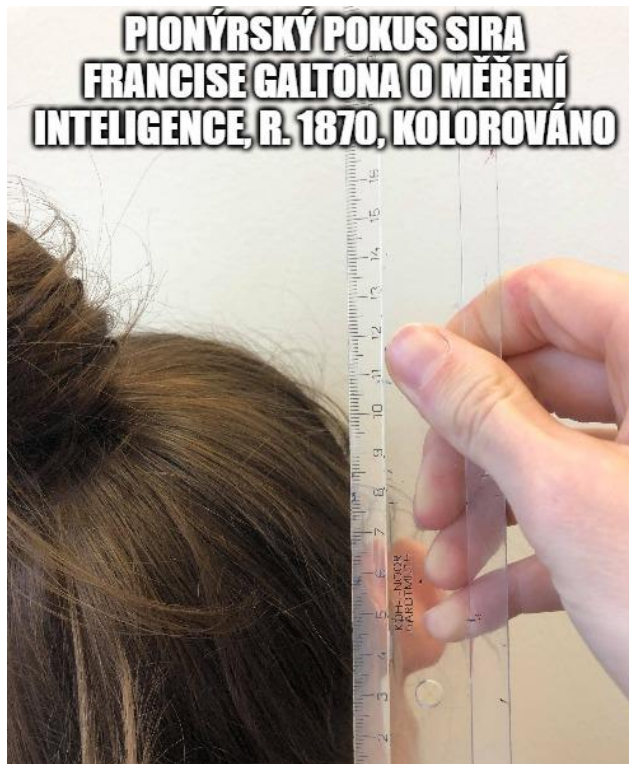


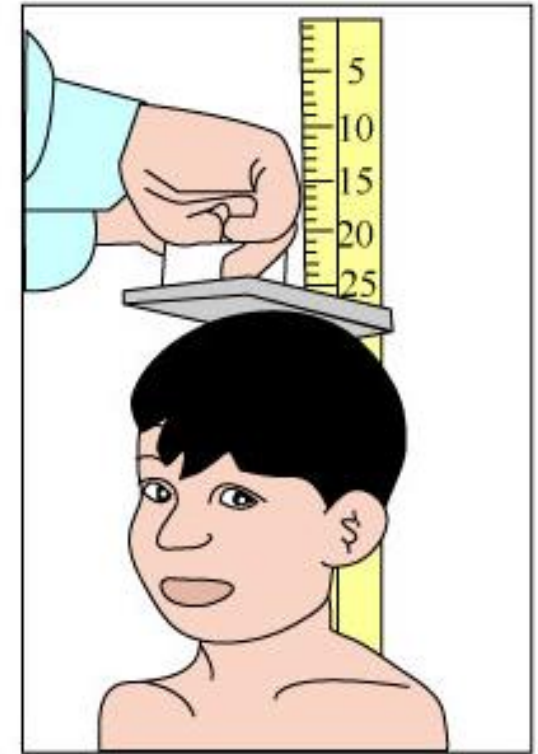
Úvod do psychometriky

PSYb2590: Základy psychometriky | Přednáška 1

24. 2. 2020 | Hynek Cígler



Zdroj: <https://www.facebook.com/vimzenicnezmerim>



fssvm6.fss.muni.cz/vyska

PŘÍPADNĚ ODKAZ VE STUDIJNÍCH MATERIÁLECH.

*„Psychometrics is a scientific discipline concerned with the **construction of assessment tools**, measurement instruments, and **formalized models** that may serve to **connect observable phenomena** (e.g., responses to items in an IQ-test) **to theoretical attributes** (e.g., intelligence). [...] Such models present **conceptual, substantive, and statistical problems** that psychometricians aim to analyze and solve. Because many of the questions that psychometricians study transcend disciplinary boundaries [...] the **boundaries of the discipline are fuzzy**; psychometrics is especially closely intertwined with methodology and statistics. Psychometric techniques are **widely used across the sciences**, and have found applications in educational testing, behavior genetics, sociology, political science, and neuroscience.“*

Denny Borsboom (University of Amsterdam)

<https://www.psychometricsociety.org/content/what-psychometrics>

*„Psychometrics is the **approximation of latent psychological processes** by means of **stochastic analysis** at both the individual and population levels.“*

Peter Molenaar (Pennsylvania State University)

<https://www.psychometricsociety.org/content/what-psychometrics>

Psychometrika

Věda o měření psychických jevů (nejen v psychologii).

- Pomocná psychologická disciplína – tvorba a hodnocení diagnostických a výzkumných metod.
- Teoretický obor i praktické „řemeslo“.

Teorie měření v psychologii a příbuzných oborech.

- Čerpá zejména z metodologie a matematické statistiky.
- Úzce provázaná s psychologií osobnosti, individuálních rozdílů a dalších.

Způsob měření často souvisí s tím, jak vnímáme některé psychologické konstrukty, či přímo náš pohled přímo utváří.

- Osobnost: Eysenck; Cattell & McCrae... Inteligence: Thurstone, Spearman, Cattell-Horn-Carroll...
- Např. extraverte jako osobnostní „**typ**“ (Jung; MBTI^a) vs. kontinuální „**míra**“ (Eysenck, NEO).

^a Tímto srovnáním v žádném případě neuvádím Junga do falešné souvislosti s pseudovědeckými diagnostickými metodami.

Proč vás psychometrika zajímá?

Využití psychometriky je všude kolem nás: v průběžných testech z psychometriky, státní maturitě, přijímačkách na VŠ, kariérním poradenství...

Psychometrické znalosti jsou nezbytné pro praktické provádění psychologické diagnostiky (zejména v klinických a poradenských oborech včetně HR).

Pochopení měření v psychologii umožňuje lépe uvažovat o psychických jevech.

Měření je základem vědy. Poznatky, se kterými se setkáváte, byly „změřeny“.

- A validita těchto výzkumů je jen tak dobrá, jak dobré bylo použité měření.

Psychometrika je nezbytná, pokud zvažujete budoucnost ve vědě (i aplikované).

- Hlubší psychometrické dovednosti jsou zárukou velmi dobré pracovní pozice (v zahraničí).

Psychometrika je vzrušující! 😊 (V současnosti extrémně dynamický rozvoj.)

Cíle kurzu PSYb2590: Základy psychometrie

- „Psychometrický“ **způsob uvažování** o psychologických konstruktech.
- Základní pojmy a koncepty** měření v psychologii (a sociálních vědách).
- Předpoklady pro budoucí psychodiagnostickou praxi.**
- „**Řemeslo**“: postupy vývoje psychodiagnostických metod.
- Statistické postupy** nezbytné pro analýzu testu.

Co se v PSYb2590 nenaučíte?

Základy metodologie a statistiky.

- PSY117, PSY252 (Statistická analýza I a II)
- PSY112 (Metodologie psychologického výzkumu).

Konkrétně:

- Sampling, formulace a testování hypotéz, úrovně měření.
- Rozdělení proměnné (střední hodnoty a rozptyl).
- Kovariance a korelace, testy středních hodnot, kontingenční tabulky a testy.
- Lineární, logistická regrese (základní pojmy).

Doporučujeme **současný zápis PSYb2550** (Individuální rozdíly).

Organizace kurzu

Přednášky (nástin, přehled)
+ **semináře** (nácvik dovedností)
+ **literatura** (znalosti).

Semináře:

- Seminární skupina 01: Hynek Cígler (sudé týdny, po přednášce).
- Seminární skupina 02: Standa Ježek (liché týdny v 10:00).
- Seminární skupina 03: Adam Ťápal (liché týdny ve 12:00).

Konzultační hodiny (zřejmě bude ještě upřesněno):

- HC, SJ středa odpoledne, kanc. 2.47. AŽ po domluvě tamtéž.
- Primární kontakt: HC. E-mailem jen stručně, raději osobně.

Orientační harmonogram

Pořadí	Datum	Téma
přednáška 1	24. 2.	Úvodní informace o kurzu. Úvod do psychometriky a měření v psychologii.
seminář 1	24. 3. / 2. 4.	Zadání seminární práce, vývoj měřicích nástrojů.
přednáška 2	9. 3.	Model měření CTT. Reliabilita . Úvod do chyby měření v rámci CTT.
seminář 2	9./16. 3.	Práce s chybou měření.
přednáška 3	23. 3.	Validita . Nové i tradiční pojetí validity.
seminář 3	23./30. 3.	Položková analýza a analýza testu v rámci CTT.
přednáška 4	6. 4.	Faktorová analýza – teorie.
seminář 4	6. 4.	Explorační faktorová analýza v praxi. Společný seminář (aula)!
přednáška 5	20. 4.	Normy a standardizace testu.
seminář 5	20./27. 4.	Konfirmační faktorová analýza v praxi.
přednáška 6	4. 5.	Úvod do dalších teorií měření .
seminář 6	4./11. 5.	Psychometrika v praxi (bude upřesněno).

Studijní zdroje

FURR: Primární učebnice pro kurz.

- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics : An Introduction*. Los Angeles, CA: Sage.

HOG: Vybrané kapitoly, více psychodiagnostické (a využijete na mgr) – lze kombinovat s FURR.

- Hogan, T. P. (2013). *Psychological Testing: A Practical Introduction*. New York: Wiley.

URB: Doplnující zdroj v češtině (ale pozor, některá sub témata chybějí).

- Urbánek, T., Denglerová, D., Širůček, J. (2011). *Psychometrika: Měření v psychologii*. Praha: Portál.

SCHULTZ: Praktické, dobrý zdroj k seminářům.

- Schultz, K. S., Whitney, D. J., & Zickar, M. J. (2014). *Measurement Theory in Action: Case Studies and Exercises (2 ed.)*. London: Routledge.

Další: U každého tématu jsou specificky uvedené zdroje včetně rozsahu.

Specifikace konkrétní literatury k jednotlivým tématům se může v průběhu semestru měnit!

FB group: [Statistika, metodologie, psychometrika](#). Prostor pro dotazy.

Požadavky k ukončení kurzu

6 průběžných testů (max. 25 bodů).

- Minimálně 13 bodů.

Skupinová seminární práce (max. 25 bodů).

- Spodní limit není, nelze přepracovat.

Písemná zkouška (max. 30 bodů).

- Minimálně 15 bodů.

Ústní zkouška (max. 20 bodů).

- 0-20 bodů.
- Nebo „nedostatečně“ (= F).

celkem bodů	hodnocení
100–91	A
90–81	B
80–71	C
70–61	D
60–51	E
50 a méně	F

Požadavky: Průběžné testy

Celkem 6 průběžných testů, každý max. 5 bodů.

- Nejhorší výsledek nebude započítán, celkem max. $(6-1) \times 5 = 25$ bodů.
- 12 bodů a méně → hodnocení „X“, nelze opravit.

Otevřené položky z povinné literatury.

- Podrobné zadání nejpozději v den po přednášce k danému tématu.

Open-book, z domova, s použitím libovolných zdrojů.

- Lze zodpovědět kdykoli v týdnu před konáním další přednášky.
- Max. jedno spuštění odpovědníku, zpravidla max. 30 minut.

Požadavky: Průběžné testy

Celkem 6 průběžných

- Nejhorší výsledek ne
- 12 bodů a méně → h

Otevřené položky z

- Podrobné zadání nej

Open-book, z domo

- Lze zodpovědět kdyk
- Max. jedno spuštění

Spolupráce?



Požadavky: Zkouška

Písemná zkouška: maximálně 30 bodů.

- Alespoň 15 bodů pro přístup k ústní zkoušce.
- 14 a méně bodů = F.
- Open book, libovolné zdroje, vlastní notebook nebo školní PC.
- Zákaz e-mailu, messengeru apod.

Ústní zkouška: 0–20 bodů nebo „F“.

- Znalosti, schopnost argumentace.

Maximálně 3 (2) písemné termíny, max. 2 ústní.

- Při nedosažení 51 bodů lze opravit ústní a/nebo písemnou zkoušku.

Statistika úspěšnosti

JS 2016:

Předmět	Celkem studentů	Úspěšně	Průměr	A	B	C	D	E	FD	FFD	FFE	F	-	FF	FFF	--
PSY259	57	85 %	2.87	1	7	16	7	9	5	1	2	1	1	4	2	1

JS 2017:

Předmět	Celkem studentů	Úspěšně	Průměr	B	C	D	E	FC	FD	FFD	FE	FFE	-	FF	FFF	FF-
PSY259	59	80 %	2.79	5	16	14	5	1	1	1	3	1	8	1	2	1

JS2018:

Předmět	Celkem studentů	Úspěšně	Průměr	A	B	C	D	E	FC	-D	FD	FFD	FE	FFE	X	-	-FF
PSY259	69	91 %	2.65	4	5	15	20	6	1	1	6	2	1	1	4	1	2

JS2019:

Předmět	Celkem studentů	Úspěšně	Průměr	A	B	C	D	E	FC	FD	FFD	FE	FFE	X	-
PSY259	60	98 %	2.72	5	9	6	14	9	1	4	1	6	3	1	1

Pozorované příčiny neúspěchu

Studium pouček bez jejich pochopení.

- Nezkoušíme dílčí znalosti, zajímá nás, jak tomu rozumíte.

Podcenění průběžných písemek.

- Nejde ani tolik o body, jako o to, že na konci se to naráz nenaučíte.
 - $r_{2019} = 0,73, r_{2018} = 0,77, r_{2017} = 0,69, r_{2016} = 0,71$).

Rezignace na matematiku, přeskokování vzorců při studiu.

- Bez matematiky to nejde; není v tom ale nic složitého. Důležité je konceptuální porozumění vzorci.

Studium psychometrie zvlášť oproti zbytku psychologie.

- Je potřeba si představovat měření ve vztahu ke konkrétním proměnným: „Jak změřím tohle?“

„Tohle tam nebude.“

- Bude.



Základy psychometriky a trocha historie

Zdroje k historii: zejm. Hogan (2013) a Buchanan ([2005](#))

Historie „mentálního měření“

Počátky měření psychických vlastností už v předvědeckých dobách.

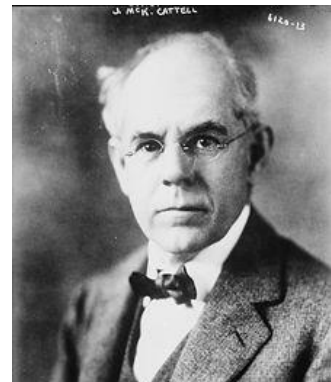
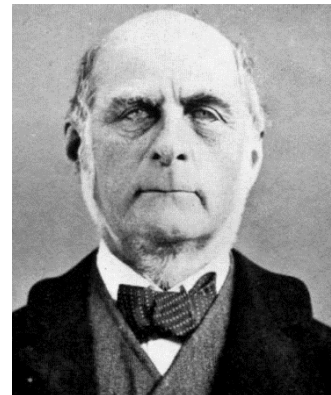
- Čínští úředníci cca 2200 př.n.l., od 13. století první ústní zkoušky na evropských univerzitách.
- 16. století: první písemné testy (Jezuité).

Základy moderní psychometrie: **Francis Galton (1822–1911)**.

- Navázal na předchozí informace o rozložení lidských fyzických charakteristik (Quetelet) či systematické interindividuální rozdíly v astronomických pozorováních (Bessel, 1816).
- V 80. letech 19. stol.: řada testů fyzických a psychických charakteristik (síla, paměť).

James McK. Cattell (1860–1944): Zkušenost z Wundtových laboratoří.

- Zaměřil se na *rozdíly*, nikoliv *obecné charakteristiky* (rozdíly → předpoklad měření).
- Mentální test, 1890. Do konce 19. století vytvořil velké množství „testů“ (reakční čas, sensorická přesnost) i komplexnější antropometrické nástroje.



Charles Spearman, Alfred Binet

Přelom 19./20. stol.: Prudký rozmach testování (zejm. didaktického).

- Predikce (životní úspěch, školní výkon) ale slabší, než očekávali.
- Doporučuji: Rogers, A. L. (1919). Tests of Mathematical ability - their scope and significance. *The Mathematics Teacher*, 11(4), 145–164. Dostupné [online](#).

Charles Spearman (1904): oslabení (attenuation) a základ CTT.

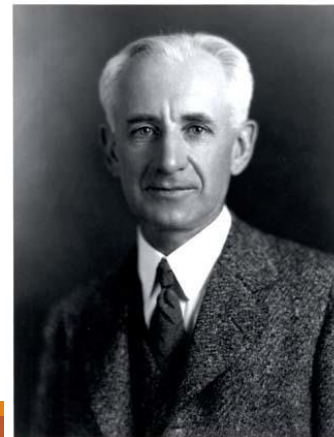
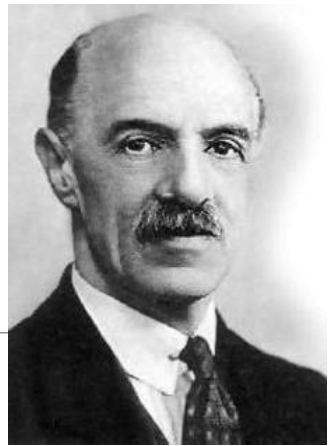
- Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, 15(2), 201. doi: [10.2307/1412107](https://doi.org/10.2307/1412107)
- Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1), 72. doi: [10.2307/1412159](https://doi.org/10.2307/1412159)

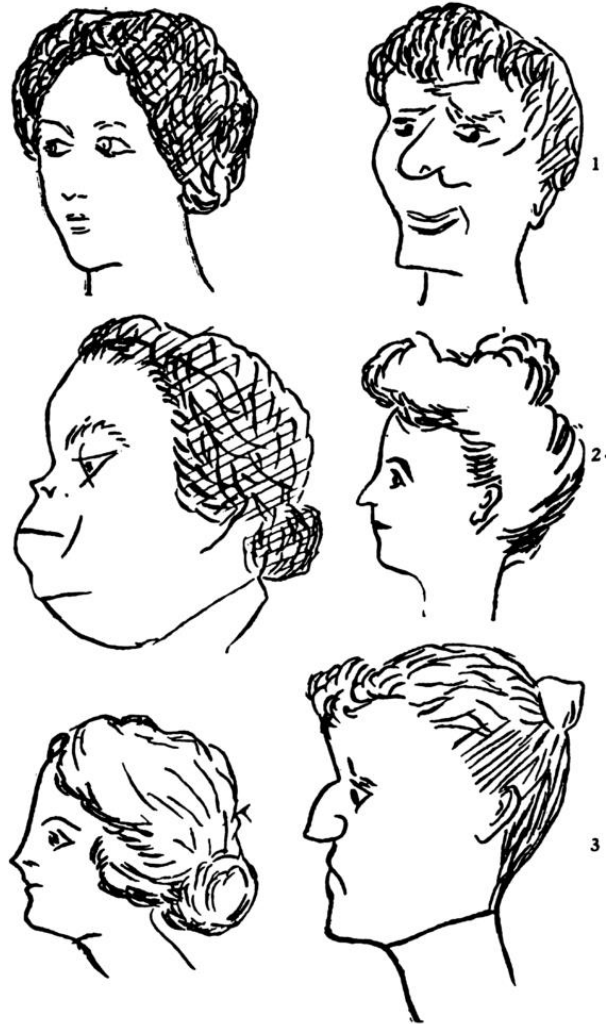
Alfred Binet (1905): testy školních předpokladů.

- Neměřil kognitivní procesy, prováděl jen *screening* a *predikoval*.
- Goddard a jiní však *index* interpretovali jako (dědičnou) inteligenci (*g*).

Binetovský formát + Spearmana teorie = podhoubí pro rozmach.

- Masivní nasazení psychodiagnostiky v USA na konci WW1 (**Yerkes, 1917**)
- Army Alfa/Beta: 1,5 milionu rekrutů.





THE PSYCHOLOGICAL CLINIC is indebted for the loan of these cuts and those on p. 225 to the courtesy of Dr. Oliver P. Cornman, Associate Superintendent of Schools of Philadelphia, and Chairman of Committee on Backward Children Investigation. See Report of Committee, Dec. 31, 1910, appendix.

Binet-Simon Intelligence Scale (1908)
 „Which of these two faces is the prettier?“

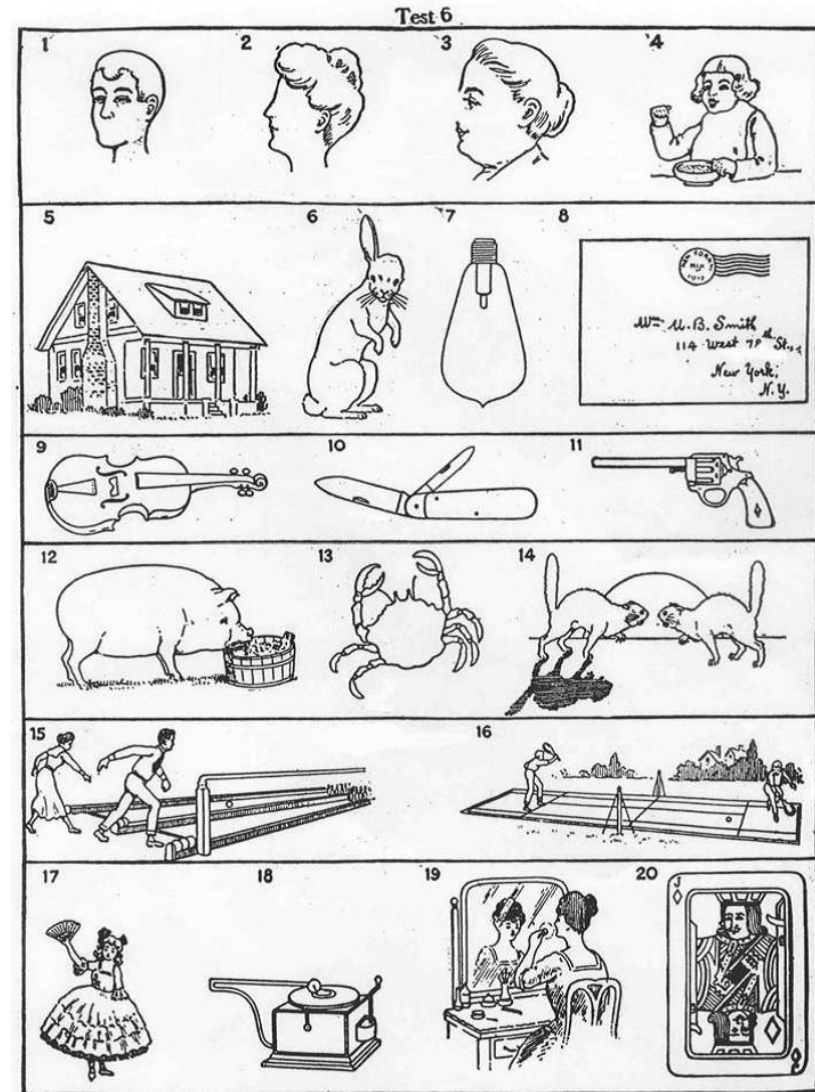


Fig. 63.—Group Examination Beta, Form 0, Test 6, Picture Completion.

Army Beta (Yerkes, 1917)
 „What is missing?“

1. pol. 20. století, cca 1915–1940

Rozvoj CTT, FA a vznik mnoha „velkých“ testů.

Fergusonova komise (Británie, 1932–1940).

- Požadavek aditivity, upozornění na nemožnost „sčítání“ (konkrétně „zřetězení“) v sociálních vědách. Závěr: psychologické měření není měření.

Reakce: „Operační teorie měření“ a „matching“ ([Stevens, 1946, s. 677](#)).

- „...*measurement, in the broadest sense, is defined as the **assignment of numerals to objects and events according to rules.***“
 - Ale: „*Measurement is a method of **assigning numbers to magnitudes***“ ([Helmholtz, 1887](#)).
 - Jsou čísla nezbytně magnitudami? ([Magnitude](#) = „*number characteristic of a quantity and forming a basis for comparison with similar quantities, as length*“).

1940–1970: Konsolidace

Konsolidace předchozích poznatků a restandardizace původních testů.

F. M. Lord a M. R. Novick (1968): *Statistical Theories of Mental Test scores*. (Dokončení CTT).

Cronbach, Nageswari a Gleser ([1963](#)): **Teorie zobecnitelnosti**.

50.–60. léta: **teorie odpovědi na položku** (IRT; zejm. Lord, [Rasch](#), Lazarsfeld).

Luce a Tukey ([1964](#)): **Teorie spojitého měření** (CM), vyvrácení závěrů Fergusonovy komise

Rozvoj konceptu validity (zejm. **konstruktová validita**; Cronbach a Meehl, [1955](#)).

Statistický vs. klinický úsudek: **Meehlovy experimenty** ([1954](#)).

- Groove a kol. ([2000](#)): 136 studií, 6 % klinická vs. 47 % statistická.

Hlavní limit: výpočetní kapacita.

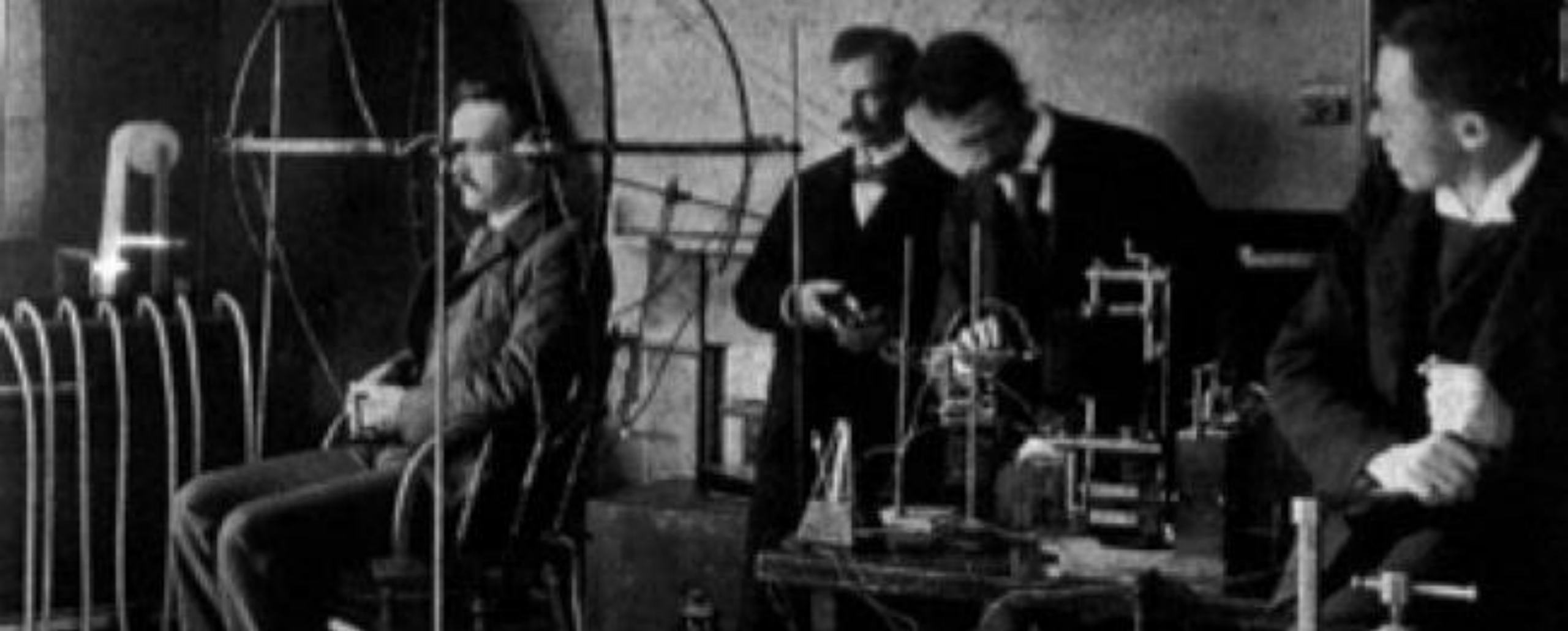
Od r. 1970: Moderní psychometrika

1970–2000: Rozvoj počítačů → vývoj nových postupů a zavádění do praxe.

- IRT, FA, SEM, počítačové modelování.
- Zavádění pokročilejších postupů (zejm. IRT) do praxe.
- Důraz na férovost, rozsáhlé důsledky pro individuální diagnostiku.
- Ale: od 60. let silná veřejná kritika psychologického testování v USA.

2000–dnes: Prudký teoretický rozvoj (nejen).

- Rozpracovávání starších přístupů, vznik/rozvoj nových (síťové modely).
- Ohromné množství nových testů, časté revize, důraz na evidence-based.
- Zavádění počítačů do praktického testování (CAT).
- Aplikace IT přístupů pro predikci chování: machine learning, neuronové sítě.



Měření v psychologii

Na fotografii interiér Wundtovy laboratoře v Lipsku.

(Pozn.: Pán vlevo má přivázané nohy a pouští do něj elektřinu, ale to se na obrazovku už nevešlo.)

Předpoklady měření (v psychologii)

4 klíčové předpoklady (Hogan, 2013):

- Měřené objekty (lidé) se liší v měřené charakteristice
 - (Proměnná, nikoli konstanta...).
- Rys či míra odlišnosti může být kvantifikována.
- Měřené atributy jsou dostatečně stabilní
 - (Vzhledem ke způsobu a účelu měření).
- Tyto atributy se projevují v pozorovaném chování.

Což ale nic neříká o povaze měřeného atributu, ani o postupu samotného měření.

Předpoklady měření (v psychologii)

Teorie měření: postuluje vztah pozorování s konstrukty.

Měření: atribut je na intervalové škále.

- Jinak jde o prostou **klasifikaci**.
- Což ale může být rovněž psychometrická výzva.

V psychologii občas „jen“ klasifikujeme → psychometrika není vědou výhradě o měření.

Co to je měření?

Měření: Množství a velikost.

Základní charakteristika měřeného pojmu/konstruktu/veličiny na (spojité?) intervalové či poměrové škále.

Už v 18. a 19. století (Kant, Leibniz aj.) definice „intenzity“:

- **Extensivní velikost** – celek se fyzicky skládá z částí (délka);
- **Intensivní velikost** – rovněž intervalové, ale projevují se „instantně“ (teplota, barva...).
- Překonané už v průběhu 19. století.

Co to je měření?

Aditivní operace: Základní předpoklad „dobré škály“.

Aditivita umožňuje převést funkci „+“ do „×“.

- $f(a+b) = f(a)+f(b)$.

Předpokladem aditivity je zejm. „řazení“ a „řetězení“.

Hodnoty lze přinejmenším sčítat a provádět běžné matematické operace.

Výsledkem je intervalová/poměrová škála se stejně velkými jednotkami a aditivní strukturou.

- Nutnost oddělit aditivitu veličiny a aditivitu měření.

Fundamentální měření

Fundamentální měření:

- **Základní (přímé):** Není odvozené z jiného měření, měří se přímo objekt za pomoci stejné veličiny *dle své definice*.
 - Délka pravítkem, hmotnost závažím, rychlost dráhou a časem.
- **Odvozené (nepřímé):** Odvozené pomocí aditivních operací z jiných naměřených hodnot (objem, teplota, síla zemětřesení na Richterově stupnici)
 - Obě tato měření jsou „**fundamentální**“, jsou to měření v pravém slova smyslu.
 - Někdy je za fundamentální označováno jen základní měření (nejasnosti v terminologii).

Podobné extensivní (základní) a intenzivní (odvozené) veličině.

- Jde ale o charakteristiku měření, nikoli veličiny.

Intermezzo

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).

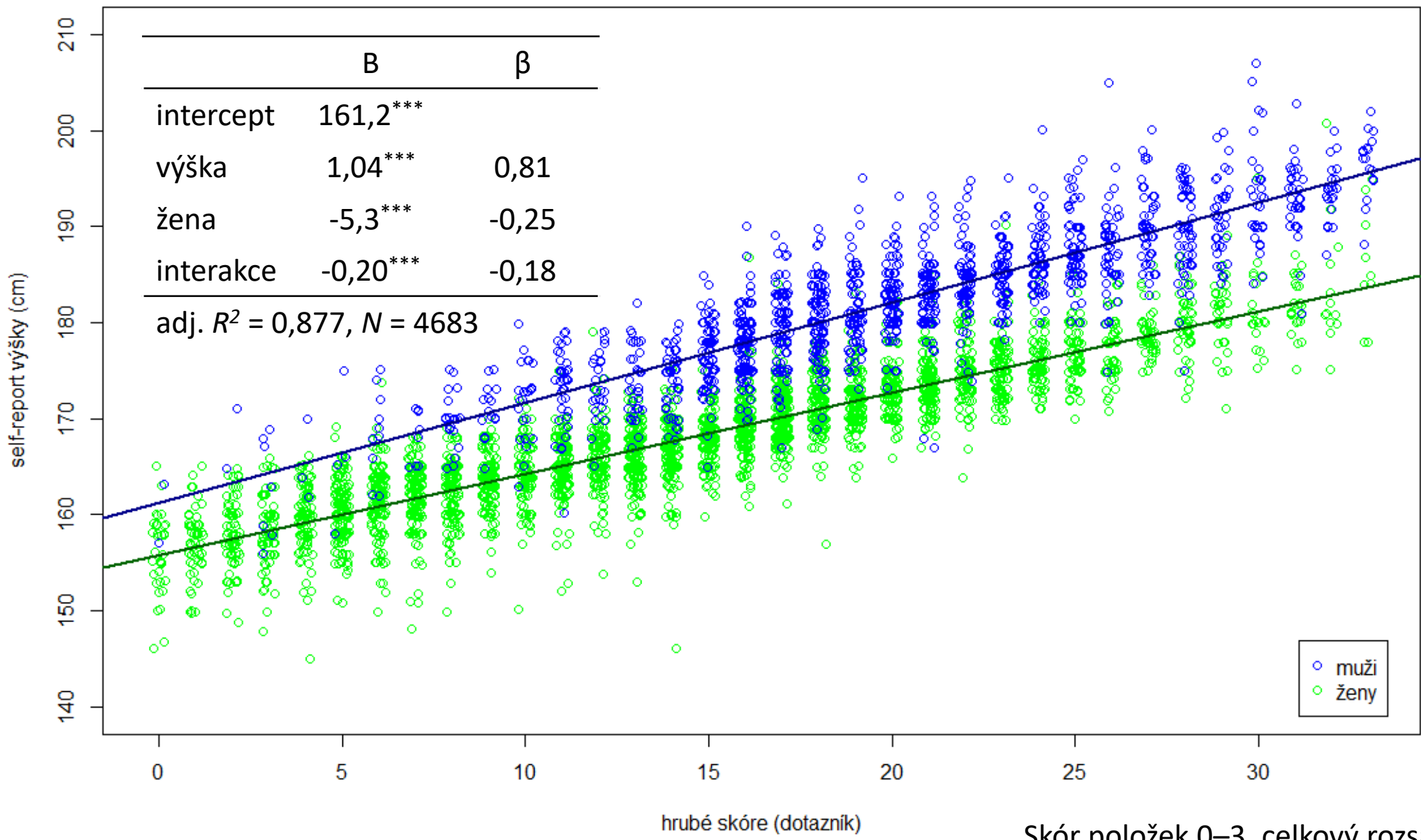
Použitou škálou je
**počet souhlasných
odpovědí** (rozmezí 0–11).

Je měření přímé či nepřímé?

- Proč ano?
- Proč ne?

Je výsledná škála aditivní?

- Proč ano?
- Proč ne?



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

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.

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.

- Operacionalismus vs. realismus.

Reflektivní vs. formativní měření.

- Kausální směr vztahu latentní a manifestní proměnné.

Úroveň latentní proměnné.

- intervalová vs. ordinální vs. nominální

Povaha intervalové proměnné

- spojitá vs. diskrétní

Intermezzo: Dotazník výšky

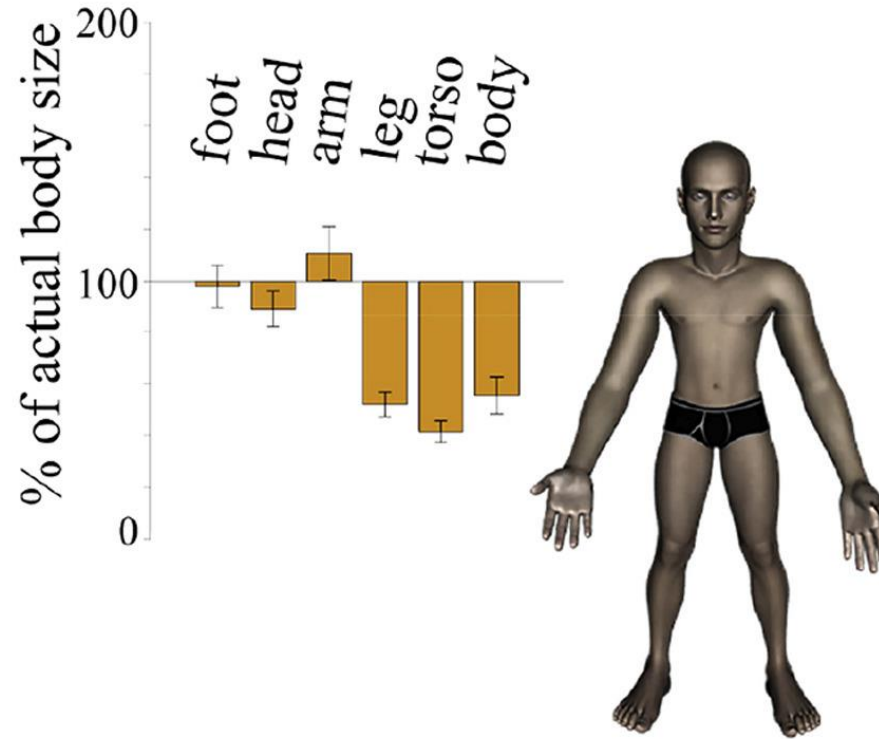
Měřit výšku dotazníkem je samozřejmě nesmysl:

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

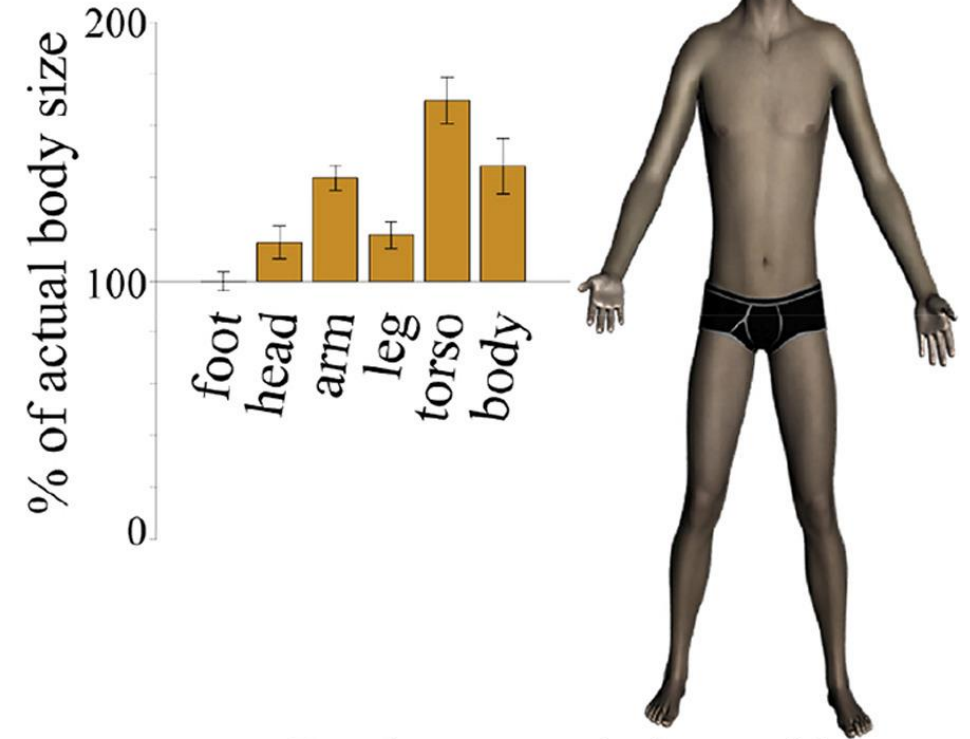
actual body



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., & Longoa, 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)

Teorie měření v psychologii

Teorie měření: 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.

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**.