

Kvantifikace jevů v sociálních vědách

Pozapomenuté dědictví Fergusonovy komise

Hynek Cígler | Katedra psychologie & INPSY
Fakulta sociálních studií, Masarykova univerzita

PhD Existence | 30.–31. 1. 2023 Olomouc

Běžný postup měření v sociálních vědách

1. Vytvořím set „položek“ (a případně odpovědí).

- WISC: „Co uděláš, když chceš uvařit čaj?“
- RSES: „Celkově vzato, jsem se sebou spokojen(a)“.
- d2, Stroopův test: specifický podnětový materiál.

2. Přípustným odpovědím přiřadím číselné hodnoty.

- WISC: špatně/správně, 0/1
- RSES: míra souhlasu, 1–6
- d2: Počet prvků z jednotku času.
- Stroop: reakční čas (resp. rozdíl reakčních časů).

Běžný postup měření v sociálních vědách

1. Vytvořím set „položek“ (a případně odpovědí).
2. Přípustným odpovědím přiřadím číselné hodnoty.
3. Číselné hodnoty považuji za čísla a sečtu je.
4. Ověřím psychometrické parametry tohoto součtu.
 - Reliabilita a validita; zpravidla s předchozím předpokladem numerické kvality.
 - Vnitřní konzistence, faktorová validita (korelace položek), souběžná validita...
5. Výsledný součet interpretuji jako odhad lidské vlastnosti.
 - Resp. jako numerickou reprezentaci psychického jevu.

Běžný postup měření v sociálních vědách

5. Výsledný součet interpretuji jako odhad lidské vlastnosti.

- Resp. jako numerickou reprezentaci psychického jevu.
- WISC: *Intelligence* (aka Guttmanova škála)
- RSES: *Self-esteem* (Likertova škála)
- d2: *Pozornost* (aka psychofyzika)
- Stroop: součet *rychlost zpracování*, rozdíl *selektivní pozornost* či *inhibice* (aka výsledky experimentálního výzkumu).

„Kde se vzala ta čísla?“

cit. dle Zbyněk Vybíral (megakonvice, někdy)

Zdroje měření v psychologii

Dvě hlavní tradice „měření“

- s výrazně odlišnou povahou psychologických konstruktů.

Psychofyzika

= matematická psychologie

- vztah stimulu a smyslového vjemu, modelování kognitivních procesů
- propojení fyziky a psychologie
- zaměření na obecně platné zákonitosti
- člověk často jako „měřicí nástroj“
- Wundt, Weber, Fechner...

Mentální testy

= psychometrika

- popis rozdílů mezi lidmi, resp. systematické vysvětlení pozorovaných rozdílů
- škálování
- zaměření na mezisubjektové rozdíly
- člověk je měřeným subjektem
- Galton, McKeen Cattell, Spearman...

Mentální měření – psychometrika

Návaznost na Galtona a měření inteligence.

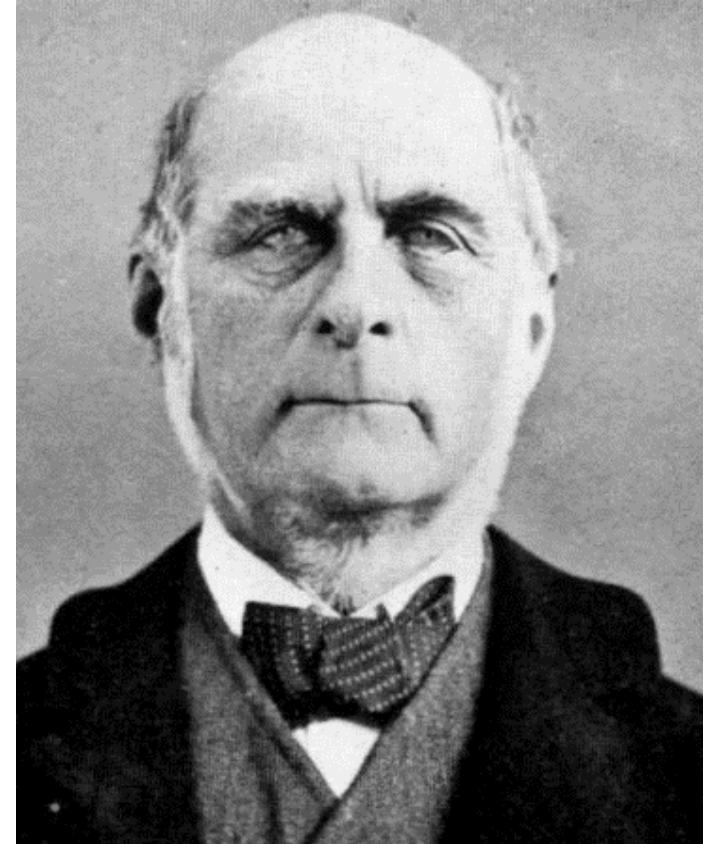
Základní předpoklad: existence mezilidských rozdílů s kvantitativní povahou.

První pokusy: součet správných odpovědí

- Řada potíží, zejména u osobnostních charakteristik.
- První tvůrci testů (např. Binet) byli velmi opatrní v užívání slova „*měření*“.

V první polovině 20. stol. proto vznikla řada kvantifikačních postupů.

- Hayes a Patterson (1921), Bogardus (1925), Thurstone (1928), Likert (1932), Guttman (1944), Osgood (1957) aj.



Psychofyzika – matematická psychologie

Psychofyzika se původně zabývala senzorymi vjemy.

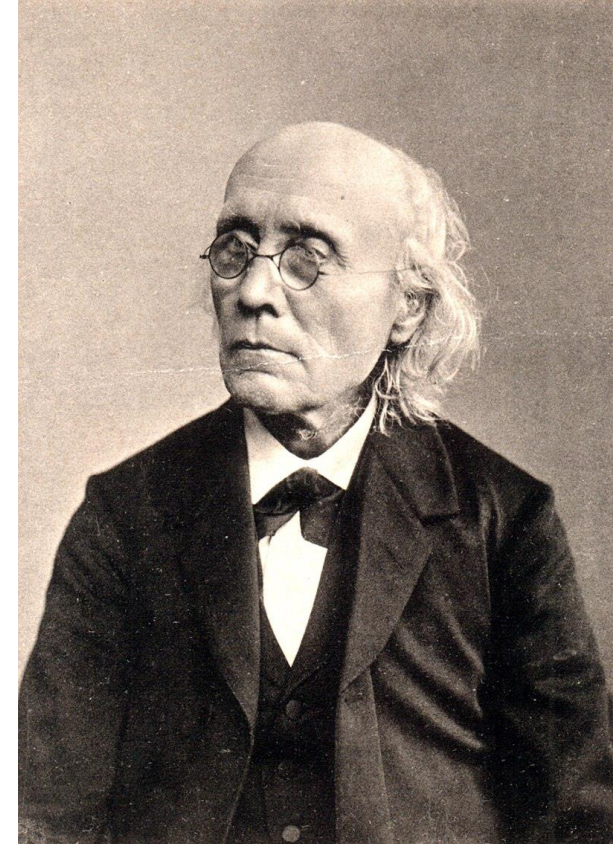
- Dnes je tradice patrná v „matematické psychologii“.

Obecně platné zákonitosti lidského vnímání, koncept „průměrného člověka“ (Quetelet, 1842).

Např. Weberův a Fechnerův zákon.

Člověk byl využíván jako „měřicí nástroj“ pro kvantifikaci jiných jevů (například síly akustického či světelného signálu).

- Psychofyzici byli mnohem méně, avšak stále ještě velmi zdrženliví se slovem „měření“.



Fergusonova komise (1932–1940)

British Association for the Advancement of Science.

- Sekce A (matematika a fyzika) a J (psychologie).
- Klíčové kvůli pronikání psychologie do fyziky.
 - son; logaritmická stupnice decibelu; kandela...

Základní otázky:

- Mají sensorické vjemy (tedy nikoli podněty) kvantitativní povahu?
- Pokud ano, je možné měřit jejich intenzitu?

Pokud totiž nelze měřit ani základní sensorické vjemy, nelze předpokládat, že by bylo možné měřit komplexnější psychologické konstrukty.

- Intelligence, osobnostní rysy, postoje...

Fergusonova komise (1932–1940)

- Ferguson, A., Myers, C. S., Bartlett, R. J., Banister, H., Bartlett, F. C., Brown, W., Campbell, N. R., ..., & Tucker, W. S. (1940). Final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *Report of the British Association for the Advancement of Science*, 2, 331–349.

10 bodů a přílohy. Hlavní závěry:

- **Shoda není možná**, komise pouze popsala rozdílná stanoviska svých členů.
- Zásadní **neshoda v definici měření**.
- Zcela klíčová role *nejmenšího rozlišitelného rozdílu a metody stejně se jevících intervalů*.
- Problémem není chyba měření; ta se zdá být (v psychofyzikálním výzkumu) přijatelná.
- Empirické pozorování není postačujícím důkazem pro existenci kvantitativního atributu.
- Lidé mohou vyjadřovat míru intenzity kvantitativně, protože jsou *naučeni* pracovat s kvantitami.
- Další matematické důkazy.

Co je to měření?

Extenzivní veličina: samotný atribut je aditivní.

- $3 \text{ cm} + 5 \text{ cm} = 8 \text{ cm}$.
- Rozdělením celku vzniknou části. Součet míry jejich atributů je roven původnímu celku.
- Umožňuje **přímé měření** srovnáním s etalonem, např. přiložením pravítka.
- Délka, hmotnost, objem, elektrický odpor, Δt .

Intenzivní veličina: atribut aditivní není, ale má kvantitativní povahu.

- $200 \text{ K} + 50 \text{ K} \neq 250 \text{ K}$.
- Každá část rozděleného celku bude mít stejnou míru atributu jako původní celek.
- Nelze „přiložit“ měřicí nástroj; umožňuje pouze **nepřímé měření**.
 - Campbel (1940): kvalita, nikoli kvantita předmětu.
- Hustota, teplota, tlak.

Měření: koordinační funkce

Funkce, která prováže pozorování s atributem.

Přímé měření: zpravidla jednoduchá lineární funkce $L = f(I) = x \cdot \delta I + I_0$

- x – naměřená hodnota; δI – jednotka; I_0 – referenční bod

Nepřímé měření: funkce využívající zpravidla více přímých a nepřímých veličin.

- Jen zřídka je lineární.
- Např. hustota: $\rho = f(m, V) = \frac{m}{V}$

Psychologové stáli před dvěma úkoly:

- 1. Vytvořit koordinační funkci.
- 2. Stanovit dostatečně spolehlivou jednotku, resp. referenční bod.

Počátky měřicích škál

Kategorické či ordinální pozorování bylo nutné provázat s domnělým kvantitativním, spojitým, intervalovým rysem.

Vizuální analogová škála (Hayes a Patterson, 1921).

- Apriori předpokládaná lineární koordinační funkce neobstála.

Metoda stejně se jevících intervalů (Thurstone, 1928).

- Namísto volby vhodné koordinační funkce využil předběžnou kalibraci podnětového materiálu tak, aby mohl výslednou funkci považovat za lineární.
- Pět různých modelů měření.
- Law of Comparative Judgement – vychází z Weberova-Fechnerova zákona.

Likertova škála (1932). Pragmatický přístup:

- **Metoda sigma:** Kalibraci na základě předpokladu normálního rozložení ve výzkumném souboru.
- „**Jednodušší**“ metoda: Z důvodu prakticky perfektní korelace začala být preferovaná.

Nejmenší rozlišitelný rozdíl

Metoda stejně se jevících intervalů

Nejmenší rozlišitelný rozdíl (JPD)

- Weber a Fechner.
- Lidem byly prezentovány podněty a měli vybrat jiný, právě rozlišitelný podnět s vyšší/nížší intenzitou:

$$\delta I = (I' - I)$$

Cílem bylo využít δI jako etalon pro měření intenzity smyslového vjemu.

- Propojila by fyzikální intenzitu a psychický jev.

Záhy se ale ukázalo, že δI záleží na míře intenzity podnětu. Fechnerův zákon:

$$\delta I = kI$$
$$I' = (k + 1)I$$

Metoda stejně se jevících intervalů (EAI)

- Rozpracoval Thurstone.
- Lidem byly prezentovány podněty o síle p, q, r . Měli vybrat podnět s , pro který by platilo

$$(q - p) = (s - r)$$

Alternativní postup pro tvorbu koordinační funkce.

- JPD: změny δI (resp. koeficient k) definuje vztah podnětu a škály.
- EAI: poměry $\frac{p}{q} = \frac{r}{s}$ definují škálu.

Existovaly ale další postupy.

Nejmenší rozlišitelný rozdíl Metoda stejně se jevících intervalů

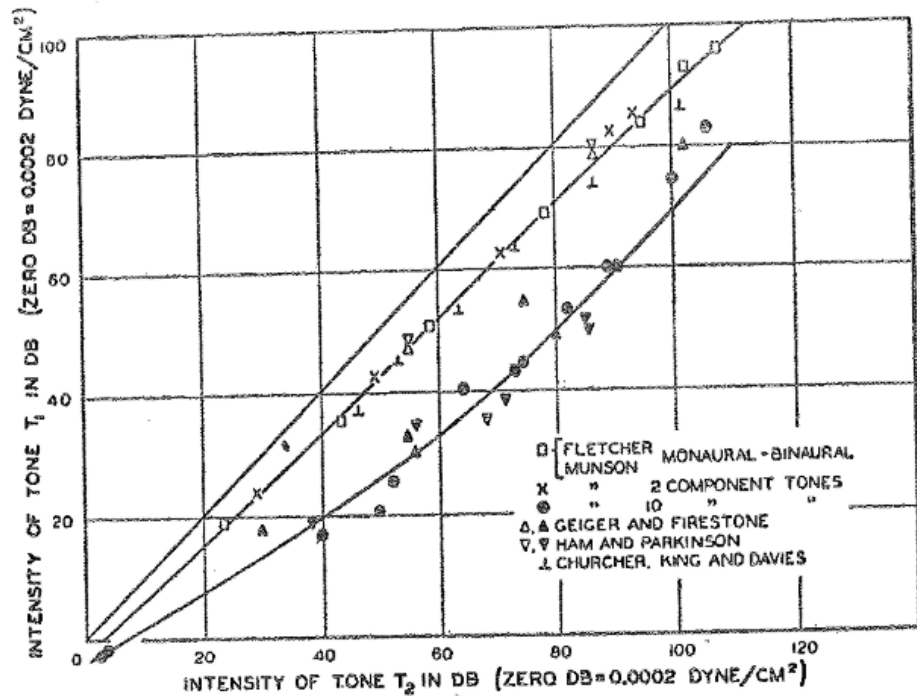


FIG. 41. The ordinate shows the intensity at which a tone (T_1) sounds half as loud (open figures) or a tenth as loud (solid figures) as another tone (T_2) whose intensity is given by the abscissa.

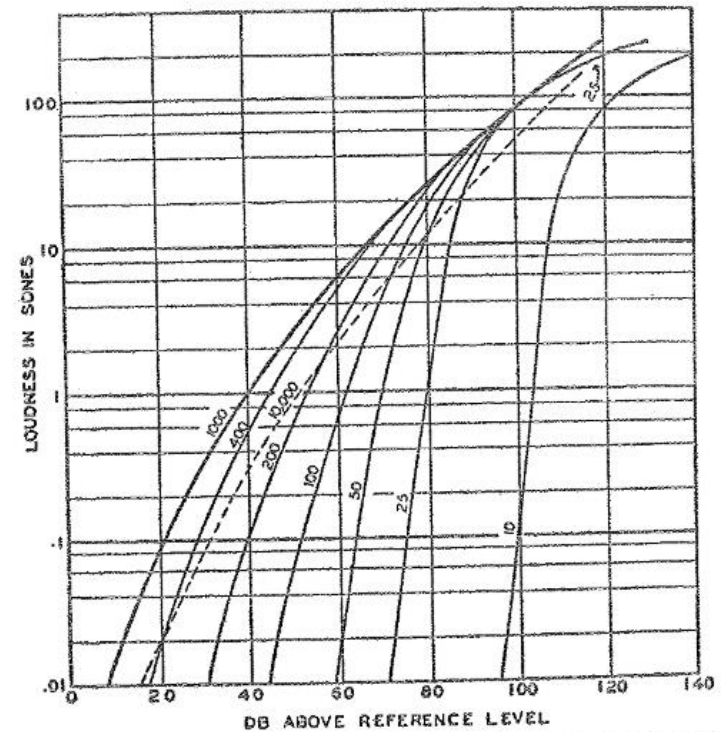


FIG. 43. The loudness-function. Showing how the perceived loudness of various tones depends upon the intensity of the stimulus. Frequency is the parameter.

Weberův-Fechnerův zákon

Weberův zákon:

$$\Psi(I) = k \cdot \ln I$$

Fechnerův zákon:

$$\Psi(I) = k \cdot \ln \frac{I}{I_0}$$

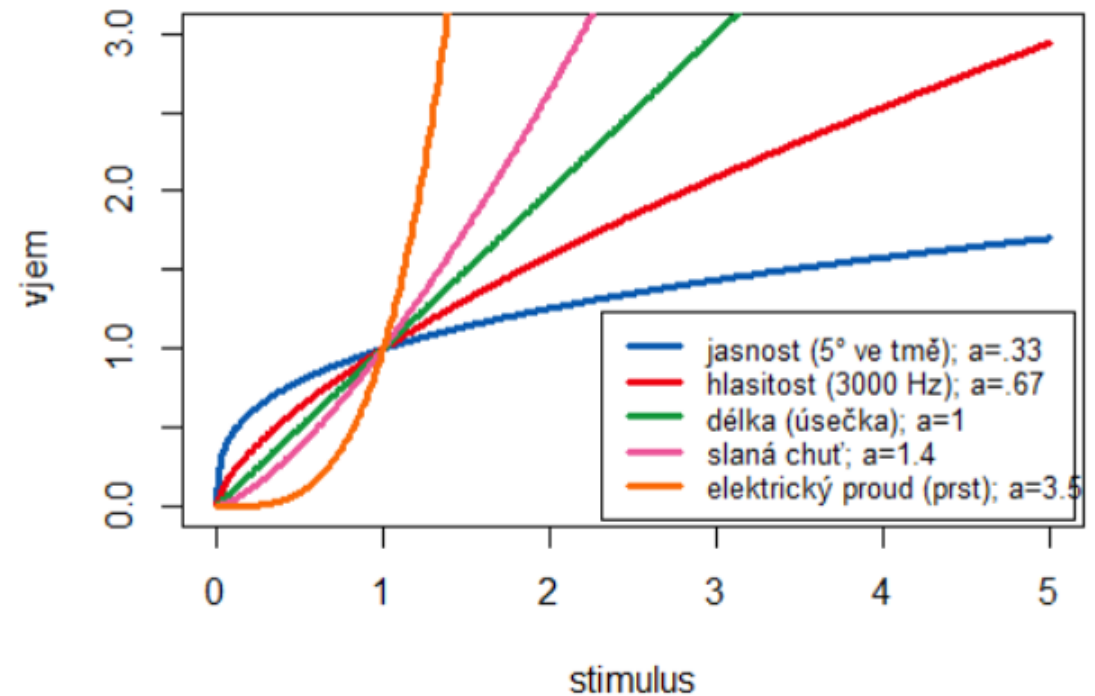
- Funguje dobře pro malé rozdíly (s velkou nejistotou odlišení podnětu)

Stevensův exponenciální zákon:

$$\Psi(I) = kI^a$$

- Funguje i pro větší rozdíly (s téměř jistou rozlišitelností vjemu).
- Umožňuje modelovat rozdílný vztah pro různé proměnné; např. délka $a=1$, hlasitost $a=0,67$,

Stevensův exponenciální zákon



Co je to měření:

Final Report, Příloha 2

„Measurement in its widest sense may be defined as the assignment of numerals to things so as to represent facts or conventions about them“.

- Final Report, Appendix 2 (**Campbell**, 1940).

1. Čísla („numerals“) jsou seřazená a číslování objektů reflektuje jejich „míru“.

2. **Podmínka A:** přiřazení musí být jedinečné.

- Pokud je jednomu objektu přiřazena magnituda, lze odhadnout magnitudu všech jiných objektů.
- Podle Campbella (1940) jediný známý způsob přiřazení: číslice (numerals) reprezentují čísla (numbers).

Podstatný důsledek: **měření není závislé na měřicím nástroji** (Campbell, 1940).

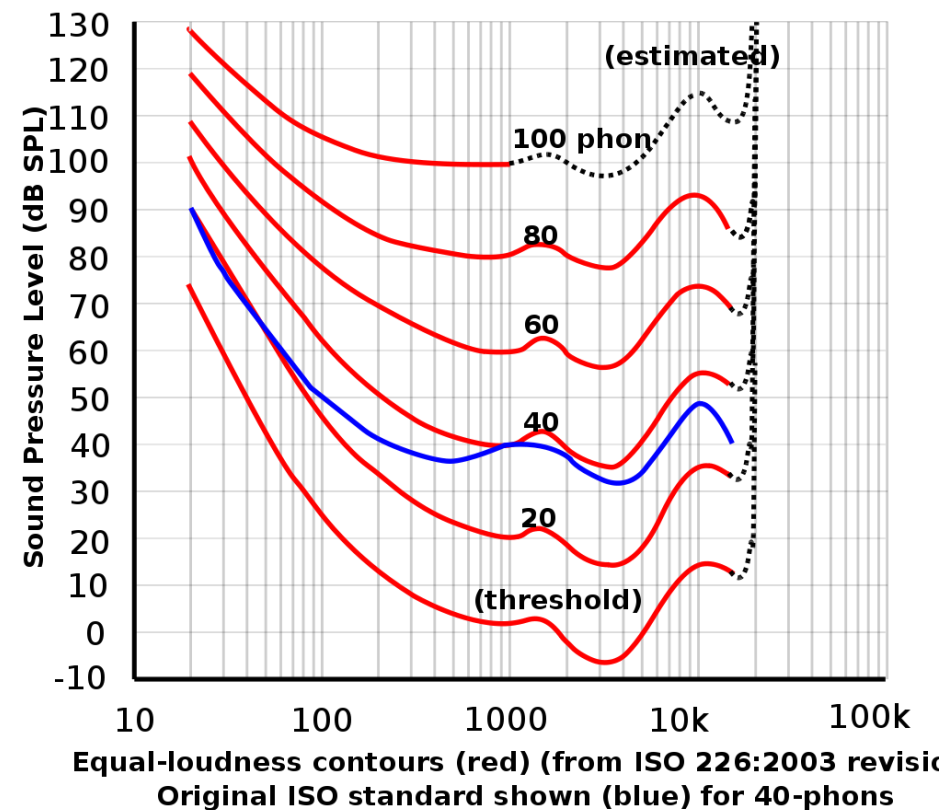
- Například teplota: teploměr funguje pomocí Boylova zákona, $pV = c$, změna v konstantně je proporčně rovna změně teploty, $\delta c = \delta T$.
- Zákon platí v určitém rozmezí teplot; předpokládáme ale, že měřená vlastnost platí daleko mimo jeho rozsah. Pokud ne, nejde o měření.
- Měřítka v různých úrovních rozsahu lze vzájemně kalibrovat.

Stanley S. Stevens

Měření hlasitosti v sonech a phonech (1936).

- Právě sony byly diskutovány v závěrečném reportu Fergusonovy komise.
- Exponenciální Stevensův zákon:
- $a = 0,3$ (referenční tón: 1 kHz)
- 0 son = hranice slyšitelnosti.

Operacionální definice měření (1946).



<https://en.wikipedia.org/wiki/Phon>

Stanley S. Stevens

Měření hlasitosti v sonech a phonech (1936).

- Právě sony byly diskutovány v závěrečném reportu Fergusonovy komise.
- Exponenciální Stevensův zákon:
- $a = 0,3$ (referenční tón: 1 kHz)
- 0 son = hranice slyšitelnosti.

Operacionální definice měření (1946).

A CLASSIFICATION OF SCALES OF MEASUREMENT

Paraphrasing N. R. Campbell (Final Report, p. 340), we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules. The fact that numerals can be assigned under different rules leads to different kinds of scales and different kinds of measurement. The problem then becomes that of making explicit (a) the various rules for the assignment of numerals, (b) the mathematical properties (or group structure) of the resulting scales, and (c) the statistical operations applicable to measurements made with each type of scale.

Operacionální definice měření

Operacionální definice měření padla na úrodnou půdu behaviorismu a logického pozitivismu...

... a vznikla **klasická testová teorie**.

Operacionalismus vůbec nezvažuje existenci měřeného atributu.

- Hodnota vědecké teorie je odvozena jinak.

Měřený atribut si tedy lze definovat libovolně...

- ... např. jako průměr/součet správných odpovědí.

Důsledkem je, že měřicí škála je neoddělitelně spjata s měřicím nástrojem.

- Měřicí nástroj definuje škálu.

Klasická testová teorie

Díky operacionalismu lze definovat libovolnou koordinační funkci. CTT volí lineární:

$$\tau = X + \epsilon$$

Protože X a ϵ spolu z definice nekorelují, vztah platí i pro rozptyl:

$$\sigma_{\tau}^2 = \sigma_x^2 + \sigma_{\epsilon}^2$$

Chyba je pak identifikovaná skrze korelaci paralelních testů (reliabilitu):

$$r_{xx'} = 1 - \frac{\sigma_{\epsilon}^2}{\sigma_x^2}$$

- Chyba měření tedy vyjadřuje pouze stabilitu měřicí procedury.
- Koncept reliability představený Spearmanem (1904)

Žádné pravé skóre ve skutečnosti neexistuje. Jde o myšlenkový experiment.

$$E(X) = \tau$$

Důsledky běžného postupu

Drtivá většina existujících pojetí validity vychází z antirealismu.

- Nezabývá se existencí měřeného atributu.

Nejstarší pojetí: prediktivní a kriteriální validita, později souběžná.

- Instrumentální využití „indexů“ a zdrženlivost s termínem „měření“.

Konstruktová validita (Cronbach a Meehl, 1955).

- Logický pozitivismus. Soulad pozorování (korelační matice) a teorie (MTMM matice).

Unifikovaná konstruktová validita (Messick, 1989, 1995).

- Instrumentalismus. Test musí sloužit svému účelu.

Různí kritici dlouhodobě upozorňují na problematičnost běžné praxe.

- Michell (1997; 1999), Borsboom (2005).
- Měření v psychologii je dominantně postaveno na CTT. Měříme tedy neexistující pravé skóry, s konstrukty však zacházíme, jako by existovaly. Reifikace.

Důsledky běžného postupu

Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197. <https://doi.org/10.1016/j.jad.2016.10.019>

Měření deprese v běžné klinické praxi.

Sedm měřících nástrojů: celkem 52 symptomů.

- Průměrný překryv nástrojů: 36 %.
- 40 % symptomů pouze v jediné metodě.

Lze srovnávat výzkumné výsledky *depressivity* s využitím takto odlišných metod?

Každý nástroj měří „jinou depresivitu“.

- Existuje tedy vůbec nějaká univerzální depresivita?

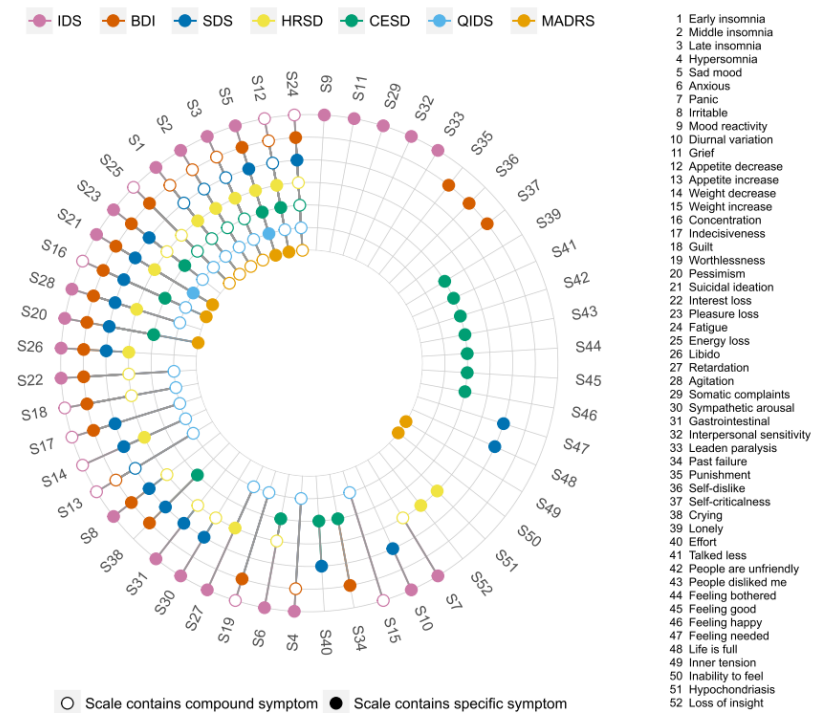


Fig. 1. Co-occurrence of 52 depression symptoms across 7 depression rating scales. Colored circles for a symptom indicate that a scale directly assesses that symptom, while empty circles indicate that a scale only measures a symptom indirectly. For instance, the IDS assesses item 4 hypersomnia directly; the BDI measures item 4 indirectly via a general question on sleep problems; and the SDS does not capture item 4 at all. Note that the 9 QIDS items analyzed correspond exactly to the DSM-5 criterion symptoms for MDD. Please see the online version for colors; in the black and white version, the circles represent (from outer to inner circle): IDS, BDI, SDS, HRSD, CESD, QIDS, and MADRS.

Důsledky běžného postupu

Lilienfeld, S. O., & Strother, A. N. (2020). **Psychological measurement and the replication crisis: Four sacred cows**. *Canadian Psychology*, 61(4), 281–288. <https://doi.org/10.1037/cap0000236>

1. Obsahová validita a spoléhání se na „název“ škál.
 - Škály se stejným názvem nemusí měřit to stejné.
 - Pro připomenutí: klasická testová teorie a operacionalismus.
 2. Ignorování chyby měření a reliability v laboratorních experimentech.
 - Přesvědčení, že pro výzkum postačuje nižší reliability (rovněž i Helmstadter).
 - Behaviorální pozorování (vysoce reliabilní) není totožné s měřeným rysem (vztah může být vágní).
 - A jaká je reliability experimentální manipulace?
 4. Důraz na konvergentní, nikoli divergentní validitu.
 - Konstruktově irelevantní rozptyl, nedostatek diferenciální validity.
 - Potíže zejména při výzkumu silně korelovaných jevů.
- (3. Náročnost sběru dat opravňuje malé velikosti vzorku.)

Další potíže: zdroj rozptylu

Borsboom (2005): Kvantitativní psychologie se zaměřuje na mezisubjektové rozdíly.

- Správná či špatná odpověď v inteligenčním testu je důsledkem konkrétních psychických procesů na úrovni jedince. Skóre v testu neříká nic o těchto procesech.
- Vysvětluje jev chování na úrovni jedince, nebo jen napříč lidmi?

Lokálně homogenní konstrukty: popisují *stejně* mezi- i vnitrosubjektové rozdíly.

- arousal, emoce

Lokálně heterogenní konstrukty: relevantní vnitrosubjektově, ale s jasnou mezisubjektovou strukturou.

- stres, pozornost, depresivita

Lokálně irelevantní konstrukty: pouze mezisubjektová struktura

- inteligence, osobnostní rysy

Další potíž: formativní konstrukty

Z hlediska kritického realismu existují dva modely latentních proměnných:

- reflektivní model
- formativní model

(Borsboom, 2005; Edwards a Bagozzi, 2000)

Z hlediska CTT přichází v úvahu jen formativní model měření.

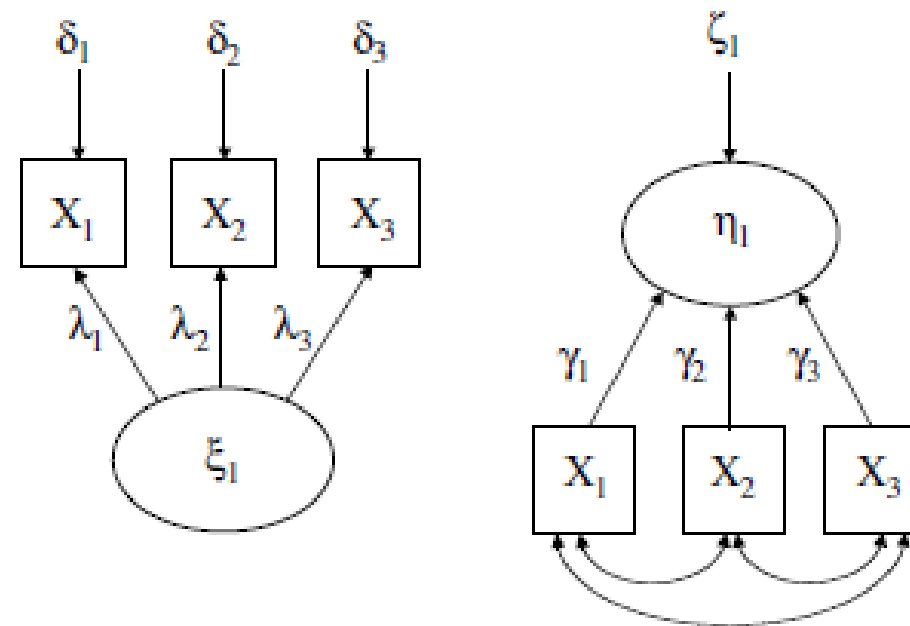


Figure 3.1. Two models for measurement. The figure in the left panel is the reflective measurement model. The X 's are observed variables, ξ is the latent variable, λ 's are factor loadings and the δ 's are error terms. The right panel shows the formative model. The latent variable is denoted η , the γ 's are the weights of the indicators, and ζ is a residual term.

Instrumentální pojetí měření

Psychologické „měření“ umožňuje tvořit predikce.

Je tedy užitečné v psychologické praxi.

- Meehlovy experimenty (1954): statistická vs. klinická predikce.
- Groove a kol. (2000): Metaanalýza 136 studií.
- Statistická predikce úspěšnější v 47 % (klinická jen 6 %).

Nevypovídá však o existenci a „kvalitě“ měřených konstruktů.

- Výsledky jsou „indexy“, které mohou sloužit k rozhodování.
- Nejde o kvantitativní reprezentaci skutečně existujících jevů.

Alternativní modely měření

Model latentních proměnných (faktorová analýza, teorie odpovědi na položku).

- Potíž: pozitivní manifold. Faktorovým modelem lze popsat cokoli.
- Ani model perfektně popisující data neprokazuje „existenci“ latentních rysů (faktorů).
- Modely zpravidla nepopisují data dostatečně dobře.

Reprezentační model měření (Debreu, 1960; Luce & Tukey, 1964)

- Alternativní reakce na Fergusonovu komisi.
- Stejně jako CTT operacionální přístup, avšak pracuje s aditivními operacemi.
- Vytvoří algebraickou strukturu z ordinálních dat.
- Příklady: Raschův model, Bradley-Terry model.
- Potíž: operacionalismus. Rovněž neprokazuje existenci atributu.

Nešťastná čísla?

Psychologické měření je skvělým nástrojem pro predikci.

- Kdo potřebuje individuální vzdělávací plán?
- Není klient rizikovým řidičem?
- Kdo bude lepším zaměstnancem?

Psychologické měření však jen málo přispívá k pochopení měřených jevů.

Východisko 1: Některé **psychické jevy mají kvantitativní povahu.**

- Psychologové by se měli zajímat o povahu měření a formalizovat své teorie, naučit se *měřit*.

Východisko 2: **Psychické jevy nemají kvantitativní povahu.**

- Psychologové by měly přestat zacházet s naměřenými čísly, jako by reprezentovaly skutečné lidské vlastnosti.

Příklad: Měření výšky

Rečka (2018): 26položkový dotazník výšky

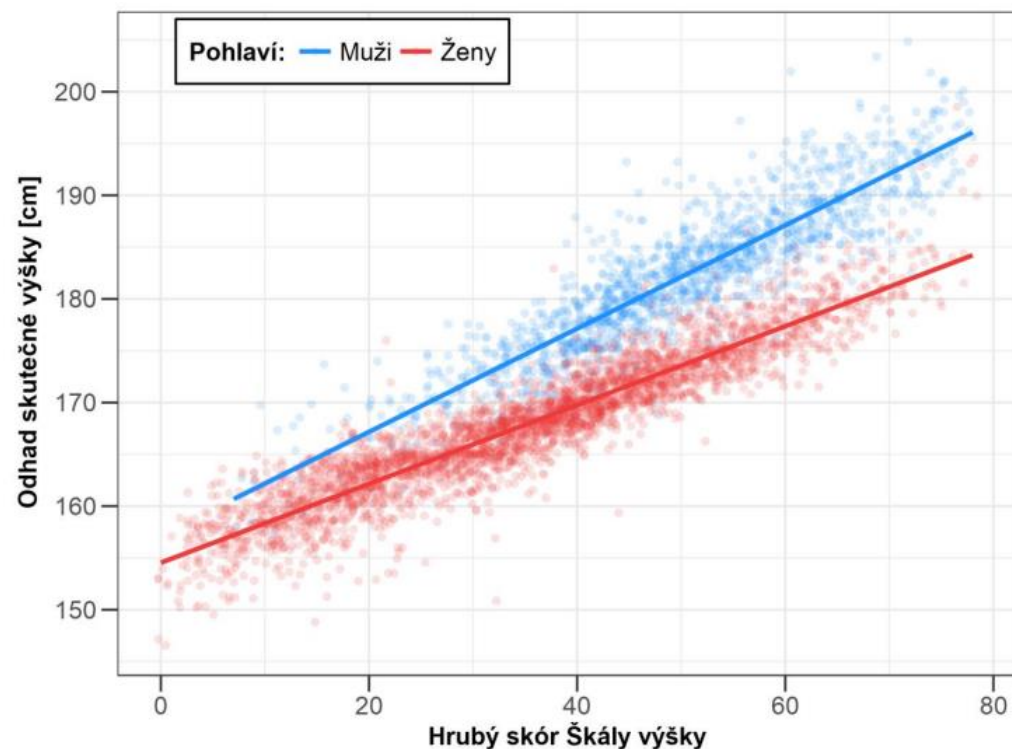
- N = 4885, 70 % žen; 13 pol. reverzních.
- I4: Mám vhodnou výšku na hraní basketbalu nebo volejbalu.
- I6: Slýchávám narážky na to, že jsem vysoký/á.
- I14: Jsem spíše nižší než muži mého věku.
- I21: V autobuse mívám dostatek prostoru pro nohy.

Vysoká vnitřní konzistence.

- Cronbachovo alfa: 0,967 (muži), resp. 0,976 (ženy).

Vysoká kriteriální validita se self-reportem výšky.

- Korelace: $r = 0,903$ (muži), resp. $0,932$ (ženy).
- Studie 1: reliabilita self-reportu $ICC > 0,999$.



Obrázek 24 Vztah mezi hrubým skórem Škály výšky a odhadem skutečné výšky

Příklad: Měření výšky

Pozitivně a negativně formulované položky vytvořily dva faktory.

- Ordinální MG CFA/EFA
- WLSMV estimátor
- lavaan
- Úpravy modelu: vyřazeny dvě položky, povolena reziduální kovariance u dvou položek (jsem větší/menší než...).

Co se stane, když do modelu vložíme skutečnou výšku v cm?

- self-report

	x2	df	CFI	TLI	RMSEA	SRMR
1F	17962,5	498	0,942	0,936	0,123	0,073
2F	7697,5	496	0,976	0,973	0,079	0,043
S-1 BI	5809,6	474	0,982	0,979	0,070	0,033
BI	2118,2	450	0,994	0,993	0,040	0,018
EFA	1771,8	452	0,996	0,995	0,035	0,015

Příklad: Měření výšky

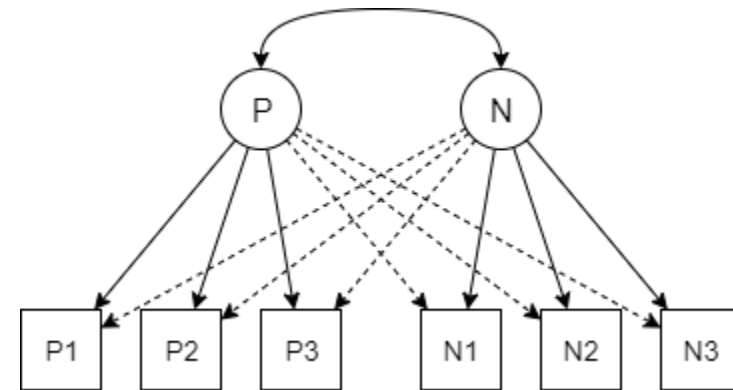
M0: Výchozím modelem je EFA se dvěma faktory.

Rotované řešení (šikmý geomin):

- Existují dva korelované latentní rysy (ženy $r = 0,727$, muži $r = 0,650$).
- Dohromady tvoří „konstrukt“ výšky.

Nerotované řešení:

- Obecný faktor výšky.
- Sekundární faktor – interindividuální rozdíly v odpovídání na pozitivní a negativní položky.

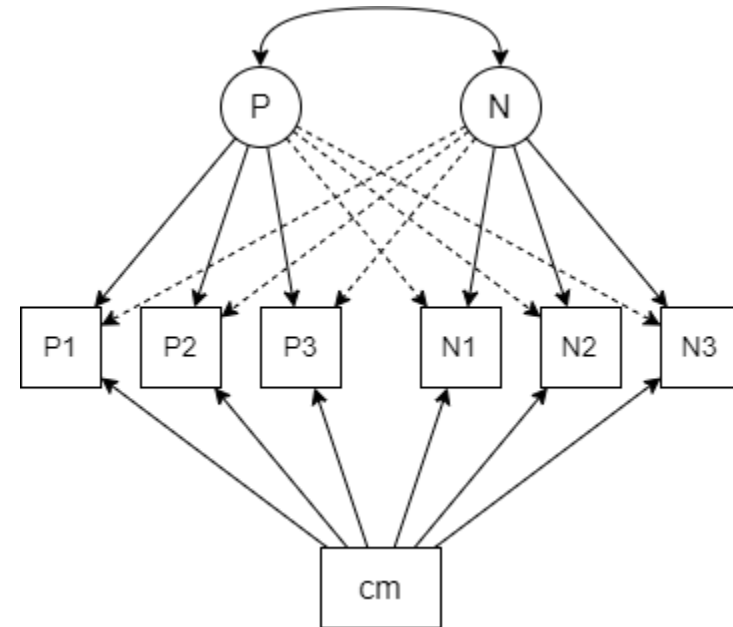


Příklad: Měření výšky

M1: Skutečná výška „sytí“ přímo položky dotazníku.

Latentní faktory popisují psychologické proměnné, které způsobují rozdíl odpovědí oproti očekávání na základě skutečné výšky.

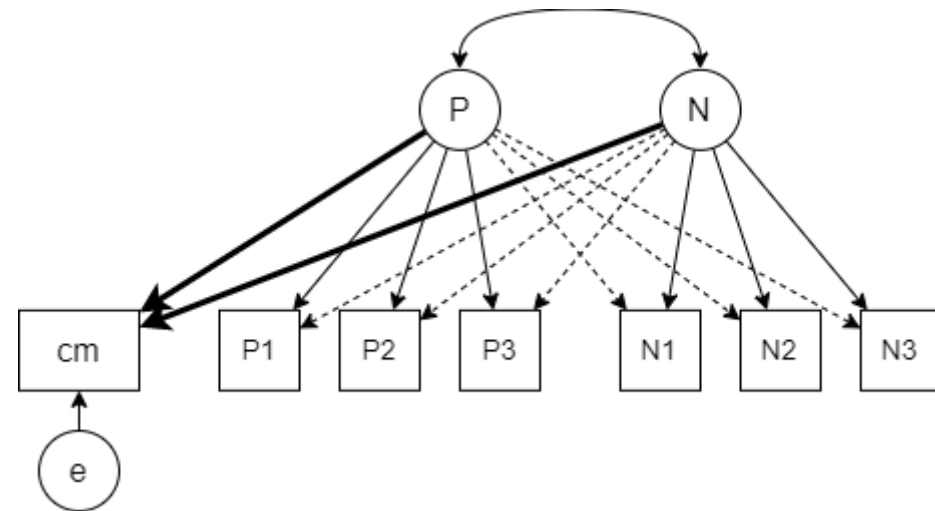
Sklon k nadhodnocování, resp. podhodnocování odpovědi.



Příklad: Měření výšky

M2: Skutečná výška je indikátorem obou faktorů (ženy $R^2=0,866$, $\beta_P=0,49$, $\beta_N=0,51$).

- Faktory P a N nyní reprezentují „skutečný atribut výšky“ ve smyslu fyzikální veličiny, nikoli však přesně.



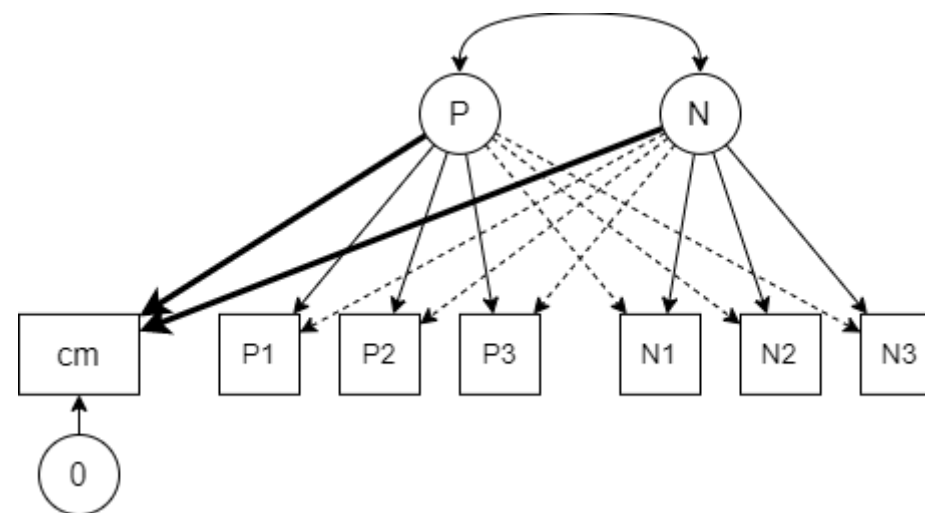
Příklad: Měření výšky

M2: Skutečná výška je indikátorem obou faktorů (ženy $R^2=0,866$, $\beta_P=0,49$, $\beta_N=0,51$).

- Faktory P a N nyní reprezentují „skutečný atribut výšky“ ve smyslu fyzikální veličiny, nikoli však přesně.

M3: Skutečná výška je bezchybným ($R^2=1$) indikátorem obou faktorů ($\beta_P=0,53$, $\beta_N=0,55$).

- Faktory P a N vysvětlí 100 % rozptylu výšky v cm.
- Faktory jsou ztotožněny s fyzikální výškou.



Příklad: Měření výšky

M2: Skutečná výška je indikátorem obou faktorů (ženy $R^2=0,866$, $\beta_P=0,49$, $\beta_N=0,51$).

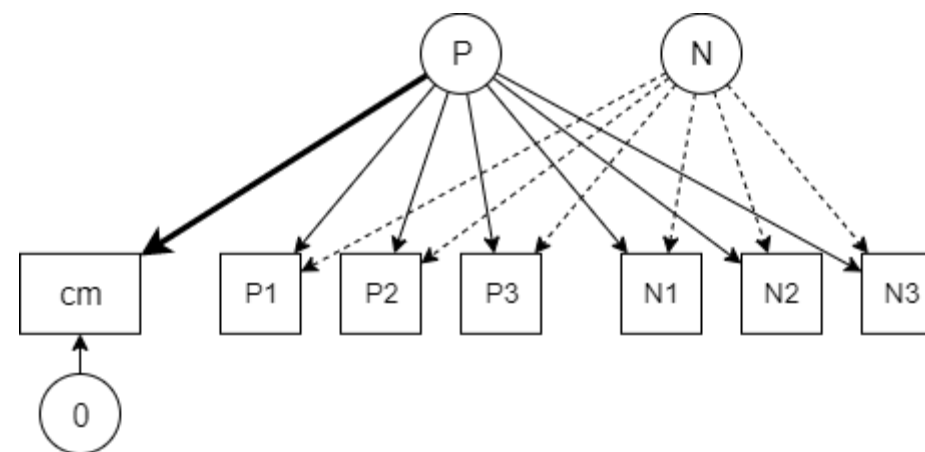
- Faktory P a N nyní reprezentují „skutečný atribut výšky“ ve smyslu fyzikální veličiny, nikoli však přesně.

M3: Skutečná výška je bezchybným ($R^2=1$) indikátorem obou faktorů ($\beta_P=0,53$, $\beta_N=0,55$).

- Faktory P a N vysvětlí 100 % rozptylu výšky v cm.
- Faktory jsou ztotožněny s fyzikální výškou.

M4: Pouze jeden faktor je ztotožněn s fyzikální výškou ($\beta_1=1$, $\beta_2=0$).

- Druhý, ortogonální faktor reprezentuje něco jiného (rozdíly v odpovídání).



Příklad: Měření výšky

M2: Skutečná výška je indikátorem obou faktorů (ženy $R^2=0,866$, $\beta_P=0,49$, $\beta_N=0,51$).

- Faktory P a N nyní reprezentují „skutečný atribut výšky“ ve smyslu fyzikální veličiny, nikoli však přesně.

M3: Skutečná výška je bezchybným ($R^2=1$) indikátorem obou faktorů ($\beta_P=0,53$, $\beta_N=0,55$).

- Faktory P a N vysvětlí 100 % rozptylu výšky v cm.
- Faktory jsou ztotožněny s fyzikální výškou.

M4: Pouze jeden faktor je ztotožněn s fyzikální výškou ($\beta_1=1$, $\beta_2=0$).

- Druhý, ortogonální faktor reprezentuje něco jiného (rozdíly v odpovídání).

	x2	df	CFI	TLI	RMSEA	SRMR
M1	1614,4	452	0,996	0,994	0,034	0,014
M2	2079,9	496	0,994	0,993	0,038	0,016
M3	2234,2	498	0,994	0,992	0,039	0,017
M4	2234,2	498	0,994	0,992	0,039	0,017

Reference

- Hayes, M. H. S., & Patterson, D. G. (1921). Experimental development of the graphic rating method. *Psychological Bulletin*, 18, 98–99.
- Bogardus, E.S. (1926). Social Distance in the City. *Proceedings and Publications of the American Sociological Society*, 20, 40-46.
- Thurstone, L. L. (1928). Attitudes Can Be Measured. *American Journal of Sociology*, 33(4), 529–554.
- Thurstone, L. L. (1928). Attitudes Can Be Measured. *American Journal of Sociology*, 33(4), 529–554.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. (1957). *The measurement of meaning*. Univer. Illinois Press.
- Ferguson, A., Myers, C. S., Bartlett, R. J., Banister, H., Bartlett, F. C., Brown, W., ..., & Tucker, W. S. (1940). Final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *Report of the British Association for the Advancement of Science*, 2, 331–349.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 5–55.
- Stevens, S. S. (1936). A scale for the measurement of a psychological magnitude: loudness. *Psychological Review*, 43(5), 405–416.
- Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684), 677–680.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302.
- Lilienfeld, S. O., & Strother, A. N. (2020). Psychological measurement and the replication crisis: Four sacred cows. *Canadian Psychology*, 61(4), 281–288.
- Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, 18(2), 5–11.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Fried, E.I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197.
- Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge University Press.
- Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88(3), 355–383.
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2), 155–174.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1), 19–30.
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of mathematical psychology*, 1(1), 1-27.
- Rečka, K. (2018). *Dotazník výšky a váhy* [Diplomová práce, Masarykova univerzita]. Institucionální repozitář MU.
- Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1), 72–101.

Děkuji za pozornost!

Hynek Cígler

Katedra psychologie a Institut pro psychologický výzkum (INPSY)
Fakulta sociálních studií, Masarykova univerzita

web: <https://inpsy.fss.muni.cz/>

e-mail: cigler@fss.muni.cz

OSF: <https://osf.io/t6ufg/>

Github: <https://github.com/hyneckcigler>

ORCID: <http://orcid.org/0000-0001-9959-6227>



<http://fssvm6.fss.muni.cz/vyska/>