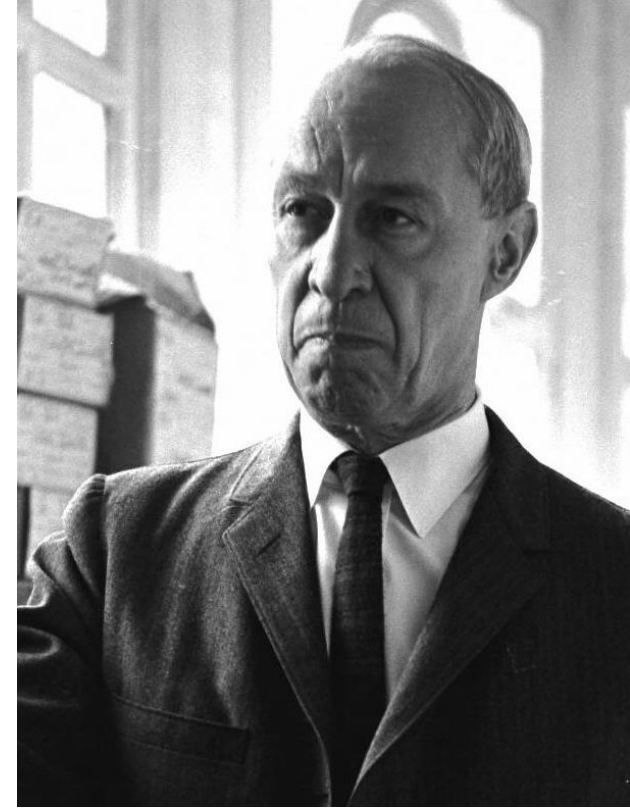


FIGURE 1  
DISTRIBUTIONS OF THE SCALE VALUES ON THE MINNESOTA SURVEY OF OPINIONS



# Škálování v psychologii

PSYb2590: Základy psychometriky | Přednáška 2

27. 2. 2024 | Hynek Cígler

# Měření v psychologii

---

Měření psychických rysů je výrazně náročnější než v přírodních vědách.

- 1. Psychické jevy **nelze přímo pozorovat**. Měření musí být nepřímé, odvozené.
- 2. Měření je zatíženo **extrémním množstvím** systematického i nesystematického **šumu**.
- 3. Veškeré informace o podstatě měřených jevech máme pouze z tohoto měření.  
Příslušná **teorie měření reifikuje** naše představy o měřeném atributu.
- 4. Neexistují žádné „**psychikální zákony**“, které by bylo možné bezchybně matematicky popsat.

V psychologii proto zcela oddělujeme **latentní proměnné...**

- Nepozorované, myšlené teoretické konstrukty, které mohou, ale nemusí mít reálnou existenci.

**... a manifestní proměnné:**

- Pozorované chování, které je chápáno jako projev těchto proměnných latentních.

Teorie měření: definice **vztahu** latentní proměnné s proměnnou manifestní a způsobu (a **přesnosti**) odhadu úrovně latentní proměnné z proměnné manifestní.

Různé teorie měření navíc implikují různé *kvality* či *významy* latentní proměnné.

# Povaha psychologického konstruktů

---

Epistemologická východiska.

- Antirealismus (zejm. operacionalismus) vs. realismus.
- A s tím související otázka existence té konkrétní proměnné.

Reflektivní vs. formativní měření.

- Kauzální směr vztahu latentní a manifestní proměnné.

Úroveň latentní proměnné.

- intervalová vs. ordinální vs. nominální

Povaha intervalové proměnné

- spojitá vs. diskrétní

# Teorie měření v psychologii

---

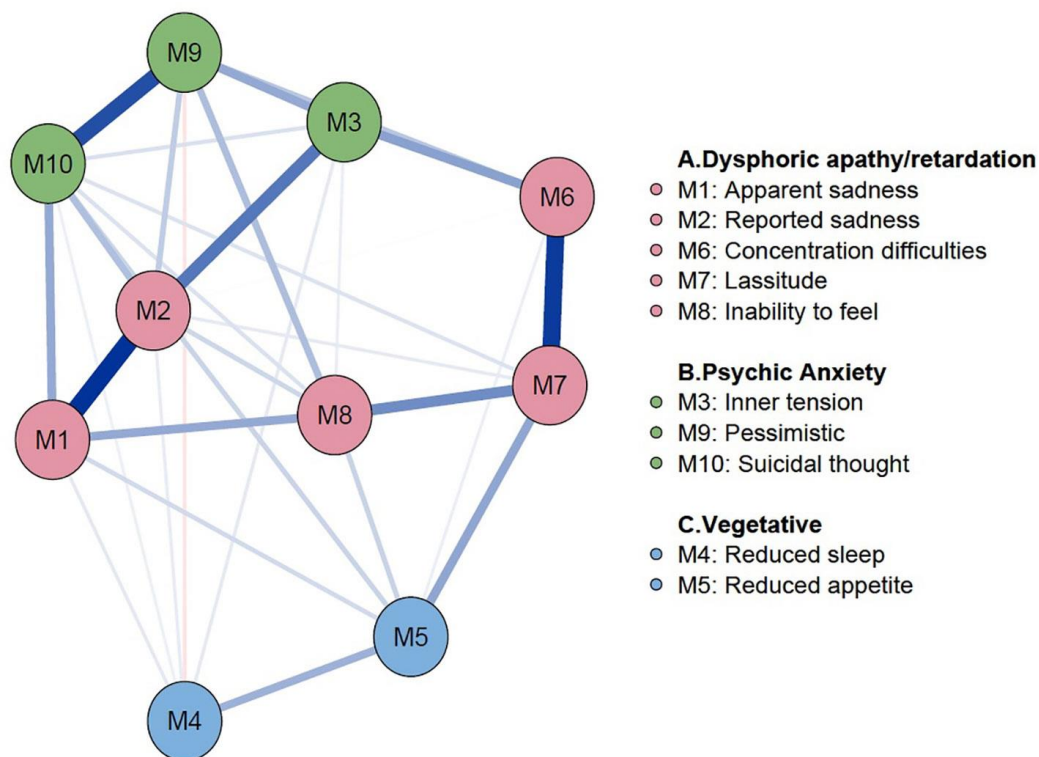
Teorie měření: rozdělení podle Borsbooma (2005):

- **Klasická testová teorie (CTT).**
- **Modely s latentními rysy (latent trait theories).**
  - Faktorová analýza (FA), teorie odpovědi na položku (IRT).
- **Fundamentální/reprezentační model měření (spojité měření, CM).**

Tyto teorie se ale překrývají.

- Např. Raschův model může být chápán jako fundamentální měření.
- Analogicky faktorová analýza je v praxi těsně propojená s CTT.

# Existuje vůbec latentní rys?



Latentní rys vůbec nemusí existovat.

To, co se nám „zdá být“ kvantitativní veličinou, může mít zcela jinou „povahu“.

Příklad: depresivita.

- Žádný rys „deprese“ nemusí existovat, ve skutečnosti může jít jen o souhrn dílčích symptomů.
- Měření „reifikuje“ naše úvahy o konstraktu.
- An, M. H., Park, S. S., You, S. C., Park, R. W., Park, B., Woo, H. K., ... Son, S. J. (2019). Depressive symptom network associated with comorbid anxiety in late-life depression. *Frontiers in Psychiatry*, 10. <https://doi.org/10.3389/fpsy.2019.00856>

# Aplikace psychometrie

---

Kromě teoretických úvah o modelu měření se psychometrie zabývá **praktickými aspekty** tohoto měření.

- Při vývoji i užití diagnostického nástroje.

Psychometrie je tedy základem jakékoliv psychologické diagnostiky.

Cílem je **kontrola chyb**, které jsou s měřením spojené.

- Jak chyby minimalizovat?
- Jak je odhadnout a jak s nimi pracovat?
- Náhodná vs. systematická chyba.

# Odhad a kontrola chyb měření: Vlastnosti psychodiagnostických metod

---

Chyby vztažené k nástroji: **validita** a **reliabilita**

Chyby vztažené k teorii: **konstruktová validita**

- Cronbach a Meehl vs. Messick.

Chyby vztažené k měřené osobě: **férovost**

- Respektive psychometrické pojetí férovosti.

Chyby vztažené k pozorovateli: zejména **objektivita**.

- **Férovost** v psychodiagnostickém pojetí.

Chyby interpretační: **normalizace, standardizace**.

*„Kde se vzala ta čísla?“*

---

cit. dle Zbyněk Vybíral (megakonvice, někdy)



# Škálování

---

Škálování může mít několik významů, mj.:

- Asociace kvalitativního pozorování s kvantitativním atributem.
- Procedura pro přiřazování čísel psychologickým jevům.
- Označení praktického postupu při vyhodnocování dotazníku.

Konkrétní odpověďový formát (**stupnice**) vs.

způsob odvození celkového skóre (a toto skóre jako takové) z dílčích odpovědí (**škála**).

Otázka škálování je pro psychologické měření zcela klíčová.

- A v první polovině 20. století proto vzniklo velké množství škálovacích procedur.

# 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 vnitrosubjektové rozdíly
- lokálně homogenní konstrukty
- 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
- lokálně irelevantní/heterogenní konstrukty
- člověk je měřeným subjektem
- Galton, McKeen Cattell, Spearman...

# Mentální měření – psychometrika

---

Návaznost na Galtona a měření inteligence.

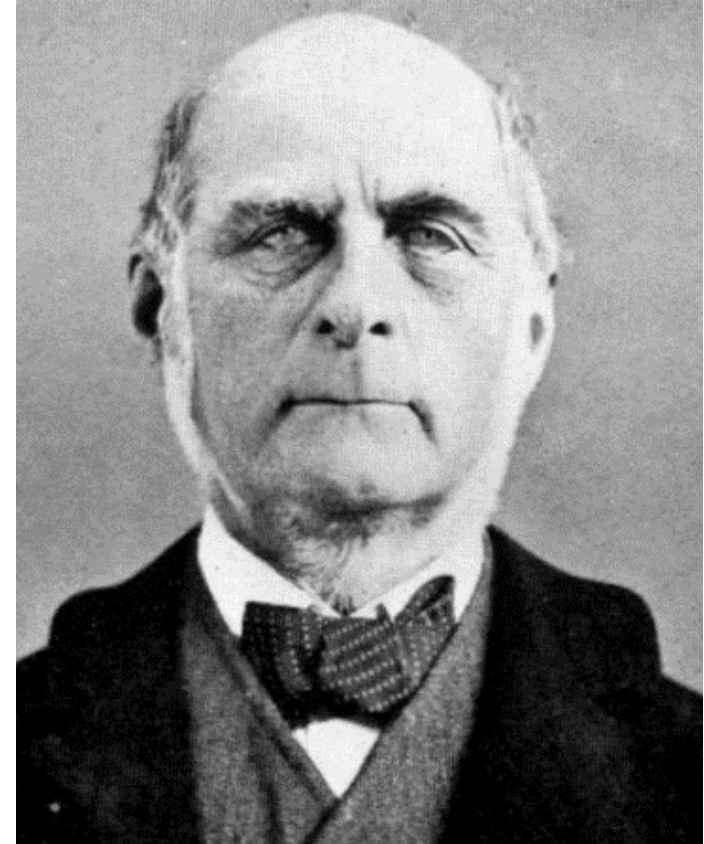
Základní předpoklad: existence mezilidských rozdílů s kvantitativní povahou.

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 sensorickými vjemy.

- Dnes je tradice patrná v „matematické psychologii“.

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

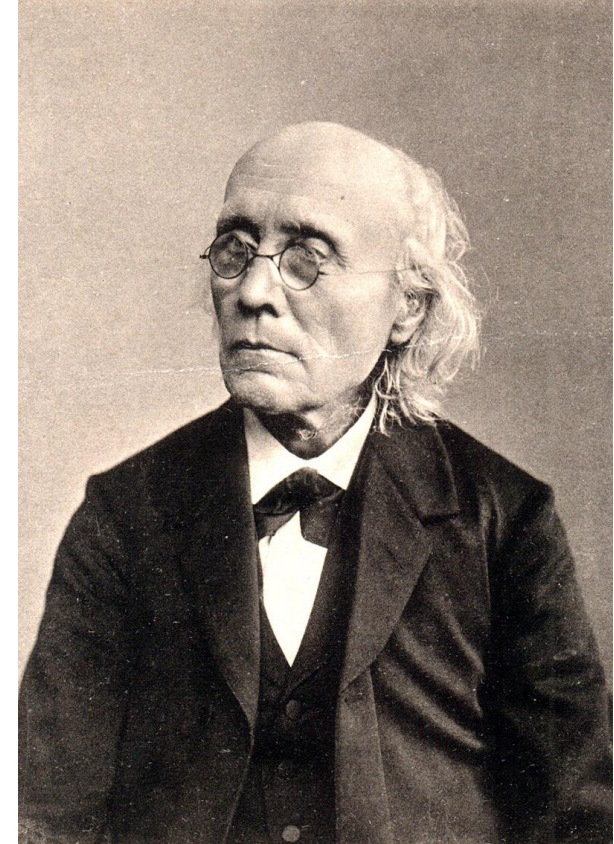
Např. Weberův a Fechnerův zákon.

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“.

Nalezení obecně platného zákona by bylo průlomové.

- Propojení fyzikálních atributů (tlak) a psychických jevů (bolest).



# 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; smyslové počítky...

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,  $\Delta t$ .

**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;  $\delta I$  – 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 (kalibrace).

Otázka: Kolik kalibrovaných bodů je nutné pro identifikaci koordinační funkce?



# Co je to měření: Final Report (Ferguson, 1940); 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. Číslo („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  $c$  je 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).

## 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.

# Škálovací postupy pro lokálně-homogenní atributy

---

# Nejmenší právě rozlišitelný rozdíl (JPD)

---

Weber a Fechner. Oba zákony formulované Fechnerem (1860).

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

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

Cílem bylo využít  $\delta I$  jako etalon pro měření intenzity smyslového vjemu.

- Propojila by fyzikální intenzitu a psychický jev.

Záhry se ale ukázalo, že  $\delta I$  záleží na míře intenzity podnětu. Weberův zákon:

$$\begin{aligned}\delta I &= kI \\ \Psi(I) &= k \cdot \ln I\end{aligned}$$

Lidé se ale liší v citlivosti na vjemy. Proto Fechnerův zákon, který zavádí referenční bod:

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

# Stevensův exponenciální zákon

Weber-Fechnerův zákon velmi špatně funguje pro velmi rozdílné podněty.

- Kde je pravděpodobnost odlišení velmi vysoká, resp. kde je vysoký rozdíl v intenzitě podnětu.
- Navíc u některých podnětů intenzita roste se silou podnětu (bolest).

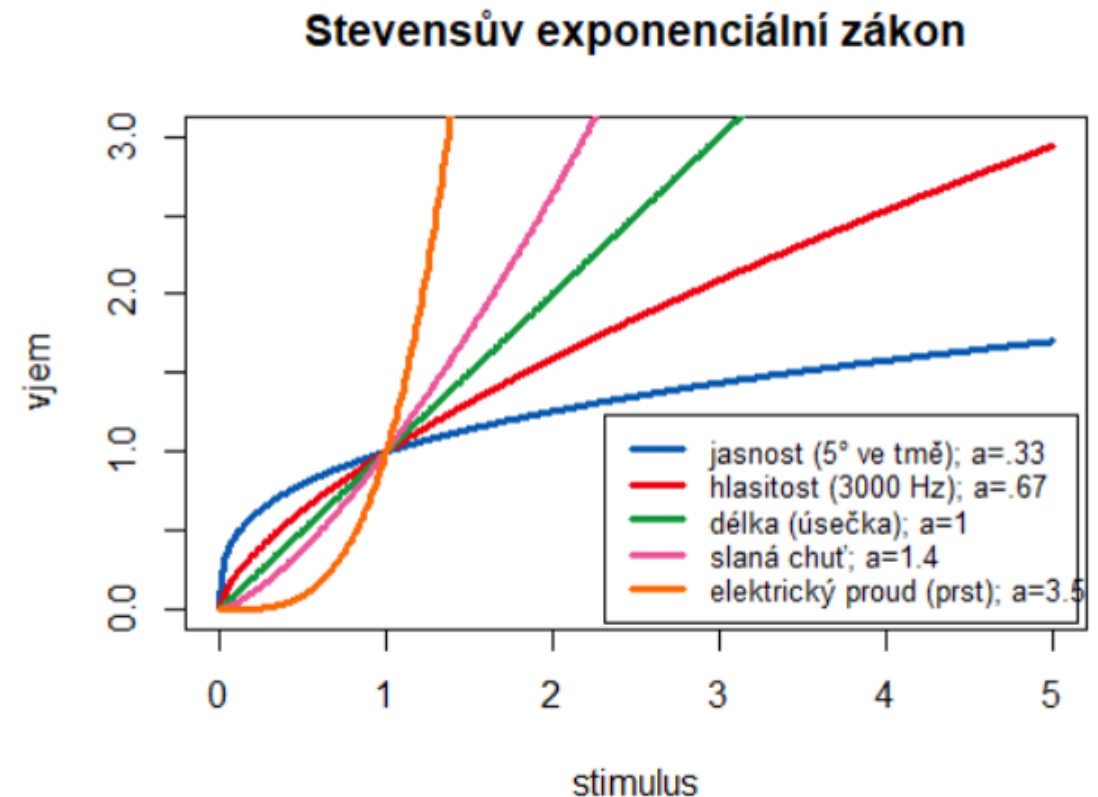
**Stevensův exponenciální zákon:**

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

- Funguje i pro větší rozdíly (s téměř jistou rozlišitelností vjemu).

Stejně jako Weber-Fechnerův zákon umožňuje modelovat rozdílný vztah pro různé proměnné

- Např. délka  $a = 1$ , hlasitost  $a = 0,67$ , bolest (způsobená elektřinou)  $a = 3,5$ .

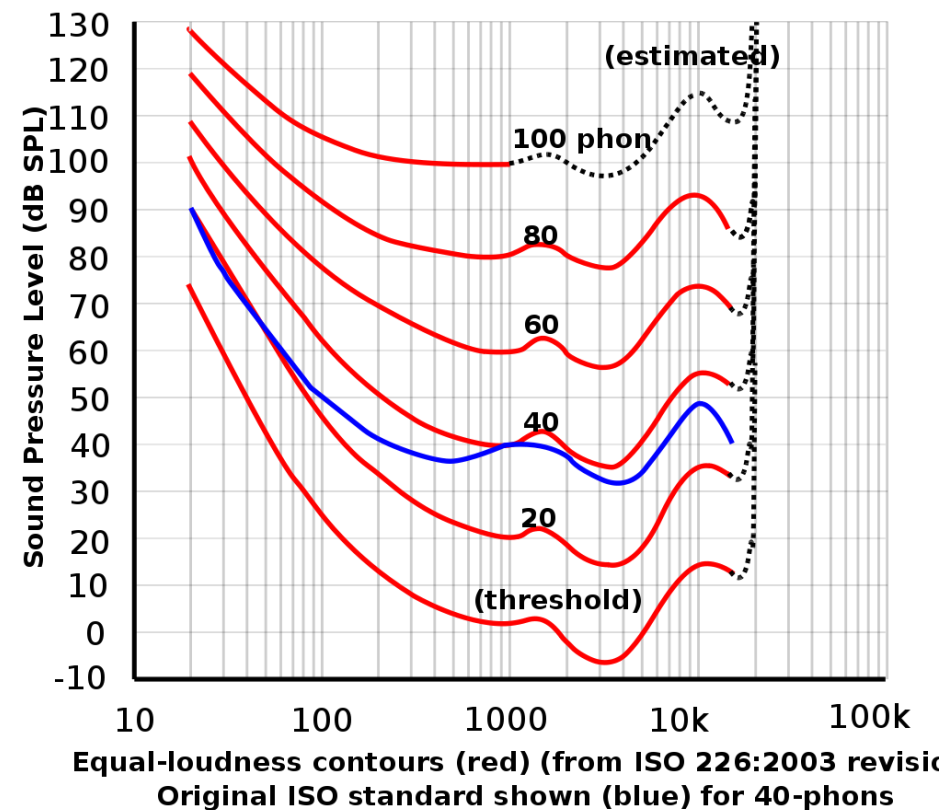


# Odbočka: Stanley S. Stevens

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>

# Metoda stejně se jevících intervalů (EAI)

---

Alternativní postup pro tvorbu koordinační funkce rozpracovaný Thurstonem.

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

Postup:

- Lidem prezentujeme tři podněty o síle  $p, q, r$ .
- Cílem je vybrat čtvrtý podnět  $s$  tak, aby intervaly byly shodné:

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

Těsně souvisí se známějším **Zákonem srovnávacího usuzování** (law of comparative judgment).

- Vzájemné srovnávání párů objektů s různou intenzitou.

# Nejmenší rozlišitelný rozdíl Metoda stejně se jevících intervalů

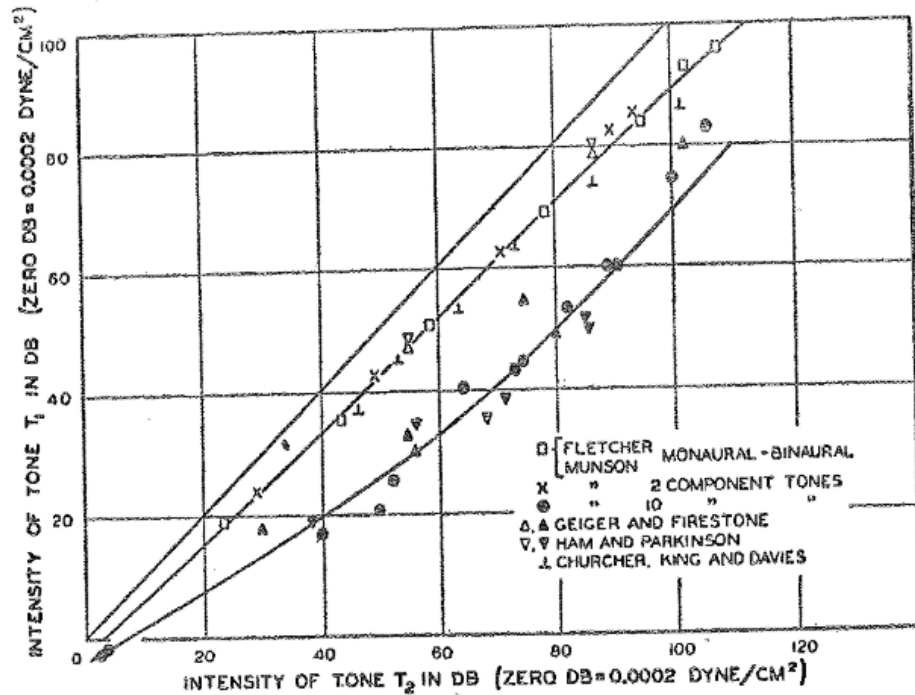


FIG. 41. The ordinate shows the intensity at which a tone ( $T_1$ ) sounds half as loud (open figures) or a tenth as loud (solid figures) as another tone ( $T_2$ ) whose intensity is given by the abscissa.

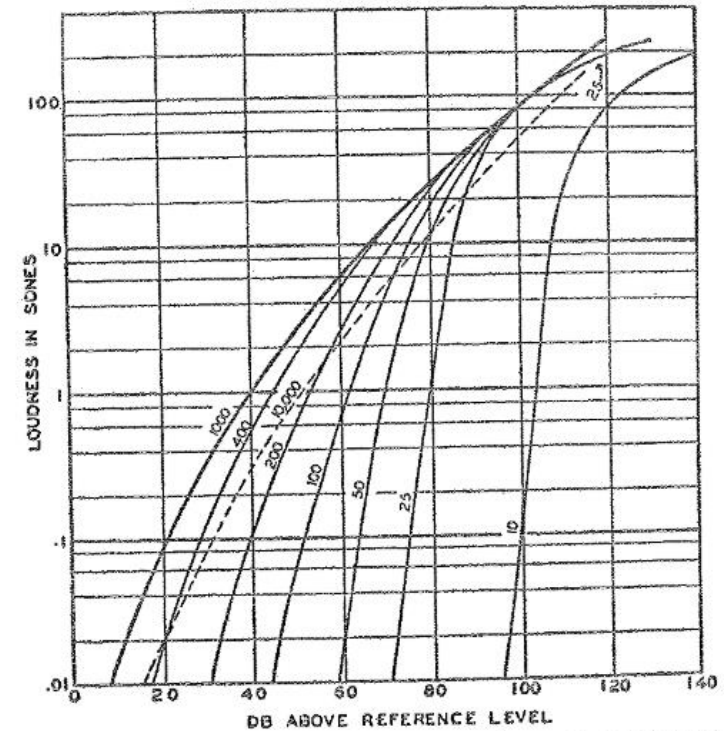


FIG. 43. The loudness-function. Showing how the perceived loudness of various tones depends upon the intensity of the stimulus. Frequency is the parameter.



# Škálovací postupy pro lokálně-irelevantní atributy

---

# Lokálně-irelevantní atributy

---

Lokálně-homogenní atributy mají vnitrosubjektovou variabilitu.

- Bolest může být velká nebo malá, a její síla se může lišit uvnitř i napříč osobami.

Lokálně-irelevantní atributy se vynořují až při srovnání osob.

- Inteligence či osobnostní rysy – jde o stabilní interindividuální charakteristiky.
- Stejně tak koncept „délky“ dává smysl až při srovnávání různě „dlouhých“ objektů.

Předchozí postupy proto nelze využít pro tyto atributy.

- Silou „vjemu“ (velikostí inteligence...) nejde variovat napříč experimenty, protože je stabilní.

Ve 20. století proto začala vznikat řada postupů pro škálování postojů, osobnostních charakteristik, inteligence a podobně.

# Vizuální analogová škála (VAŠ)

Hayes a Peterson (1921).

- Naivní předpoklad lineární funkce mezi VAŠ a silou postoje.
- Dneska se používá spíše zřídka (až na několik „standardizovaných“ použití).

Příklad škály:

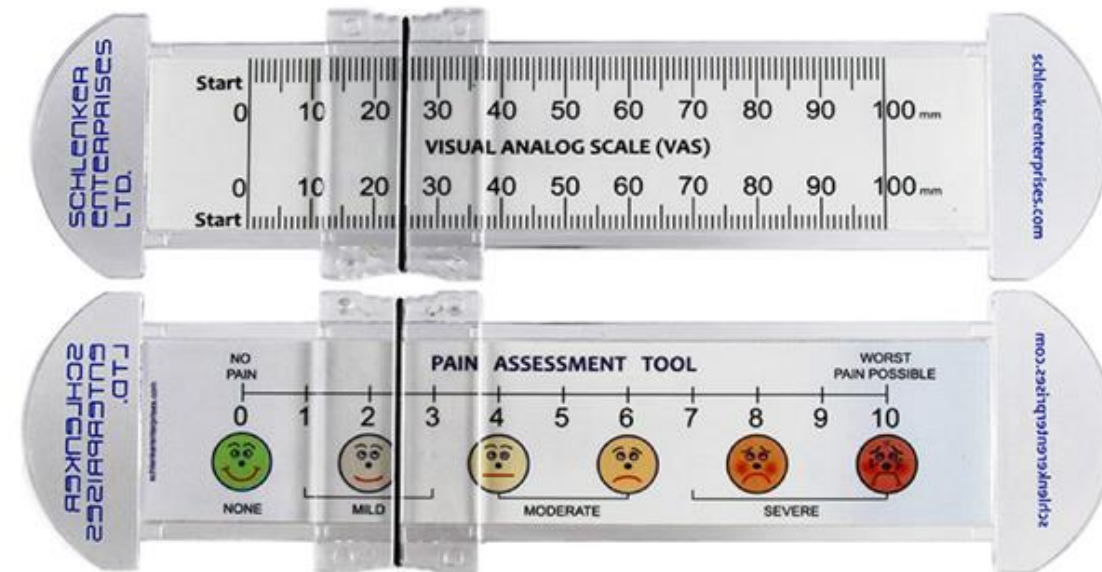
4. I worry about what will happen to me



Not worried



Very worried



# Bogardusova škála sociální distance

---

Bogardus (1926) sestavil sadu ordinálně seřazených výroků pro měření sociální distance.

Míra je reprezentována výrokem „nejsilnějším“ výrokem, se kterým respondent souhlasí.

Původní podoba škály: „Určitou osobu [...] byste ochotně přijal/a:

- do blízkého příbuzenstva skrze manželství (1)
- do svého klubu jako blízkého přítele (2)
- za souseda v ulici (3)
- svého zaměstnání za spolupracovníka (4)
- jako občana své země (5)
- jen jako návštěvníka své země (6)
- vyloučil by ji ze země (7)

# Thurstonovy škály

---

Thurstone (1928) vyvinul tři různé škálovací postupy (původně postoje k náboženství).

Založené na metodě stejně jevících se intervalů.

- Překonává potíž Bogardusovy škály, která je ordinální, a snaží se vytvořit intervalovou stupnici.

Každé tvrzení je ohodnoceno větším množstvím posuzovatelů. Tvrzením jsou pak přiřazeny číselné hodnoty reprezentujícími jejich průměrné hodnocení.

- Položky mohly být binární i ordinální.

Tyto čísla jsou použity při sčítání jednotlivých dotazníkových položek.

Položky pro finální respondenty vypadají stejně jako likertovské položky, ale liší se postup statistického zpracování a „škálování“.

Postup byl ale drahý (původní doporučení – ideálně přes 200 posuzovatelů).

# Likertova škála

---

Likert (1932): dva podobné škálovací postupy.

- **Metoda sigma:** Na rozdíl od Thurstonovy škály jsou číselné hodnoty derivovány přímo z finálního vzorku za předpokladu normálního rozložení.
- **Zjednodušená metoda:** Číselné hodnoty jednotlivých položek jsou přiřazeny arbitrárně.

Výsledky obou postupů spolu silně korelovaly ( $r > 0,98$ ), a proto Likert doporučil používat zjednodušenou metodu.

- A korelace rostla spolu s počtem položek.

Likertova škála následně byla převzata jako standard pro měření postujů a osobnostních charakteristik.

# Sémantický diferenciál a další rozšíření Likertovy škály

---

Osgood (1957) navrhl namísto míry souhlasu pojmenovat oba krajní póly Likertovy škály antonymními výroky.

Průběžně dochází k dalším dílčím úpravám Likertovy škály.

- Např. namísto míry souhlasu označují jednotlivé verbální kotvy časovou frekvenci a podobně.

Je proto nutné odlišit Likertovu škálu jako:

- formát předložení jedné položky („likertovská stupnice“, míra souhlasu);
- způsob tvorby celkového skóre (arbitrární očíslování odpovědí a jejich součet).

# Guttmanova škála

---

Guttman ve 40. létech rozpracoval Bogardusovu škálu pro výkonové testy.

Test sestává z úkolů se vzrůstající obtížností.

Sledujeme, „jak daleko“ se respondent dostane.

Každá položka je skórována správně (= 1) a chybně (= 0).

Celkové skóre je součtem bodů (počet správně odpovězených položek).

Škála je potom definovaná právě jednotlivými úkoly.

Guttmanův deterministický model byl jedním z hlavních zdrojů pozdější stochastické teorie odpovědi na položku.



# Závěrem: Jiné teorie měření

---

Závěr: Prostý součet arbitrárně skórovaných položek je uspokojivý.

- Má přijatelné psychometrické vlastnosti.
- Výkonové testy (Guttmanova škála) i „škálovací postupy“ (Likertova škála).

Ale pozor! Všechny uvedené škálovací postupy jsou těsně spojeny s CTT.

- Jiné teorie měření (reprezentační model, teorie odpovědi na položku) využívají stejné formáty položek, ale číselnou strukturu z nich odvozují jiným způsobem.
- CTT je ale teorií „celkového skóre“. Jen někdy je důležité, jak toto skóre vzniklo.

# Položky

---

# Položka

---

Nejzákladnější „skórovaná“ jednotka měřicího nástroje (dotazníku, testu...).

Struktura:

- „Item stem“ – kořen/kmen položky. Stimulus.
- „Response options“, odpověďová stupnice – formát toho, jak respondent odpovídá.
  - Distraktory – nabídka chybných odpovědí u multiple-choice testů.
- „Response“ – pozorovaná odpověď respondenta.
- „Scoring function“ – způsob skórování odpovědi (vztah odpovědi a měřeného atributu).

# Typy položek a testových metod

---

Z hlediska úrovně měření:

- Nominální, ordinální, intervalové.
- (Nemusí souviset s úrovní měření škály).

Z hlediska formátu odpovědi:

- míra souhlasu (likertovská stupnice).
- otevřená odpověď, doplňování odpovědi
- vícenásobná volba (multiple-choice)
- nucená volba
- dichotomická položka
- spojování, párování...
- řazení
- párové srovnání
- a mnoho dalších

Typy psychodiagnostických metod:

- Testy schopností
- Osobnostní dotazníky
- silové, rychlostí vs. kapacitní testy
- výkonové vs. introspektivní metody
- „objektivní“ testy
- projektivní testy

V principu: existuje „správná“ či „diagnostická“ odpověď?

# Závěrem

---

# Dotazník výšky

---

1. Mám vhodnou výšku na hraní basketbalu nebo volejbalu.
2. Když mluvím s jinými dospělými a chci se jim dívat do oči, častěji na ně spíš vzhlížím nahoru.
3. Lidem, kteří na koncertě stojí za mnou, většinou má postava dost brání ve výhledu.
4. Často musím stát na špičkách, abych lépe viděl/a.
5. Když chci někoho obejmout, většinou se musím sklonit.
6. Často potřebuji stoličku, abych dosáhl/a na něco, na co jiní lidé dosáhnou normálně.
7. Jednou z prvních věcí, které si na mně lidé všimnou, je to, jak moc jsem malý/á.
8. Často si musím dávat pozor, abych se neuhodil/a hlavou např. o nízký strop nebo rám dveří.
9. V autobuse mívám dostatek prostoru pro nohy.
10. Slýchávám narážky na to, že jsem vysoký/á.
11. Kvůli mé menší výšce lidé hádají, že jsem mladší, než ve skutečnosti jsem.

Škála je skórovaná  
ano=1, ne=0 (po otočení  
reverzních položek).

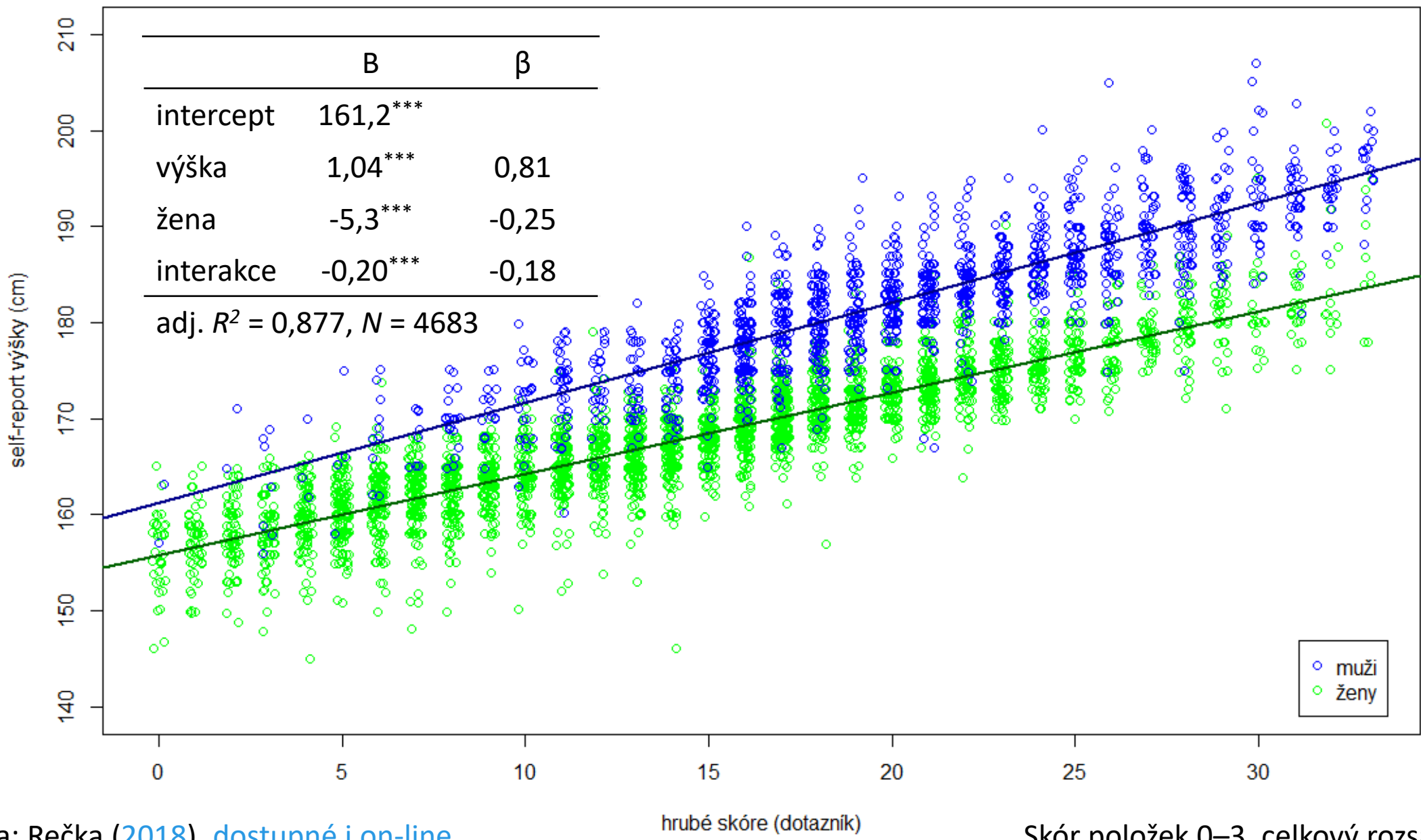
Použitou škálou je  
**počet souhlasných  
odpovědí** (rozmezí 0–11).

Je měření přímé či nepřímé?

- Proč ano?
- Proč ne?

Je výsledná škála aditivní?

- Proč ano?
- Proč ne?



Data: Rečka (2018), [dostupné i on-line](#).

Skór položek 0–3, celkový rozsah 0–33 b.

# Intermezzo: Dotazník výšky

---

Měřit výšku dotazníkem je samozřejmě nesmysl:

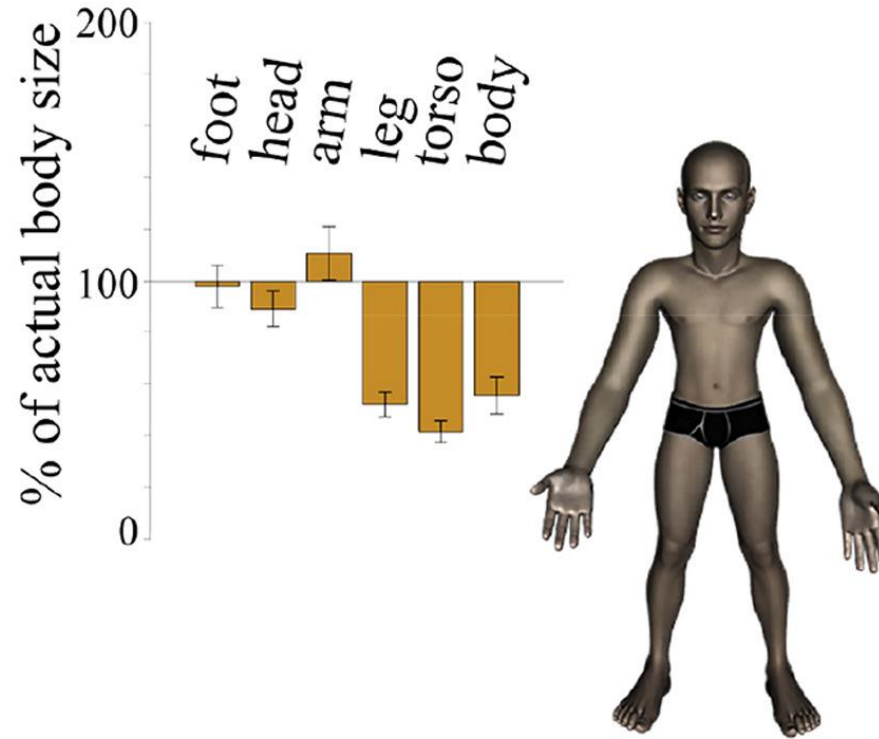
1. Máme lepší nástroje měření.
2. Výšku přece vidíme, není to žádná latentní proměnná?



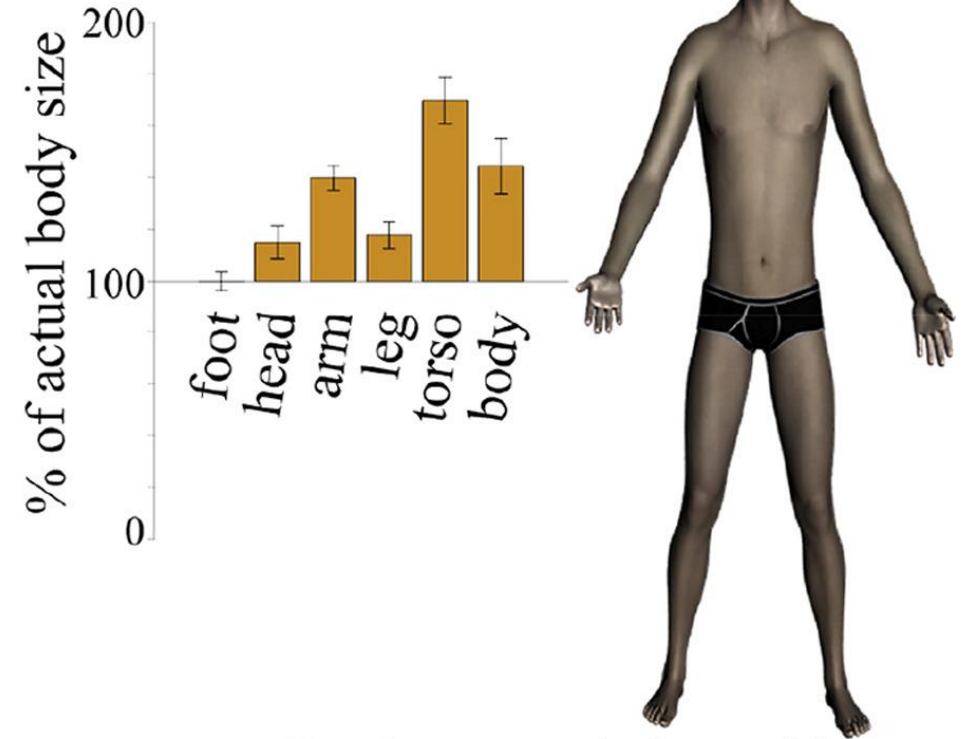
## actual body



## volume perception



## length perception



Error bars are standard error of the mean

*„Participants [...] estimated how many units would fit in a perceived size of body segments and the whole body. [...] The body parts with a smaller actual surface area relative to their volume were underestimated more. There was a tendency for body parts underestimated in volume to be overestimated in length.“*

Sadibola, R., Ferrè, E. R., Linkenauger, S. A., & Longo, M. R. (2019). Distortions of perceived volume and length of body parts. *Cortex* 111, 74–86. doi: [10.1016/j.cortex.2018.10.016](https://doi.org/10.1016/j.cortex.2018.10.016)