

TVORBA ŠKÁL VE SPOLEČENSKÝCH VĚDÁCH

Ondřej Machek^a 

Robert F. DeVellis: *Scale Development*

Los Angeles: Sage, 2017. 4. vydání. 257 stran. ISBN: 978-1-5063-4156-9

Škály jsou široce používaným nástrojem pro měření konstruktů ve společenských vědách. Setkáme se s nimi v nejrůznějších typech odborných prací, od vysokoškolských kvalifikačních prací po studie ve významných vědeckých časopisech. Obecně jejich využití nalezneme v situacích, kdy je nutné měřit nepozorované, latentní proměnné, například lidské potřeby, podnikatelskou orientaci či sdílenou vizi. Existují samozřejmě i situace, kdy škála, kterou výzkumník potřebuje, dosud neexistuje. Publikace *Scale Development* je primárně zaměřena na postupy, jak škálu vytvořit vlastními silami, sekundárně pak na to, jak škály využít k měření. Autor knihy Robert F. DeVellis, emeritní profesor The University of North Carolina at Chapel Hill, USA, se věnoval problematice měření proměnných naprostou většinu svého profesního života. Odborně se zabýval především tím, jak ovlivňují sociální faktory chování a psychologii jedinců v souvislosti se zdravím. Hodnocená kniha je jedním z jeho neznámějších a nejcitovanějších děl.

V souladu se svou profesní historií autor v první kapitole argumentuje, že měření je základní činností vědce. Zároveň ovšem již v úvodu odrazuje od očekávání, že „správně vypadající“ otázky budou poskytovat spolehlivé výsledky. Náklady, způsobené špatným měřením, mohou vysoce převýšit jeho možné přínosy. Škálu autor definuje jako soubor položek zkombinovaných do kompozitního ukazatele, jejichž účelem je odhalit proměnné, které nejsou přímo pozorovatelné. Naměřené hodnoty jednotlivých položek jsou výsledkem působení určitého konstraktu. Škálu autor důsledně odlišuje od indexu, což je soubor položek, jež určuje výslednou hodnotu konstraktu. První kapitola se dále zabývá historií měření a čtenář, jehož zajímá praktické využití škál, pravděpodobně rychle přejde ke kapitole další.

Druhá kapitola se zaměřuje na pojem latentní proměnné, kterou autor chápe jako konstrukt, který nabývá určitých hodnot, ale který není přímo pozorovatelný. Latentní proměnná je příčinou hodnot naměřených v rámci jednotlivých položek škály a ve stejném

a Vysoká škola ekonomická v Praze, Fakulta podnikohospodářská, Česká republika
E-mail: ondrej.machek@vse.cz

duchu je znázorňována graficky v diagramu cest (angl. *path diagram*).² Zbytek kapitoly se zabývá kauzálními modely a konvencemi při jejich kreslení, tedy základy strukturálního modelování. Čtenář, který by se zajímal o úvod do strukturního modelování, by si ale měl sehnat jinou, specializovanější publikaci, protože postupy zde prezentované jsou popsané velmi zběžně. Ačkoliv zde autor pracuje s pojmem „measurement model“, překvapivě chybí jakákoli informace o konfirmační faktorové analýze (s tou se čtenář setká až v kapitole 6, a to pouze letmo).

Třetí kapitola se věnuje reliabilitě škály. Reliabilita je v dané kapitole chápána mnoha způsoby. Tím prvním je reliabilita coby podíl skutečné a pozorované hodnoty. Zvláštní důraz je kláden na vnitřní konzistence škál vyjádřenou prostřednictvím Cronbachova alfa. Autor podrobně rozebírá jeho definici, ale také jeho slabiny: Cronbachovo alfa bylo vyvinuto pro spojité proměnné (kdežto proměnné ve škálách jsou typicky ordinální), a tento ukazatel navíc s počtem položek ve škále automaticky roste (je možné ho zvyšovat prostým zvýšením počtu otázek). Autor uvádí přehled možných přístupů, nicméně zde není příliš návodný, a čtenář, který má zájem o větší vhled do dané problematiky, je tak odkázán na příslušné reference. Souhrnně se autor domnívá, že Cronbachovo alfa zůstává standardem a vznik alternativního mainstreamového přístupu není na pořadu dne, což bude pro čtenáře, který pracuje se škálami, zřejmě představovat hlavní přidanou hodnotu této kapitoly. Další části kapitoly se věnují souvisejícím tématům, a to zejména inter-rater reliabilitě, která je užitečná v situaci, kdy je určitý konstrukt hodnocen dvěma nebo více hodnotiteli, a test-retest reliabilitě, podle které by mělo platit, že opakované měření by za stejných podmínek mělo poskytovat stejné výsledky. Kapitola uvádí mnoho souvisejících teorií, což je jistě přínosné, ale zároveň to vede k tomu, že si čtenář odnese obecnou, neurčitou představu o tom, co vlastně reliabilita znamená.

Čtvrtá kapitola je zaměřena na pojem validity. Autor se zabývá třemi druhy validity. Prvním z nich je obsahová validita (*content validity*), která je teoreticky splněna, pokud je možné položky škály vnímat jako výsledek náhodného výběru mezi všemi možnými otázkami. V tomto ohledu by otázky vzhledem ke zkoumanému jevu neměly být ani příliš úzké, ani příliš široké. Autor uvádí mnoho praktických příkladů, neuvádí však praktické metody, jak obsahovou validitu zaručit. Druhým, velmi zběžně diskutovaným typem je tzv. kritériová validita (*criterion-related validity*), která je splněna, pokud výsledky měření odpovídají určitému kritériu nebo „zlatému standardu“. Praktická doporučení k hodnocení kritériové validity však ani zde nejsou uvedena. Třetím diskutovaným typem validity je konstruktová validita (*construct validity*), která se zabývá teoretickým vztahem

2 Česká terminologie je značně nejednotná. *Path diagram* také překládán jako „úsekový diagram“, „cestičkový diagram“, či se nepřekládá a používá se rovnou anglický výraz.

mezi zkoumanou proměnnou a jinými proměnnými. Autor doporučuje hodnotit tuto validitu prostřednictvím korelací. Kupodivu v celé kapitole není zmíněna konvergentní ani diskriminační validita, jejichž hodnocení bývá v kvalitních časopisech často vyžadované.

Pátá kapitola je mnohem rozsáhlejší než předchozí kapitoly a zabývá se praktickými doporučeními pro tvorbu škál. Proces tvorby škál je rozdělen do osmi kroků: jasné určení toho, co má být měřeno, vytvoření souboru otázek, určení formátu měření, expertní hodnocení souboru otázek, zvážení přidání otázek, které mají za cíl vyhodnotit možná zkreslení, poskytnutí souboru otázek respondentům, vyhodnocení položek škály a optimalizace délky škály. Po tomto procesu výzkumník disponuje reliabilní škálou a může zvážit vyřazení nadbytečných otázek, kdy lze s výhodou využít statistického softwaru. Pokud se čtenář zajímá o tvorbu vlastní škály, je možné pátem kapitolu považovat za nejdůležitější část knihy, protože obsahuje i mnoho praktických příkladů. Čtenář, který potřebuje pouze využít existující škály k dosažení výzkumného cíle, naopak tuto kapitolu může bez újmy přeskočit.

Šestá kapitola se zaobírá faktorovou analýzou jako metodou, která plní několik funkcí: odhalení počtu konstruktů (faktorů), které tvoří základ souboru otázek, redukce počtu položek ve škále a porozumění obsahu latentních proměnných. Kapitola poskytuje netechnický úvod do explorační faktorové analýzy (EFA). Autor popisuje, jak postupovat při volbě počtu faktorů, které budou extrahovány, jaké typy rotací s ohledem na řešený problém zvolit a jaké jsou rozdíly mezi faktorovou analýzou a analýzou hlavních komponent (PCA), které jsou v praxi často nesprávně zaměňované a mají jiný účel. Poslední část kapitoly se věnuje konfirmaci faktorové analýzy (CFA). Je však nutno dodat, že úvod do CFA je velice stručný (čítá přibližně dvě strany) a z praktického hlediska nepříliš použitelný, což je škoda, protože CFA bývá v kvalitních zahraničních časopisech vyžadována k prokázání konvergentní a diskriminační validity. Z hlediska „uživatele“ škál je absence rozsáhlejšího textu věnovaného problematice CFA poměrně zásadním nedostatkem této publikace.

Sedmá kapitola publikace je zaměřena na tzv. *item-response theory* (IRT, česky nejspíše „teorie odpovědi na položku“), která se zaměřuje na jednotlivé položky spíše než na škálu jako celek, zkoumá, jak konstrukt ovlivňuje jednotlivé položky, a využívá grafické způsoby zobrazení vlastností jednotlivých položek a škály jako celku. IRT je v knize popisována na 27 stranách; autor ji považuje za „atraktivní“, ale uzavírá, že se jedná o teorii značně subjektivní a neustálou, s čímž lze naprostoto souhlasit. Vzhledem k omezenému využití IRT v ekonomických, zejména manažerských vědách, je možné obsah kapitoly považovat za nepříliš užitečný, a to především pro čtenáře, který potřebuje škály tvorit či jen prakticky využívat.

Poslední, osmá kapitola se věnuje měření v širším kontextu. Mnoho zde uvedených informací mohlo být prezentováno již dříve. Autor například před započetím tvorby vlastní

škály zdůrazňuje nutnost provedení rešerše existující literatury, ujištění se, že vhodná škála dosud neexistuje, a ověření, že cíloví respondenti konstruktu porozumí. Dále autor připomíná, že způsob sběru dat by měl být konzistentní: data sbíraná prostřednictvím tištěného dotazníku mohou být fundamentálně odlišná od dat sbíraných prostřednictvím rozhovorů. Výzkumník by si měl být vědom toho, že výsledky často nelze zobecnit, a tuto skutečnost by měl přiznat jako omezení svého výzkumu. Tato kapitola tedy shrnuje základní principy empirického výzkumu v souvislosti se škálami a její obsah mohl být vyřčen již na začátku knihy.

DeVellisova kniha je poměrně čтивá a srozumitelná i čtenářům bez hlubších znalostí vědecké práce. Publikaci je možné doporučit začínajícím výzkumníkům, např. doktorandům, protože poskytuje mnoho praktických příkladů zaměřených na vytváření a používání škál. Kniha je primárně koncipována jako jakási „kuchařka“ pro tvorbu vlastní škály. Popsané postupy a příklady však mohou být užitečné i v případě, kdy je třeba korektně využít již existující škály jako prostředku pro dosažení výzkumného cíle. Čtenář zde například nalezne netechnický úvod do reliability, validity, nebo explorační faktorové analýzy. Je ale třeba podotknout, že pokud by výzkumným cílem bylo vytvoření a validace vlastní škály, bylo by vhodné se podívat po jiné publikaci. Ohledně mnoha důležitých kroků tvorby škál, například využití konfirmační faktorové analýzy, totiž předkládaná kniha není příliš návodná, ani nenabízí doporučení ohledně relevantní literatury.