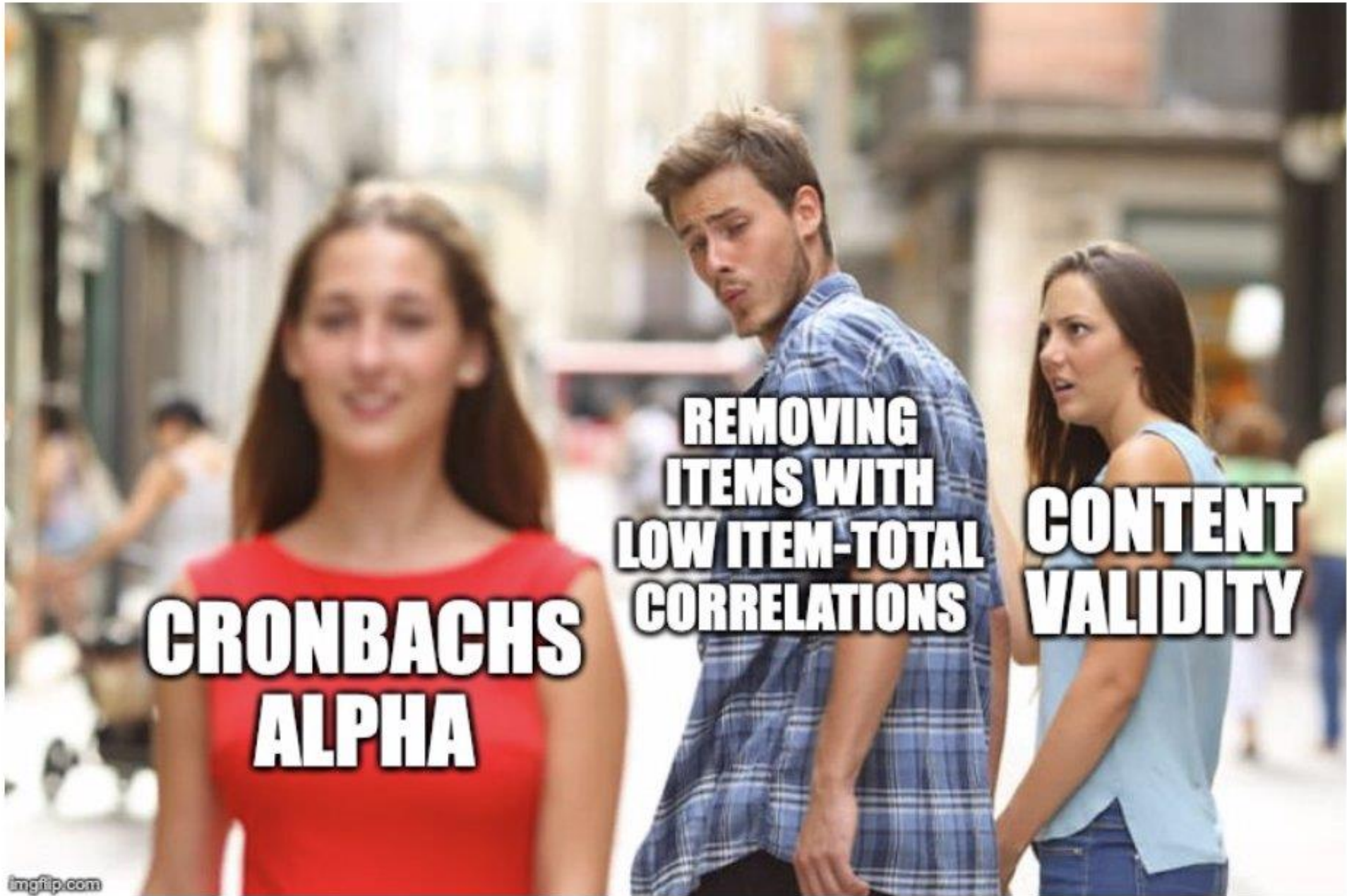


Přednáška 1: Úvod

14. 9. 2021 | PSYn4790 | Psychometrika: Měření v psychologii
Katedra psychologie, Fakulta sociálních studií MU

Hynek Cígler | cigler@fss.muni.cz



**CRONBACHS
ALPHA**

**REMOVING
ITEMS WITH
LOW ITEM-TOTAL
CORRELATIONS**

**CONTENT
VALIDITY**



Richard Feynman: „*Maybe some day they will...*“

Záměr kurzu PSYn4790

Poskytnout psychometrické a epistemologické zázemí k validnímu zkoumání člověka.

- **Epistemologická východiska** (týden 4–5).
- **Psychometrika:** Vybrané modely měření (týden 6–10).
 - Bližší představení vybraných teorií měření a jejich propojení s praxí a výzkumem.
- **Další témata** (týdny 1–3, 11–13).
 - **Úvod kurzu:** replikovatelnost, metaanalýza, krize zobecnitelnosti.
Práce s chybou měření (samostudium).
Proč stojí za to se orientovat v modelech měření?
 - **Závěr kurzu:** Férovost, shoda posuzovatelů, aplikace modelů měření pro praxi

K čemu vám kurz bude?

Problémy s měřením jsou jedním z příčin *replikační krize*.

Lepší orientace v modelech měření.

Přehled pokročilejších psychometrických postupů pro případné pozdější využití.

Propojení teorie měření a výzkumu.

Nástrahy sociálních věd, výzkumná „scestí“.

Aplikace psychometrických postupů do praktické diagnostiky, využití znalostí z bakalářského studia v praxi.

Stručný harmonogram

#	datum	téma	přednášející	úkol
1.	14. 09.	úvod, zadání seminární práce	HC	P1
2.	21. 09.	meta-analýza, replikovatelnost, replikační krize, open science	HC	P2
3.	28. 09.	<i>samostudium/opakování (chyba měření)</i>	<i>státní svátek</i>	
4.	05. 10.	epistemologie 1	HŠ	P3
5.	12. 10.	epistemologie 2	HŠ	test 1
6.	19. 10.	klasická testová teorie	HC	P4
7.	26. 10.	teorie zobecnitelnosti	HC	P5
8.	02. 11.	IRT 1	HC	P6
9.	09. 11.	IRT 2	HC	seminárka
10.	16. 11.	síťové modely	ECh	P7
11.	23. 11.	férovost	HC	P8
12.	30. 11.	shoda posuzovatelů	HC/AŘ	P9
13.	07. 12.	Závěrečné setkání: příklady z praxe	HC	test 2

Intermezzo: Product placement



Intermezzo : Product placemenet

The banner features a dark blue background with a large, multi-faceted, crystalline structure on the left side, rendered in shades of red, orange, and pink. The text is positioned to the right of this structure.

METASCIENCE
2021 CONFERENCE

A GLOBAL VIRTUAL CONFERENCE

September 16-18 and 23-25, 2021

08:00-22:00 EDT | 05:00-19:00 PDT | 14:00-04:00 CEST | 22:00-12:00 AEST (+ 1 day)

Intermezzo : Product placement

The Metascience 2021 Conference is a global virtual gathering to connect the study of science across disciplines, methodologies, and regions. It follows the inaugural [Metascience 2019](#) Symposium held at Stanford University. Metascience 2021 is an initiative of the Center for Open Science (COS), the Association for Interdisciplinary Meta-Research and Open Science (AIMOS), and the Research on Research Institute (RoRI) and is generously supported by the Templeton World Charity Foundation and the RoRI consortium.

www.metascience2021.org

Students/postdocs	General Admission	Discount
\$ 10	\$ 20	Free
<small>Sales end on Sep 25, 2021.</small>	<small>Sales end on Sep 25, 2021.</small>	<small>Tickets for those who would like to attend but do not have available funding.</small>

Psychology's crisis of confidence: Measurement edition
[Jessica K. Flake](#), [Eiko Fried](#), [Andrea Helena Stoevenbelt](#)
Moderator: [Esther Maassen](#)

*In psychology's crisis of confidence, issues related to measurement did not receive much attention. However, ignoring challenges of measurement—one of the foundations of science—has **serious consequences for the validity of the inference, replicability, and reproducibility of research**. In our symposium, we highlight empirically studied measurement issues and present steps researchers can take to prevent them. [...] Our aim is to emphasize to researchers that **measurement is a crucial part of the research process** that should not be neglected in the broader quest to **improve research practices in the social sciences**.*

<https://metascience2021.org/events/psychologys-crisis-of-confidence-measurement-edition/>

Požadavky na ukončení kurzu

2 písemné testy: $2 \times 25 = 50$ bodů

- Na konci každého bloku, termíny v ISu (15. 10. a 10. 12.).
- Minimum 10 b. (X). Opravné termíny ve zkuškovém.

A: 100–91

B: 90–81

C: 80–71

Seminární práce: 25 bodů.

- 10 a méně bodů → možnost přepracovat (-5 bodů).

D: 70–61

Ústní zkouška: 25 bodů.

- Minimum 1 😊

E: 60–51

F: 50–0

Nebodované přípravy.

- Alespoň 7 z 9 (jinak X)

Požadavky na ukončení kurzu

Písemné testy

- Open-book, otevřené otázky.
- On-line odkudkoli.
- Lze využít libovolné materiály.
 - Ty ale nestačí, je potřeba se orientovat.
- Jednodenní časové okno (pátek?).
 - Na konci každého bloku.

Průběžné přípravy

- Slouží pro vás.
- Nebodováno, bez individuální zpětné vazby.
 - Nepište ale pitomosti, budeme to procházet 😊
- Není potřeba se nic učit.
 - Stačí přečíst literaturu.
- Třídenní časové okno (st–pá?).
 - Mimo týdny s testem nebo seminární prací.
- Nečíst detailně, není to beletrie.

Vaši předchůdci „vzkazují“:

Jde to zvládnout! 😊

Obtížné, pokud jste neabsolvovali bakalářskou psychometriku.

Nezbytné studovat průběžně.

Nestačí přednášky, bez literatury to nepůjde.

- Nespoléhejte, že na přednáškách zazní všechna témata.
- Bez přednášek chybí kontext.
- Podívejte se na tipy v interaktivní osnově ke čtení literatury.

Věnujte pozornost zadání SP (struktura recenze).

Věnujte přípravě **adekvátní** čas (ani moc, ani málo).

Studijní zdroje

[Interaktivní osnova](#) a IS MU, **předmětová diskuze**.

FB: [Psychologické testování](#)

- Odborná diskuze o psychodiagnostice a metodách

FB: [Statistika, metodologie, psychometrika](#)

- Prostor pro řešení běžných problémů, dotazy ke studiu.

FB: [Kvantitativní kroužek](#)

- Advanced skupina o statistice a metodologii v sociálních vědách.

Časopis [Testforum](#).

Povinná a rozšiřující literatura vždy ke každému tématu.



Konzultace

Najdete u každého vyučujícího v IS MU.

Organizační záležitosti: Hynek Cígler.

Hynek Cígler: kanc. 2.47, středa 14.00–15.30.

- Bez garance.
- Možnost on-line konzultace bez předchozího ohlášení (stejný čas, MS Teams).

Preferujeme cokoli, jen ne e-mail! 😊



Zadání seminární práce

Recenze zvolené psychodiagnostické metody

Zadání seminární práce

Recenze komerčně dostupné psychodiagnostické metody určené pro diagnostiku v běžné psychologické praxi.

- V „běžném“ formátu odpovídajícím tuzemským zvyklostem.
- Formát recenze by měl odpovídat „manuskriptu“ recenzního článku.

Cíle SP

Projít si a ujasnit si, jaké všechny vlastnosti testů nás zajímají.

Zkusit tyto vlastnosti objektivně popsat a zhodnotit.

A to v kontextu nějakého konkrétního testu, který vás zajímá a jehož konkrétní použití si dovedete představit.

Integrovat tyto soudy.

A vypracovat text, který by si mohl vzít do ruky uživatel testu – prakticky orientovaný psycholog – jako podklad pro výběr vhodného testu pro nějaký účel.

- Možnost publikace v Testfóru.

Seminární práce nenavazuje přímo na obsah přednášek, ale doplňuje zbytek kurzu o nová témata (s návazností na psychologickou diagnostiku).

Postup tvorby SP

Studium Burosových ročenek, EFPA manuálu.

Výběr vhodné metody (neváhejte konzultovat).

Prostudování manuálu, podnětového materiálu a dalších součástí.

- Možnost využít cvičné administrace do kurzu PSYn4790 (Psychodiagnostika dospělých)

Doplnění pomocí relevantních empirických zdrojů (existují-li).

Vypracování EFPA formuláře.

Převedení formuláře do souvislého textu.

Odevzdání formuláře i textu (do 15. 11.).

- text: „jméno příjmení_recenze.docx“
- formulář: „jméno příjmení_formulář.docx“

Doporučený harmonogram: Propojení psychometrie a diagnostiky

První měsíc: Zamyslete se, *jaká metoda vás zajímá* (recenze i vyšetření).

20. 10.: Termín 1 v PSYn4020 (Psychodiagnostika): Téma & metody

- Vyberte si ideálně stejnou metodu pro vyšetření i recenzi.
- K tomuto termínu byste měli do metod již nahlédnout.

Druhý měsíc: Seznamte se s metodou, pročtěte si všechny materiály.

- Můžete již realizovat vlastní vyšetření.

15. 11.: Termín pro recenzi testu do PSYn4790 (Psychometrie).

- Zjistíte limity a omezení dané metody.
- Můžete využít poznatky z administrace, pokud ji stihnete.

10. 12.: Termín 2 v PSYn4020 (Psychodiagnostika): Vlastní zpráva z vyšetření.

- Při interpretaci výsledků využijte limity a omezení zjištění při tvorbě recenze.

Důležité zdroje a inspirace

Evers a kol. (2013): článek o EFPA modelu, doi: [10.7334/psicothema2013.97](https://doi.org/10.7334/psicothema2013.97)

Urbánek, T. (2010): článek o recenzním procesu, doi: [10.5817/TF2010-1-1](https://doi.org/10.5817/TF2010-1-1)

- Představení recenzního modelu českému publiku.

Kabinet diagnostických metod: <https://psych.fss.muni.cz/studenti/kabinet-diagnostickych-metod>

- Primárně vybírejte metody dostupné na katedře.

Testfórum: <https://testforum.cz/about/editorialPolicies#sectionPolicies>

- Informace o recenzích v Testfóru, [formulář recenzního modelu EFPA](#) (verze CZ 3.42)
- Inspirace: dosud publikované metody: <https://testforum.cz/pages/view/recenze>

Burosovy ročenky, dostupné v knihovně (D2-732; D2-732a)

- Inspirace ohledně struktury a obsahu samotného textu.

EFPA manuál

POPIS (NEHODNOTÍCÍ)

obecný popis

klasifikace

skórování

generované zprávy

dodavatel a náklady

ZHODNOCENÍ METODY

kvalita osvětlení teoretických východisek

kvalita materiálů

psychometrické parametry

- normy
- reliabilita
- validita
- (kvalita generovaných zpráv)

závěrečné zhodnocení a hlavně doporučení

literatura

Vlastní text: Struktura

Řidíte se (nezávazně) strukturou Burosových ročenek a recenzí v Testfóru.

Záhlaví (vybrané důležité informace, autoři, distributor aj.).

Popis testu, jeho určení, cílové populace atp. (nehodnotící, vychází z informací autora).

Vývoj metody a kvalita technických materiálů (popis s hodnotícím komentářem).

Technické parametry (kritické zhodnocení na úrovni faktů, mělo by obsahovat i další studie v manuálu neobsažené).

- zejm. validita, reliabilita, normy, počítačové zprávy (existují-li)

Komentář (zhodnocení faktů uvedených výše).

Shrnutí a závěr s konkrétním doporučením.

Literatura, zdroje

Doporučení

Otevřená pole formuláře mohou obsahovat identický (copy-past) text s vlastní recenzí. Formulář je jen doplněk, primárně je hodnocen text.

- Otevřená pole klidně vynechte.

Doporučujeme pracovat se stejnou metodou, jako v Psychodiagnostice.

- Zvážení psychometrických parametrů → výhoda v diagnostice.
- Zkušenost s administrací → výhoda v recenzi (nejen hodnocení testového materiálu).

Nevybírejte metodu na první dobrou.

- Některé metody je snadnější recenzovat než jiné.
- Hodně špatná metoda → velmi náročné hodnocení.

Nevybírejte příliš „velké“ metody. Poradte se s obsluhou.

Chcete-li vybrat metodu, která není v KDM, domluvte se osobně.

Seminární práce: Tipy

Věnujte pozornost struktuře recenze (opravdu mrkněte na [Burosovy ročenky](#)).

Využívejte **externí empirické zdroje**.

- Manuály (zejm. českých) metod nejsou dostatečné a přiměřeně kritické.

Využijte Messickovo/Lissitzovo pojetí validity („shoda skóru s interpretací“).

Kombinujte originální a české zdroje.

- Lze přenést důkazy validity? Lze se domnívat, že něco funguje stejně?

Věnujte pozornost normám.

- Odpovídají účelu metody? Reprezentuje vzorek cílovou populaci?

Budte kritičtí, zdrojujte tvrzení („Podle autorů metody...“).

- „Metoda je vhodná...“, to je už váš závěr. „Autoři tvrdí, že je vhodná...“ je v úvodu vhodnější.

Alternativní zadání

Recenze testu vypracovaná ve dvojici v publikační kvalitě.

- Musí jít o recenzi metody využívané v české praxi.
- Je nutný **individuální souhlas** a konzultace s HC (cca do 14 dnů).

Datum odevzdání: Co nejdříve (listopad?), protože:

- **25 bodů:** práce odevzdána v kvalitě manuskriptu recenze a text je (**po případných revizích**) přijat k publikaci v Testfóru **do konce zkouškového období** (13. 2. 2022).
- **15 bodů:** splnění požadavků na individuální SP, nedosáhla publikační úrovně.
- **0 bodů:** nepovedlo se to. Jeden pokus na vypracování individuální SP za běžných podmínek.

Psychometrika



Psychometrika

Co to je psychometrika:

- Pomocná disciplína všech základních psychologických věd.
- Na pomezí psychologie (zejm. interindividuálních rozdílů), kognitivní psychologie (psychologie dotazování) a matematické statistiky.

Cíle a oblasti použití psychometriky:

- Praktická činnost, tedy ověřování a konstrukce testů.
- Teorie: vývoj statistických postupů k lepšímu měření člověka.
- Nedílná součást **praktické** psychologické diagnostiky.

*„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 studuje...

... vztah manifestních a latentních proměnných.

... podstatu a charakteristiku latentních psychických procesů.

... možnosti a chybovost usuzování na latentní proměnné z manifestních.

... a vysvětluje a predikuje chování lidí.

Za tímto účelem pracuje s různými **teoriemi**, které postulují statistické **modely**.

Manifestní proměnná: pozorované chování.

Latentní proměnná/proces: Měřený atribut / jev.

Latentní procesy a proměnné

Způsob definice závisí na teorii měření.

- Pravý skór (operacionalismus, CTT).
- Latentní rys (realismus, IRT, FA, LCA).
- Symptomatologická síť (network modely).
- Spojitá, „kvantitativní“ (FA, IRT) vs. nominální, „kvalitativní“ (LCA, network) povaha.

Tradiční psychodiagnostický pohled:

Latentní proměnná je **interpretace testových výsledků v širším kontextu.**

- Konstruktivismus, antirealismus.
- Messickovo pojetí validity, Lissitzův model pro hodnocení testu.

Více viz blok přednášek z epistemologie.

Základní koncepty psychometrie

Opakování z bakaláře:

nezbytné statistické pojmy

práce s rozptylem

validita a související pojmy

reliabilita, způsoby odhadu

chyba měření a intervaly spolehlivosti

já jsem ale po napsání
své bakalářky strašný
psychometrický skeptik..

už chci odejít do ústraní
a dělat jen kvalitu 😄 😄



Základní statistické pojmy

Úrovně měření.

- Stevensonova definice, fundamentalita měření...
- Nominální, ordinální/pořadová, metrická (intervalová vs. poměrová).

Míry centrální tendence, rozptyl, kovariance a korelace, rozložení proměnných.

Statistické analýzy: korelace, parciální korelace

ANOVA, lineární a logistická regrese, neparametrické statistiky.

Explorační a konfirmační faktorová analýza, logika strukturních modelů.

Práce s rozptylem, zejm. [sčítání rozptylových komponent](#).

- Jak z kovarianční matice položek odhadneme rozptyl součtového skóru?

Položky, konstrukce položek

Tvorba položek

- teorie faset, dekompozice obsahového univerza
- rozdíl fasety vs. dimenze/faktor
- složení položky: podnět vs. odpověď (stimulus vs. response)
- druhy položek: doplňování, dichotomické, multiple-choice, řazení, volná odpověď (plus její skórování), mnohonásobná volba...
- odpověďové škály (Likertova, sémantický diferenciál, Guttmanovo škálování...)

Podoba metody/testu:

- druhy škál: typologie/ipsativní, ordinální, intervalové...
- druhy metod: testy schopností, osobnostní dotazníky, silové vs. rychlostní vs. kapacitní, introspektivní vs. výkonové, objektivní, projektivní...

Položková analýza, faktorová analýza

POLOŽKOVÁ ANALÝZA V CTT

Popularita, obtížnost

Korigovaná korelace položky s celkovým skóre

Cronbachova alfa po vyřazení položky

ULI (upper-lower index)

deskriptivy...

FAKTOROVÁ ANALÝZA

Explorační faktorová analýza

- Metody odhadu počtu faktorů – paralelní analýza, Kaiserovo pravidlo, Cattelův scree-plot
- Metody extrakce, eigenvalue
- Metody rotace (šikmé a ortogonální)
- Faktorový náboj, komunalita, unicity
- Heywoodův případ

Konfirmační faktorová analýza

- Shody modelu s daty, χ^2 , SRMR, RMSEA, CFI/TLI..., inspekce reziduí, M.I.
- Definice a odhad modelu.

Předpoklady faktorové analýzy.

Reliabilita: Metody odhadu v CTT

Paralelní formy: korelace, ICC.

Test-retest: korelace, ICC.

Shoda posuzovatelů: řada ukazatelů, samostatná přednáška.

Vnitřní konzistence:

- Split-half: Spearman-Brownův vzorec, Guttmanova λ_4 ...
- „Kovarianční matice“: Cronbachovo alfa (= KR-20), omega...

Reliabilita

$$r_{xx'} = \frac{\sigma_{\tau}^2}{\sigma_x^2} = \frac{\sigma_{\tau}^2}{\sigma_{\tau}^2 + \sigma_e^2} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$$

Podíl vysvětleného rozptylu měřicího nástrojem měřeným atributem.

- A tedy odmocnina z korelace manifestní a latentní proměnné.

Očekávaná korelace paralelních testů.

(Reliabilitě se více budeme věnovat v 6. týdnu kurzu.)

Normy

Způsob tvorby standardizačního vzorku a výběru respondentů, typy norem.

- Lokální, národní, institucionální...

Kriteriální, ipsativní, normativní skóry a normy.

- klasická vs. kontinuální (vyhlazená) normalizace.

Typy standardních skóru běžně používaných v psychologii a vzájemný převod.

- z-skóry, T-skóry, IQ-skóry, steny, staniny, percentily, percentilové pořadí...

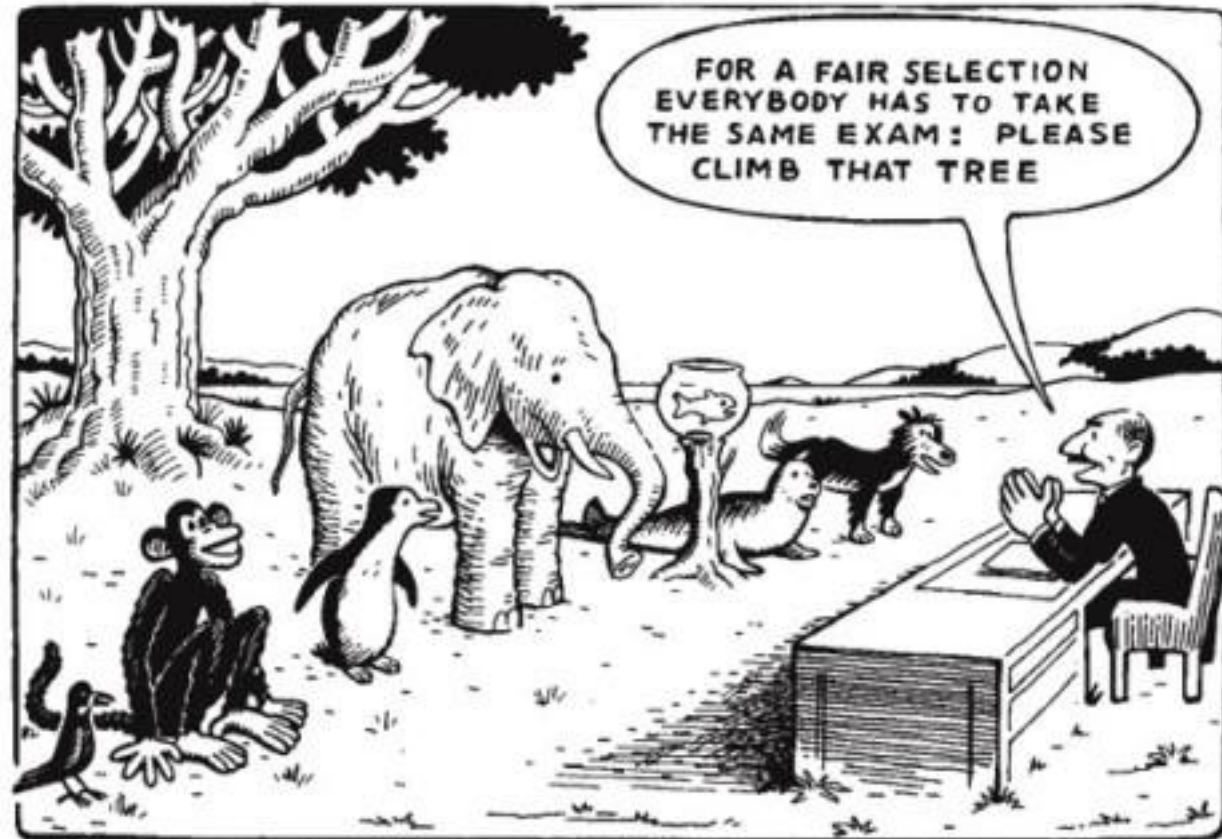
Problémy ze zešikmením, normalizace rozložení (McCallova plošná transformace, rozdělení podle mediánu...).

Kontinuální normy (vertikální a horizontální normy).

Validita

Messickovo pojetí validity

Lissitzův model validity



Validita

Unifikovaná konstruktová validita (Messick)

Ontologické pojetí (Borsboom)

Tradiční pojetí: obsahová, empirická, konstruktová

Klíčové pojmy z tradičního uvažování o validitě:

- Výběrová v., face-validity (zjevná, zdánlivá), faktorová.
- prediktivní vs. postdiktivní, inkrementální
- souběžná, konvergentní, divergentní/diferenciální, kriteriální
- Cronbach a Meehl ([1955](#)): Construct Validity in Psychological Tests, nomologická síť.
- Campbell a Fiske ([1959](#)): Multitrait-multimethod matrix (MTMM).

Definice konceptu validity

Messick (1989, s. 20): „... *an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment.*“

- Konstruktivismus, operacionalismus.

Borsboom (2004): „A test is valid for measuring an attribute if (a) the **attribute exists** and (b) variations in the attribute **causally produce variation in the measurement outcomes.**“

- Validita je vlastnost testu, kterou hodnotíme.
- Realismus, ontologické pojetí.

Unifikovaná konstruktová validita

Důraz na hodnocení a použití testu v diagnostice.

- Validita je jediným, multifasetovým konstruktem.
- Validita je integrativním shrnutím dílčích důkazů.
- Integrována ve Standardech pro pedagogické a psychologické testování (AERA, 2014).

Zdroje důkazů:

- Obsah testu
- Vnitřní struktura testu
- Odpověďové procesy
- Souvislost s kritériem
- Konsekvence testování

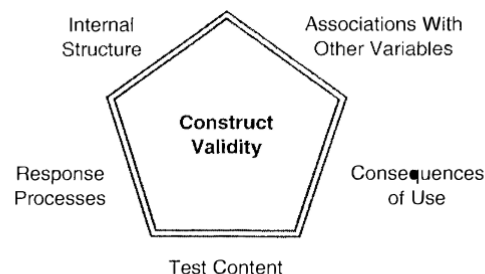
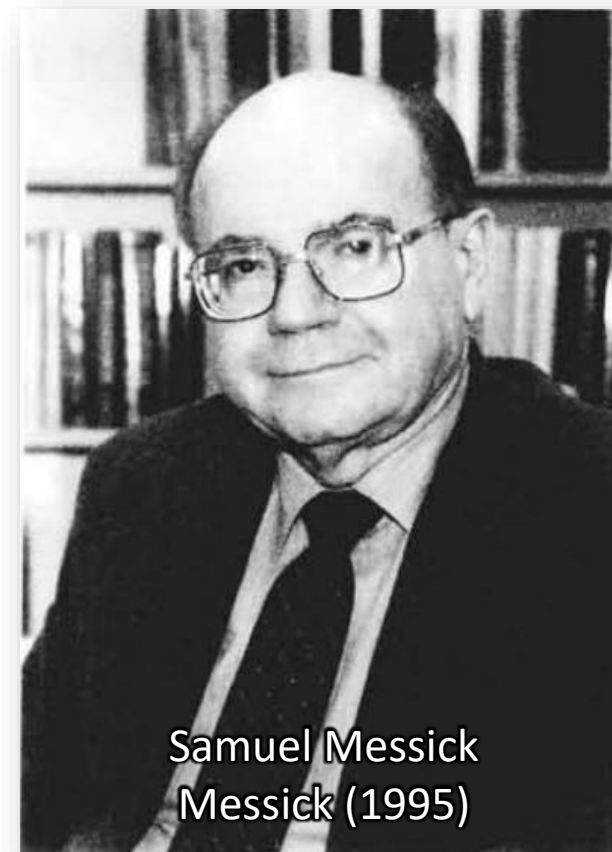


Figure 8.1 A Contemporary Perspective of Types of Information Relevant to Test Validity



Samuel Messick
Messick (1995)

Otázky spojené s Messickovým pojetím validity (Lissitz & Samuelsen, 2007)

Znamenají nízké korelace s kritériem skutečně nízkou validitu měření?

Jak v diagnostické praxi zhodnotit „globální užitečnost“ nástroje?

Jsou skutečně všechny aspekty důležité při hodnocení metody součástí validity?

Mají všechny atributy metody při hodnocení stejnou váhu?

Jak souvisí teorie a specifikace konstruktů při konkrétním měření?

Pokud je reliabilita podmínkou či součástí validity, proč ji Messick explicitně nezmiňuje?

Termín „validita“ by měl být vyhrazen pro schopnost metody zachytit určitý „rys“.

- Otázka „obsahové validity“ a reliability.

Validita jako hodnocení testu

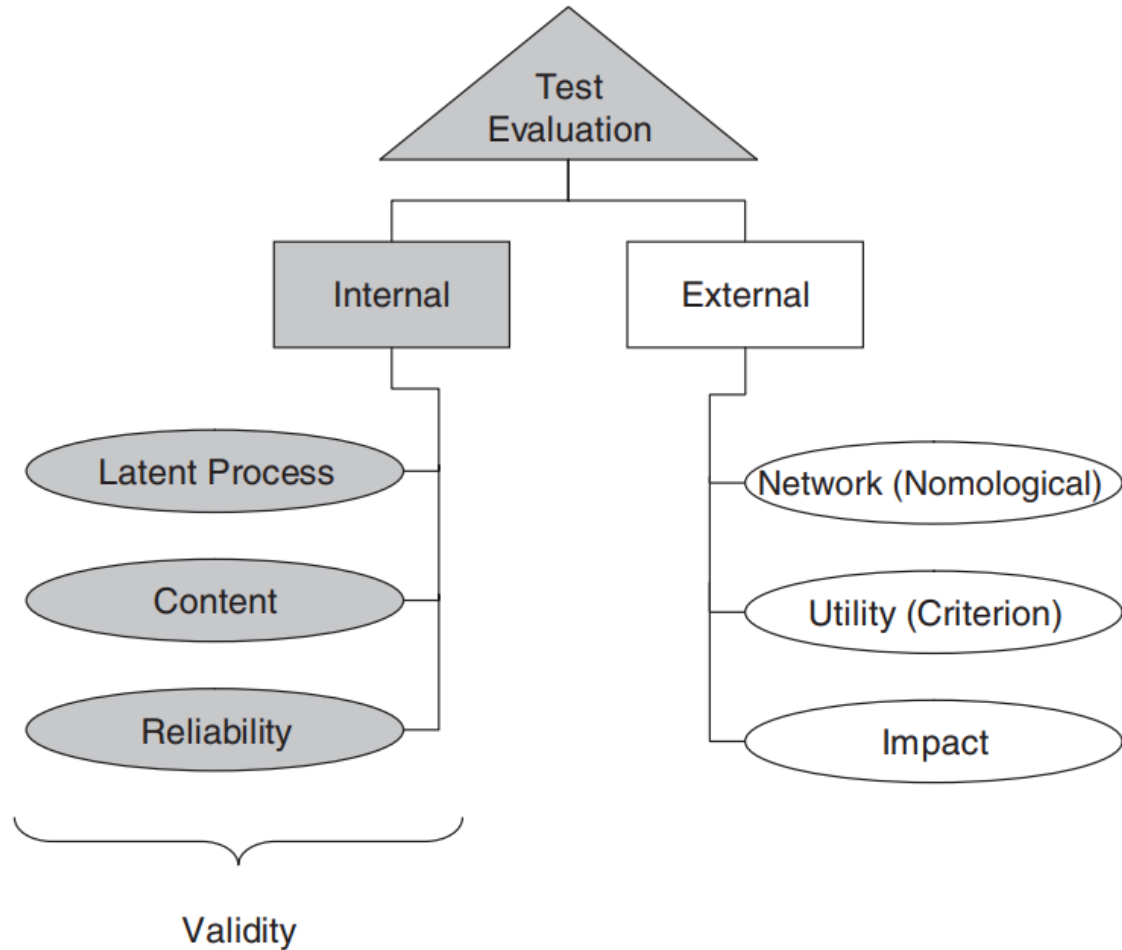


FIGURE 1. *The structure of the technical evaluation of educational testing.*

		Perspective	
		Theoretical	Practical
Investigative Focus	Internal	Latent Process	Content and Reliability
	External	Nomological Network	Utility and Impact

FIGURE 2. *Taxonomy of test evaluation procedures.*

Lissitz, R. W., & Samuelsen, K. (2007). A Suggested Change in Terminology and Emphasis Regarding Validity and Education. *Educational Researcher*, 36(8), 437–448. doi:[10.3102/0013189x07311286](https://doi.org/10.3102/0013189x07311286)