měřitelné, a učíš měřitelným to, co měřitelné není.“ O Galtonovi se tvrdí, že počítal tahy štetcem ve dvou obdobích svého života, kdy byl portrétován, počítal trny na květinách a stromech, projevy nervozity osob sedících na nudné přednášce, počet vynikajících příbuzných význačných osob, znaky tváře, charakteristiky otisků prstů, žížaly na chodníku po dešti, počet případů chrápání atd. (Ludlow, 1998). V roce 1883 vyšla jeho kniha *Inquiries into Human Faculty and Its Development*, která představuje v podstatě počátky vědeckého zkoumání individuálních rozdílů a mentálních testů.

Kolem roku 1884 založil Antropometrickou laboratoř pro měření fyzických a senzomotorických charakteristik. Například na Mezinárodní výstavě zdraví měřil návštěvníkům „rychlost reakce, ostrost zraku a sluchu, rozlišování barev, schopnost odhadu délky, sílu tahu a stisku, sílu úderu, výšku, váhu, rozpětí paží, sílu dechu a kapacitu plic“ (údajně se nechalo otestovat 9337 osob; viz Hunt, 2000). Kromě toho se mu připisuje autorství metody korelační analýzy, kterou později matematicky zformuloval jeho žák Karl Pearson (1857–1936). Byl dokonce možná prvním autorem dotazníku v historii (Hunt, 2000), když zkoumal členy Královské společnosti (Royal Society).

Tato doba znamená počátek používání statistických metod jako hlavního matematického nástroje pro vědecké zkoumání individuálních rozdílů. Všichni další psychometrickí a statističtí pracovníci v oboru psychologie, jako byli v Anglii Karl Pearson, Charles Spearman, Cyril Burt a Geoffrey Thompson nebo ve Spojených státech amerických Truman Lee Kelley, Edward Lee Thorndike, Joy Paul Guilford spolu se svými žáky a mnoho dalších až do současnosti, vycházejí z této tradice, která započala v druhé polovině 19. století.

James McKeen Cattell¹ (1860–1944) absolvoval doktorské studium ve Wundtově laboratoři v Lipsku a dá se říci, že v jeho osobě došlo k propojení německy důkladné experimentální metodologie na jedné straně s americkým pragmatismem na straně druhé. McKeen Cattell si byl velmi dobře vědom důležitosti rigorózní kontroly intervenujících proměnných při pozorování chování, nicméně nesouhlasil s Wundtovou představou, že hlavním (respektive jediným) cílem psychologie je formulovat zobecňující závěry o normální dospělé

1 O J. McK. Cattellovi jsou kritické zmínky i níže v této kapitole, části Vývoj přístupů k měření v psychologii.

psychice. Podle Cattella odchylky od těchto zobecnění (tehdy považované pouze za obtěžující chyby operující ve výzkumech) nezmizí ani při použití té nejrigoróznější experimentální kontroly. Spíše tyto rozdíly považoval za projevy vlivu reálných proměnných, které jsou podle jeho názoru natolik důležité, že stojí za vědecké zkoumání. Takže přes Wundtovy námitky byla Cattellova doktorská disertace zaměřena na psychometrické studium individuálních rozdílů v reakčním čase (obhájil ji v Lipsku v roce 1886; Michell, 1999).

Cattell se kolem roku 1888 v Anglii seznámil s Galtonem a jeho metodami a po návratu do Spojených států, kdy se v pouhých 28 letech stal profesorem psychologie na University of Pennsylvania (dokonce o rok dříve než William James), se věnoval výzkumu individuálních rozdílů v inteligenci (Hunt, 2000), ke kterému přešel od výzkumu původně psychofyzikálně orientovaného. Ovšem tehdejší testy vycházely z představ, podle kterých veškerá psychika je dána zkušenostmi. Převládající teoretické názory pak ve shodě s těmito předpoklady tvrdily, že komplexní projevy psychiky nejsou nic než kombinace jednodušších sensorických zkušeností. Proto se rané pokusy o měření psychických jevů zaměřovaly na sensorické schopnosti a jiné jednodušší psychické funkce (např. sílu stisku, rychlost pohybu ruky, reakční čas na zvuk, rozlišování hmotností; Hunt, 2000). Výsledky tohoto snažení byly ale skličující – nebyl zjištěn žádný významný vztah mezi výsledky těchto testů u studentů prvního ročníku Columbia University a jejich známkami.

Tuto nesnáz pomohl vyřešit až třetí z výše uvedených „praotců“ **Alfred Binet** (1857–1911), francouzský experimentální psycholog se zájmem o jedince odlišné od normy, a zvláště o výzkum retardovaných dětí. V roce 1896, po mnoha letech pečlivého studia rozdílů v chování retardovaných, normálních a nadaných dětí, prezentovali Binet a jeho kolega Victor Henri sérii testů „komplexních“ psychických procesů. Postupně bylo stále jasnější, že přímý přístup k měření obecné schopnosti bude užitečnější než měření jednodušších funkcí. V roce 1904 se Binet stal členem komise zaměřené na studium výuky retardovaných dětí a spolu s Theodorem Simonem v roce 1905 vytvořili první formální škálu pro přímé měření komplexnějších schopností (Binetovu-Simonovu škálu), kterou v letech 1908 a 1911 ještě revidovali. Nejprve výsledky v těchto testech formulovali jako tzv. mentální věk a v roce 1912 německý psycholog William Stern vytvořil index nazývaný dnes „klasický inteligenční kvocient“ (viz kap. 10) (Hunt, 2000).

Binetova-Simonova škála se záhy stala velmi populární ve Spojených státech amerických, kde byla nejprve v roce 1910 adaptována Henrym Goddardem a později v roce 1916 revidována Lewisem M. Termanem – tato revize je známa jako Stanfordská Binetova škála (Hunt, 2000). Její další verze a další nově vytvořené testy byly používány nejprve ve školách a později, po vstupu Spojených států do první světové války, pro testování téměř dvou milionů branců – jednalo se o testy Army Alpha (pro gramotné) a Army Beta (obrázkový test pro negramotné nebo pro ty, kteří neuspěli v testu Army Alpha; u něj byly instrukce čteny nahlas asistentem), vytvořené Robertem Yerkesem a jeho spolupracovníky. Podobně jako v případě mnoha jiných technologií i v případě testových metod válka urychlila vývoj v důsledku toho, že armáda rozpoznala efektivitu testů pro rychlé rozřazení branců z hlediska jejich schopností, a tudíž pro jejich potenciální úspěšnost coby poddůstojníků a důstojníků.

Již před touto válkou byly testy inteligence zneužívány při testování přistěhovalců do Spojených států – tato praxe byla popsána (a kritizována) např. v knize Stephena J. Goulda *Jak neměřit člověka* (*Mismeasure of Man*, česky 1998). Je ale otázkou, zda toto dřívější ideologicky zatížené a zjevně diskriminující použití inteligenčních testů bylo důsledkem vlastností testů jako takových či tehdejšího pojetí inteligence, jak tvrdí Gould ve své knize, nebo důsledkem rasistických názorů některých tehdejších psychologů.

Pozdější vývoj testování už je natolik spletitý, že je možné z něj vybrat jen několik nejdůležitějších faktů. Prvním osobnostním dotazníkem byl v roce 1919 Woodworthův Personal Data Sheet, který nebyl v podstatě ničím jiným než standardizovaným psychiatrickým interview. O dva roky později byla publikována nejznámější psychologická projektivní metoda, Rorschachův test, v roce 1935 Tematický apercepční test a v roce 1943 Minnesotský multifázový osobnostní inventář (MMPI). Už od dvacátých let 20. století byly podnikány pokusy měřit zájmy, postoje a hodnoty. Například Kurt Lewin se zabýval experimentálním výzkumem konfliktů, frustrace a aspirační úrovně. Ve dvacátých a třicátých letech vznikalo mnoho jednofaktorových (i když tehdy se uvažovalo spíše o počtu škál než dimenzí) a vícefaktorových nástrojů pro diagnostiku rysů osobnosti (normální i patologické). Ve čtyřicátých a padesátých letech docházelo k rozvoji aplikovaných oblastí spolu s vývojem metod konstrukce, skórování a interpretace psychometrických nástrojů. V padesátých a šedesátých letech se prováděl výzkum specifických osobnostních proměnných, které také vyžadovaly nové diagnostické metody. Byly

tak vytvořeny např. metody pro měření rysů autoritářské osobnosti, závislosti/nezávislosti na poli, výkonové motivace, represe, senzitivace nebo místa kontroly.

Významným faktorem vývoje nejen psychometrie, ale zprostředkované i psychodiagnostiky se stala výpočetní technika, která se stává postupně stále přístupnější. Mnoho psychometrických modelů tak mohlo být ověřováno na reálných datech a propracovanější psychometrické postupy mohly být aplikovány při tvorbě nových metod.

V tabulce 1.1 je uveden stručný přehled nejdůležitějších událostí psychodiagnostiky, jak je uvádí Aiken (1999).

Tab. 1.1 Nejdůležitější události psychodiagnostiky (dle: Aiken, 1999)

4. stol. př. n. l.	Teofrastos popisuje 30 osobnostních nebo charakterových typů
2. st. n. l.	Galenos dává do souvislosti Hippokratovu teorii tělesných tekutin a temperament
1800	Němečtí lékaři Franz Joseph Gall a Johann Spurzheim zakládají pseudovědeckou frenologii, která dává do vztahu hrboly na lebce a osobnostní vlastnosti
1884	Francis Galton popisuje slovní asociační metodu, výběr chování a jiné metody pro měření charakteru
1892	Německý psychiatr Emil Kraepelin používá slovní asociační metodu v klinickém kontextu
1896	Kraepelin navrhuje novou klasifikaci mentálních poruch
1905	Carl Gustav Jung používá slovní asociační test pro analýzu mentálních komplexů; publikován první praktický test inteligence, Binetova-Simonova škála od Alfreda Bineta a Theodora Simona
1906	Corneille Jean Francois Heymans a Enno Dirk Wiersma vyvíjejí seznam symptomů svědčících o psychopatologii
1910	Publikovány Kentovy-Rosanoffovy seznamy slov
1919	Používán první standardizovaný osobnostní inventář Personal Data Sheet Roberta S. Woodswortha; publikován X-O Test Sidneyho Presseyho
1921	Poprvé publikován psychodiagnostický test Hermanna Rorschacha v knize <i>Psychodiagnostik</i>
1925	Ernst Kretschmer popisuje svoje pozorování o vztazích tělesné stavby k osobnosti a mentálním chorobám
1926	Publikován test Kresby lidské postavy Florence L. Goodenoughové
1927	První vydání Vocational Interest Blank for Men od Edwarda Kellogga Stronga
1928	<i>Zprávy o studiích charakteru</i> Huga Hartshornea a Marka Arthura Maye
1930	David M. Levy a Samuel J. Beck popularizují Rorschachův test ve Spojených státech amerických

- 1935 Henry Murray se svými spolupracovníky vyvíjí Tematický apercepční test (TAT); publikována Humm-Wadsworthova temperamentová škála
- 1938 Henry Murray publikuje *Explorations in Personality*, popisující teoretické základy Tematického apercepčního testu; publikován Benderův Vizuomotorický tvarový test pro diagnostiku osobnosti a organického poškození mozku; první vydání *Mental Measurement Yearbook*
- 1939 Publikovány Kuderova metoda Preference Record a inteligenční test Wechsler-Bellevue
- 1940 Jay P. Guilford používá faktorovou analýzu pro konstrukci osobnostního inventáře
- 1941–1945 U. S. Office of Strategic Services vyvíjí situační testy
- 1942 William Herbert Sheldon a Stanley Smith Stevens zveřejňují výzkum o vztazích mezi tělesnou stavbou a temperamentem
- 1943 Publikován Minnesotský multifázový osobnostní inventář od Starka R. Hathawaye a Johna Charnleye McKinleye
- 1949 Publikován dotazník 16 PF od Raymonda B. Cattella
- 1953 William Stephenson vyvíjí metodu Q-třídění
- 1954 Publikována práce Paula E. Meehla *Clinical Versus Statistical Prediction (Klinická versus statistická predikce)*
- 1955 Publikována práce George A. Kellyho *Psychology of Personal Constructs (Psychologie osobních konstruktů)*
- 1958 Lawrence Kohlberg publikuje první verzi Škály morálního usuzování
- 1970–1990 Zvyšující se využití počítačů při vývoji, administraci, skórování a interpretaci nástrojů a postupů diagnostiky osobnosti
- 1972 Publikován Self-Directed Search J. Hollanda pro měření zájmů a osobnosti; ve Spojených státech amerických publikováno a široce přijato pravidlo Model Penal Code pro vymezení právní nezpůsobilosti, díky kterému se u soudů ve Spojených státech amerických začala používat psychologická diagnostika
- 1974 Publikován a popularizován výzkum osobnosti typu A, náchylné ke koronárním onemocněním, od Meyera Friedmana a Raye H. Rosenmana
- 1975–1980 Rozvoj metod behaviorální diagnostiky
- 1985 Publikovány Standardy pro pedagogické a psychologické testování
- 1987 Publikována revize Kalifornského psychologického inventáře
- 1989 Publikován MMPI-II
- 1994 Publikováno čtvrté vydání Diagnostického a statistického manuálu (DSM-IV)
- 1997 Publikována třetí verze WAIS-III
- 1998 Publikován Thirteenth Mental Measurement Yearbook (13. Burosova ročenka mentálního měření)

Předchozí popis vývoje psychologického testování je téměř prostý kontroverzí a slepých uliček, jaké bychom při vývoji libovolné vědecké disciplíny oprávněně očekávali. Lze se tedy domnívat, že autoři podobných výkladů historie měření v psychologii tyto kontroverze a slepé uličky buď nevnímají, nebo se jejich existenci snaží zakrýt. Naštěstí existuje Michellova (1999) kniha, která je pro zájemce o peripetie vývoje přístupů k měření v psychologii podstatně zajímavější. Na následujících stranách z ní budeme poměrně detailně čerpat.

Michell (1999) uvádí, že zakladatelem kvantitativní tradice v psychologii byl bezesporu Gustav Theodor Fechner, původně fyzik věnující se psychologickým otázkám v rámci tzv. psychofyziky, který v roce 1860 publikoval knihu *Elemente der Psychophysik*. Jeho přístup se zakládal na práci se dvěma základními kategoriemi – kauzalitou (kterou se pokoušel zkoumat experimentálně) a kvantitou (pro jejíž zkoumání potřeboval měření). Svými postupy v psychofyzice v podstatě ustanovil modus operandi kvantitativní psychologie, ve kterém krátce po něm pokračoval např. Wilhelm Wundt, který v Lipsku v roce 1879 založil svoji laboratoř.

Myšlenka kvantitativního zkoumání, respektive přímo měření nějakých vlastností psychických jevů přitom v tehdejší době nebyla zdaleka samozřejmá a měla množství autoritativních odpůrců. Mnozí filosofové možnost kvantifikace duševních jevů popírali. Například Thomas Hobbes, John Locke, Immanuel Kant nebo René Descartes, jejichž myšlenky jsou díky jejich následovníkům dodnes součástí proudu tzv. fenomenologické psychologie, zastávali názor, že žádné psychologické atributy nejsou kvantitativní. Tomuto názoru říká Michell (1999) „**námítka vůči kvantitativnosti**“ (angl. *quantitativity objection*), a pokud by byla oprávněná, používání jakéhokoli měření v psychologii by nepřipadalo v úvahu.

Vědecká revoluce v 17. století ale vytvořila ideál vědy, který byl kvantitativní. A zakladatelé psychologie chtěli, aby byla kvantitativní vědou patřící do úspěšného a sociálně oceňovaného proudu přírodovědy; koneckonců její zakladatelé byli většinou přírodovědci – Wundt byl fyziolog, Fechner fyzik, Galton polyhistor (viz výše). Takže se pokusili po vzoru přírodních věd o vytvoření vlastních postupů psychologického měření na základě předpokladu, že jimi zvolené psychologické atributy měřit lze. Tento názor nazývá Michell (1999) „**teze měřitelnosti**“ (angl. *measurability thesis*).

To bylo základem mnoha neshod. Zatímco např. již zmíněný Wundtův a Galtonův pokračovatel James McKeen Cattell (1890) přísahal v psychologii na nutnost používat měření a experimentální výzkumné plány, německý filosof a psycholog Franz Clemens Brentano (1874) již dříve ve své knize *Psychologie vom empirischen Standpunkte* (*Psychologie z empirického hlediska*) tvrdil, že nepotřebujeme používat měření a matematiku tam, kde nenacházíme měřitelné veličiny.

Součástí převládající vědecké ideologie 19. století (v případě mnoha autorů dodnes) nejen v přírodovědě, ale i v psychologii byl názor, že vše ve světě je tvořeno čísly. Tento metafyzický předpoklad se nazývá **pythagoreismus** a spolu s dalším názorem, že metody úspěšně používané v jednom vědním oboru budou úspěšně použitelné i v ostatních oborech – tzv. **scientismem**, vedl k tzv. **kvantitativnímu imperativu**, který lze shrnout do tvrzení, že věda znamená měření. To ale samozřejmě není pravda – i mezi nekvantitativními atributy mohou panovat kvantitativní vztahy (a naopak) a obory přírodovědy se nevyhýbají používání kvalitativních postupů. Dokonce ani samotná matematika není zdaleka výhradně kvantitativní obor. Přesto tento zcela iracionální postoj v psychologii byl a stále je velmi silný i v současnosti.

Vědecká psychologie přitom původně kvantitativní nebyla – např. David Hume formuloval asociační zákony, které sice byly inspirovány newtonovskou fyzikou (zákony gravitace), ale byly chápány jako kvalitativní. Problém kvantitativní psychologie, který podle Michella (1999) dodnes nebyl adekvátně vyřešen, tedy spočívá v přehlížení faktu, že vyvrácení námitky vůči kvantitativnosti psychologických atributů je úkol pro empirický výzkum. To znamená, že kvantitativní povahu psychologických atributů nelze brát jako samozřejmost, ale je nutné ji empiricky prokázat. Jedině tak je možné podpořit tezi měřitelnosti.

Předtím, než budeme pokračovat ve výkladu historie měření v psychologii, je třeba zdůraznit, že někteří autoři rozpoznali tuto potřebu kritických empirických důkazů kvantifikovatelnosti psychologických atributů. Michell (1999) uvádí, že prvním z nich byl skotský filosof Thomas Reid, který už v roce 1748 vydal *Essay on Quantity*, ve kterém popíral možnost měření atributů, které postrádají kvantitu. První rigorózní diskusi o této otázce obsahuje esej Hermanna von Helmholtze z roku 1887 nazvaný *Zählen und Messen erkenntnistheoretisch betrachtet* (*Číslování a měření z epistemologického hlediska*). To se už nacházíme v historii matematiky. Bez matematiky se ale při úvahách o povaze měření obejít nelze.

Když se totiž zabýváme měřením v psychologii, nelze skočit rovnýma nohama k testům a dotazníkům. Dávno předtím se první psychologové, dnes nazývaní jako psychofyzici nebo psychofyzikové², snažili měřit intenzity počtů. A právě z hlediska Helmholtze a dalších tyto pokusy o měření nebyly považovány za uspokojivé. K Fechnerovi se ještě záhy vrátíme, avšak detailní diskuse o nedostacích jeho názorů na měření (a názorů podobných) by byla příliš rozsáhlá. Můžeme rovněž konstatovat, že námitky nebyly ze strany kritizovaných uznány.

Teoreticko-metodologická odbočka

Celá Michellova kniha se odkazuje na kritický bod v dějinách kvantitativní psychologie, a to rok 1951, kdy americký experimentální psycholog a psychofyzik Stanley Smith Stevens vydal knihu *Handbook of Experimental Psychology*, v jejíž první kapitole formuloval svoje názory na měření. Tato práce měla neuvěřitelný vliv na myšlení psychologů až do dnešní doby, a to, jak se ve shodě s Michellem domníváme, vliv značně neblahý. Předtím, než se budeme dále zabývat vývojem, který k přijetí Steversoných názorů v psychologii vedl, rozvedeme poněkud šířeji Michellovu argumentaci.

Podle Michella (1999) má kvantifikace ve vědě dvě stránky – jednou z nich je tzv. **vědecký úkol** (mohli bychom ho ztotožnit s „důrazem na ontologii“ u Borsbooma, Mellenbergha a van Heerdena, 2004³ – viz kap. 6), na který Stevens podle jeho názoru rezignuje (viz níže), a tzv. **instrumentální úkol**, spočívající pouze v řešení konkrétních metodických nebo metodologických problémů bez dostatečného teoretického odůvodnění (u Borsbooma et al., 2004, jako „důraz na epistemologii“ – viz kap. 6). Problém měření v psychologii podle Michella (1999) a Borsbooma et al. (2004) spočívá v tom, že psychologické teorie nejsou natolik dobře zformulované, aby jejich součástí byly vlastní definice měření příslušných (kvantitativních) atributů, které budou nedílnou součástí těchto teorií. V současné době psychologové na mnoho úkolů, které by měli plnit sami, „najíímají lidi“ z vlastního oboru nebo jiných oborů – a na řešení otázek měření mají psychometriky, kteří část svých

2 Rozhodli jsme se ve slově „psychofyzika“ a slovech příbuzných zdůraznit část -fyzika – užíváme proto jak výraz „psychofyzici“, tak „psychofyzikové“, avšak pouze variantu „psychofyzikální“ apod.

3 Dále jako Borsboom et al. (2004).

povinností přesunuli na statistiky. To je však iluze. Práci na tvorbě kvantitativních psychologických teorií a z nich vyplývajících postupů měření nemůže udělat nikdo jiný než právě psychologové.

Například Fechner se domníval, že svými postupy na námitku vůči kvantitativnosti odpověděl, ale ve skutečnosti obešel vědecký úkol kvantitativnosti a přešel rovnou k instrumentálnímu. Jeho definice měření ještě byla zcela v souladu s pojetím měření přírodních věd – „měření kvantity spočívá v zjištění, kolikrát je v něm obsažena jednotková kvantita stejného druhu“ (Fechner, 1860, s. 38; cit. dle: Michell, 1999). Fechner přitom stále ještě nevěřil tomu, že probandi dokážou přímo posoudit kvantitativní strukturu svých vjemů, Stevens už ale ano. Jinými slovy, Fechner byl poměrně opatrný v tom, co nazýval měřením, ale pro Stevense už bylo měřením téměř jakékoli přiřazování čísel řízené nějakými pravidly.

Do čtyřicátých let 20. století totiž v psychologii stejně jako v ostatních vědních oborech převládala **klasická koncepce měření**. Nejednalo se ovšem o tzv. klasickou teorii testů (viz kap. 5, 6, 9 a 10), ale o koncepci měření, která byla běžná v přírodních vědách. Teprve v padesátých letech převládla podpora Stevensova pojetí měření, pravděpodobně díky tomu, že kvantitativní psychologie zcela přijala modus operandi vyhovující této definici – to znamená, že námitka vůči kvantitativnosti byla obecně ignorována a bylo rutinně prováděno množství procedur generujících čísla, prohlašovaných za měření. Psychologie v té době (a bohužel ani v současnosti) zřejmě neměla morální ani intelektuální zdroje pro redefinici tohoto centrálního vědeckého pojmu.

Praktikalismus kvantitativní psychologie

Jedním z důležitých faktorů, který podle Michella (1999) ovlivnil přijetí defektního pojetí měření v psychologii, je tzv. **praktikalismus**, názor, že úspěch vědy leží především v praktických aplikacích (Anderson, 1962; cit. dle: Michell, 1999). Podle Michella (1999) může praktikalismus podvracet vědu tím, že hledá snazší cíle a ignoruje obtížnější. Praktikalismus zastávali převážně psychologové ve Spojených státech amerických, kde dominovala pragmatická filosofie, rozvíjená i mnohými z tamních psychologů. Michell (1999) uvádí, že existoval zásadní rozdíl mezi německými univerzitami fungujícími v humboldtovské tradici, které se věnovaly i zkoumání nepřilíš praktických otázek, oproti univerzitám ve Spojených státech, kde byl větší zájem o praktičtější otázky

psychologie. Například přední americký představitel pragmatismu William James údajně kritizoval psychofyzikální výzkumy jako zcela odtržené od života.

Dalším významným autorem ovlivněným praktikalismem byl britský psycholog Charles Spearman, který naplňoval kvantitativní výzkumný program inspirovaný Galtonovým výzkumem dědičnosti intelektových schopností v knize *Hereditary Genius* (1869). Byl prvním, kdo propojil výkony v mentálních testech s mentálními schopnostmi v rámci kvantitativní teorie (Spearman, 1904). Jeho myšlenky mají podle Michella (1999) pro testování v psychologii stejný význam jako Fechnerovy práce pro psychofyziku, ale současně nesou pečeť Fechnerova kvantitativního modu operandí. Základ Spearmanovy teorie je formulován na základě Galtona a Pearsona s pomocí pojmů „regrese“ a „korelace“.

Spearman studoval u Wundta v Lipsku, podobně jako James McKeen Cattell tam obhájil i doktorát, nějakou dobu strávil u Külpeho ve Würzburgu a u Müllera v Göttingenu (Spearman, 1930). Jeho odborný trénink tedy vycházel z psychofyziky a to mimo jiné vedlo k tomu, že ve své práci vycházel z Fechnerova modelu. Údajně to platí i o Karlu Pearsonovi. Oba byli podle Michella ovlivněni kvantitativním imperativem, oslepujícím před teoretickými tématy.

Jak již bylo uvedeno, praktikalismus byl mnohem běžnější ve Spojených státech – např. James McKeen Cattell a jeho žák Edward Lee Thorndike zdůrazňovali důležitost aplikovaného měření pro školství, průmysl a armádu. Dalšími představiteli prakticky používaného měření byli američtí psychologové Henry H. Goddard, Lewis M. Terman a Robert M. Yerkes (viz např. Gould, 1998). Dá se pochopit, že procedury připomínající měření posílily pozici psychologie ve společnosti, protože pojem měření je vždy spojován s přesností a objektivitou, ať už to je, či není pravda. Michell (1999) konstatuje, že kvantitativní psychologie nabízející pomoc při řešení sociálních otázek se stala vysoce obchodovatelnou komoditou.

Trajektorii praktikalismu v kvantitativní psychologii lze ilustrovat na práci představitelů tří generací americké psychologie – Jamese McKeen Cattella, Edwarda Lee Thorndika a Trumana Lee Kelleyho.

James McKeen Cattell byl podle mnoha zdrojů jedním z prvních profesorů psychologie ve Spojených státech s obrovským vlivem na vývoj americké psychologie. Po studiu u Wundta (v letech 1883–1886) a kontaktech s Galtonem se snažil od psychofyzikálního měření přejít ke studiu intelektových schopností. Jeho přesvědčením byl výše zmíněný kvantitativní imperativ

a klasické přírodovědné pojetí měření (tedy nikoli tzv. klasická teorie testů), které spočívalo v určení hodnoty nějaké veličiny v pojmech standardní jednotky (např. metru, litru nebo kilogramu). Místo pro měření v psychologii bylo podle jeho názoru stále otevřené a obor byl otevřen kvantitativním i kauzálním (tehdy se používal pojem „genetickým“) vysvětlením.

Edward Lee Thorndike obhájil v roce 1898 doktorát na Columbia University právě u Cattella, poté se stal součástí hnutí mentálního testování a stal se jeho vůdčí postavou (Boring, 1957; cit. dle: Michell, 1999). V roce 1904 publikoval pro tuto oblast klasickou knihu *An Introduction to the Theory of Mental and Social Measurements*. V roce 1918 formuloval své „Credo“⁴, ve kterém se podle Michella (1999) vyznal z víry v pythagoreismus a kvantitativní imperativ a vyjádřil v něm množství pseudovědeckých, kvazináboženských a metafyzických myšlenek. Přesto Thorndikovy práce podle Michella (1999) slibovaly pokrok v psychometrickém myšlení.

Thorndike se zabýval počtem správně zodpovězených položek jako hlavním indikátorem testového výkonu. Jenže tento celkový skóre obsahuje jen velmi kusé informace – důležité je také vědět, jak probandi odpovídali na jednotlivé položky. Dva probandi totiž mohou podat jiný výkon z hlediska jednotlivých položek, a přitom získat stejný celkový skóre. Měření na základě relativní pozice je pouze monotónní transformace pozorovaných skóre (Michell, 1999), nikoli měření v klasickém smyslu slova. Psychologové ale byli připraveni Thorndikovi věřit.

Pochybnosti o takto prováděném měření (a zájem o vědecká témata spojená s měřením) vyjadřoval např. Edwin G. Boring, nejprve experimentální psycholog a později jeden z prvních historiků psychologie, který byl na Harvardu školitelem autora kritizované definice měření S. S. Stevensa. Samotná pořadí celkových skóre podle Boringova názoru nestačí a „psychologie ještě není připravena na psychologické měření ve striktním slova smyslu“ (Boring, 1920, s. 32; cit. dle: Michell, 1999).

Další představitel americké tradice testování Truman Lee Kelley byl Thorndikův žák a určitou dobu vedoucí americký psycholog-statistik. Na Boringovu kritiku psychologického měření reagoval požadavkem solidního statistického základu – jak podotýká Michell (1999), nikoli logického základu. Teorie testů

4 „Cokoli existuje, existuje v nějakém množství. Abychom to kompletně poznali, je nutné poznat kvalitu stejně jako kvantitu“ (Barrett, 2003).

se v té době začala postupně stávat součástí statistiky, což mělo podle Michella dva důsledky. Tím prvním bylo, že kvantifikace přestala být považována za logickou, věcnou nebo teoretickou, ale namísto toho za čistě statistickou otázku, což možná i v důsledku toho, že psychologové většinou nejsou právě kompetentní statistici, vedlo k dalšímu důsledku, že základní otázky měření se v psychologii nadále přestaly řešit. Kelley se odkazoval k příkladu fyziky bez ohledu na to, že psychologie nebyla v podobné situaci jako fyzika. Odkazování k fyzice a dalším přírodním vědám jako kanonickým příkladům věd, ve kterých se používá měření, je argument, který se používá dodnes.

Důsledky praktikalismu byly patrné už na začátku 20. století, což lze ilustrovat následujícím příkladem. Do roku 1930 bylo vytvořeno mnoho testů, které měly měřit různé mentální jevy (viz Aiken, 1999). Mnoho úsilí bylo věnováno instrumentálnímu úkolu kvantifikace, ale téměř žádné úkolu vědeckému. To vedlo ke vzniku velkého množství vzájemně neslučitelných východisek pro tvorbu testů. Symptomatická byla diskuse o různých definicích inteligence, kterou v roce 1921 uspořádali lidé kolem časopisu *The Journal of Educational Psychology*. Pozvali do ní 17 předních výzkumníků, kteří se shodli pouze na jediné vlastnosti inteligence, kterou byla všemi předpokládána kvantitativní struktura tohoto konstruktu. Následující citát je z vědeckého hlediska vlastně velmi legrační a představuje triumf praktikalismu nad kritickým vědeckým zkoumáním: „Naše mentální testy něco měří, můžeme nebo nemusíme se starat co, ale je to něco, co je pro nás výhodné měřit, protože to posiluje naše znalosti o tom, co můžeme očekávat, že budou lidé dělat v budoucnu. Nástroj měření jako míra něčeho, co je žádoucí měřit, přichází první, a co se měří, přichází potom.“ (Kelley, 1929, s. 86; cit. dle: Michell, 1999). Kritické hlasy (např. Johnson, 1936) se sice vyskytly, ale mainstream jimi nebyl zasažen. A tzv. klasická teorie testů, dodnes dominující v psychologii, je teorie přijatá díky tomu, že odpovídá na otázky, místo aby je kladla.

Reprezentacionismus ve filosofii vědy a v psychologii

Koncem čtyřicátých let 20. století došlo k dramatické změně postavení klasického pojetí měření v psychologii, pravděpodobně také z toho důvodu, že začal klesat podíl kvantitativních psychologů první generace, kteří měli zkušenosti s měřením v jiných vědních oborech než jen v psychologii. Současně docházelo ke změnám ve filosofii matematiky, kde se především britská

a americká filosofie posunuly anti-realistickým směrem (Michell, 1999). V důsledku toho se oddělilo pojetí čísla a veličiny a došlo k reinterpretaci pojmu měření. V evropské filosofii vědy v té době dominoval logický pozitivismus a ve Spojených státech operacionismus (viz část Povaha měřených atributů v kap. 3.), dva dnes překonané názory.⁵ V roce 1966 vyšla Ellisova kniha *Basic Concepts of Measurement* (Ellis, 1966; cit. dle: Michell, 1999), která zcela opomíjela klasické pojetí měření (zmíněné několikrát v předchozím textu) a posouvala se směrem k **reprezentacionistickým pojetím**. Jejich tvůrci byli tři nejvýznamnější teoretici měření té doby – Bertrand Russell, Norman Robert Campbell a Ernest Nagel (Michell, 1999).

Není možné zde uvádět celou, značně detailní Michellovu (1999) argumentaci, která vychází z děl uvedených teoretiků měření. Pro naše účely snad stačí, když se zmíníme o základních myšlenkách. Britský filosof matematiky Bertrand Russell se zajímal mimo jiné o odlišení pojmů veličiny a čísla a tvrdil, že klasická teorie měření je nekoherentní. Přišel s novým pojetím veličin jako ordinálních (uspořádaných) struktur, pro které stanovil šest axiomů (Russell, 1903, s. 168; cit. dle: Michell, 1999). Spojení veličin a čísel pak podle jeho názoru nastává díky tomu, že je vytvořeno jedinečné a reciproční spojení (korespondence) mezi veličinou určitého druhu a čísly. Jedná se tedy o izomorfismus (zjednodušeně o relaci 1:1) mezi hodnotami veličiny a podmnožinou nějakého numerického systému (např. kladných reálných čísel). Podle Michella (1999) jde o první explicitní formulaci reprezentačního pojetí měření. Klasické pojetí měření, pracující s poměrem každé naměřené hodnoty k jednotce měření, kterým je reálné číslo jako součást měřeného systému, a předpokládající izomorfismus zachovávající uspořádání veličin a čísel, je vlastně pokusem poznat veličiny pomocí numerických vztahů. Podle Russella však nejsou čísla ve vědě nutná a zavádějí se z praktických důvodů. Russellův systém měření obsahuje pouze ordinální vztahy, nikoli aditivní strukturu. Tím je ale neuspokojivý, protože nevysvětluje, jak je možné, že se dá počítat, odčítat, násobit a dělit (Michell, 1999).

Britský fyzik a filosof vědy Norman Robert Campbell je další reprezentacionistický teoretik měření, který rozlišil tzv. fundamentální a odvozené měření.

5 Hlavní námitka proti logickému pozitivismu je ta, že významem věty nemůže být její pravděpodobnostní hodnota – rozumíme i větám, které jsou zjevně lživé nebo jejichž pravdivostní hodnotu neznáme. Hlavní námitka proti operacionismu je ta, že jev nemůže být definován operacemi jeho měření – jevy přece existují nezávisle na vědě nebo na jejich měření.

Každé **fundamentální měření** je podle jeho názoru umožněno tím, že určité rozsahy fyzikálních atributů se podobají číslům, takže např. numerická aditivita (operace sčítání) reprezentuje fyzikální aditivitu (kladení předmětů k sobě nebo za sebe, proto je možné skládat hodnoty jejich atributů). Takové měření zachovává uspořádání a aditivitu fyzikální veličiny, která je komutativní a asociativní. Při fundamentálním měření je tedy nutné (a současně možné) prokázat fyzikální aditivitu experimentálně, zatímco při **odvozeném měření** jsou fyzikální veličiny vyjádřeny jako funkce fundamentálních fyzikálních veličin.

Příkladem odvozeného měření je např. hustota, daná jako jednotka hmotnosti na jednotku objemu (příčemž hmotnost i objem jsou fundamentálně měřené veličiny). Tento typ odvozeného měření by se velmi hodil v psychologii, kde neexistuje jediný atribut, u něhož bychom mohli uvažovat o přímém (fundamentálním) měření, ale museli bychom přitom vyvinout nějaký jiný, nepřímý přístup k detekci aditivních vztahů mezi nepřímo měřenými veličinami. To však Campbell odmítal, což Michell (1999) považuje za zvláštní, protože k takovému možnému chápání schází jen málo.

Protože Campbellova koncepce měření bude zmiňována v souvislosti s tzv. Fergusonovou komisí (viz následující část Jednání Fergusonovy komise ve Velké Británii), zmíníme se zde o ní ještě poněkud detailněji. Campbell rozeznával tzv. veličiny A, které byly numerickými reprezentacemi experimentálně doložených analogií numerického sčítání (např. délka a jiné geometrické veličiny, hmotnost, objem), a veličiny B, které jsou definovány na základě objevení konstant v numerických zákonitostech týkajících se veličin A.

Jeden z nejvýznamnějších filosofů vědy Ernest Nagel, který pocházel z Nového Města nad Váhem, odkud se jeho rodiče přestěhovali na začátku 20. století do Spojených států, vycházel ve svém pojetí měření z Russella (1903) a Campbella (1920), ale i starších prací, jako např. Helmholtze (1887) a Höldera (1901) (vše cit. dle: Michell, 1999). Nesouhlasil s Russellem v tom, že měření nevyžaduje aditivní vztahy mezi hodnotami, a s Campbellem, že fundamentální a odvozené měření se v podstatě neliší. Byl však silně ovlivněn logickým pozitivismem a operacionismem, takže se filosofováním příliš nezabýval a rovnou „definoval měření jako numerickou reprezentaci empirických vztahů mezi objekty“ (Michell, 1999, s. 132). Problém s jeho pojetím spočívá podle Michella (1999) v tom, že odmítnutím jednotícího pojetí veličiny došlo k pozitivistické redukci pojmů na „empirické“ významy, čímž došlo ke ztrátě koherence u odvozeného měření, kde je velmi těžké nějaký „empirický“ význam najít.