

- a) Zadání seminární práce
 - b) Výběrová validita
 - c) Psychologické škálování
-

PSYb2590: Základy psychometrie | Seminář 1

13. 2. / 20. 2. 2023 | Hynek Cígler, Petr Palíšek a Petra Hubatková

A man with a long brown beard and a red baseball cap is walking towards a large group of people on a paved road. The man is wearing a dark blue jacket and black pants. The group of people, representing a class, is standing in a line on the road ahead of him. The background is a dry, hilly landscape with sparse vegetation.

My class

**Me, teaching on a
topic I knew nothing
about two weeks ago**

Zadání seminární práce

Seminární práce 1: Téma práce

Téma práce: Návrh postupu tvorby a validizace psychodiagnostické metody určené k měření zvoleného konstruktů včetně pilotního ověření

Cíle seminární práce:

- Naučit se tvořit položky jednoduché psychologické škály.
- Osvojit si uvažování o obsahové a konstruktové validitě.
- Vyzkoušet si příslušné analýzy (FA, položková analýza...).
- Vyzkoušet si napsání výzkumného článku v odpovídajícím formátu.

Skupiny po 4 osobách (výjimky možné po schválení).

Průběžné odevzdání, průběžné konzultace, zpětná vazba spíše ústně.

Nelze opravit, není minimální počet bodů.

Seminární práce 2: Struktura práce

Struktura odpovídá formátu IMRAD. Doporučený rozsah do 15 NS bez příloh a tabulek.

Úvod: Podrobný popis měřeného konstruktů a teorie, ze které vychází, související pojmy, účel metody, očekávaná struktura. Původní i recentní literatura.

- Východiska pro ověření konvergentní, divergentní, faktorové validity.

Metoda: Popis vzorku a jeho výběru, odhad a zdůvodnění jeho velikosti.

- Návrh způsobů ověření validity a reliability (nikoliv všechny musí být realizovány v rámci SP).
- Není-li možné zajistit dostatečně velký vzorek z důvodu časově náročné administrace, lze data nasimulovat na základě několika málo (alespoň 10) administrací.
- Způsob tvorby položek, nepovinně včetně kognitivního interview.

Výsledky: deskriptivy, položková analýza, odhad reliability, faktorová analýza, další ověření validity (pokud byly provedeny).

- Případné vyřazení nefungujících položek a návrh upravené verze instrumentu.

Diskuze: zhodnocení fungování metody a návrh dalších kroků.

- Včetně kvalitativního zhodnocení svých zkušeností.

Seminární práce 3: Organizace

Skupiny: 4 studenti (jiný počet jen výjimečně).

- **Konzultanti:** Hynek, Petr, Jakub, Tomáš, Karel

Zpětná vazba výhradně/spíše ústně.

- Konzultace po domluvě.

Termín A (26. 2.): Vytvoření témat v ISu

- Včetně přihlášení členů týmu.
- Stručný popis tématu.

Termín B (27. 3.): Teorie

- Až po tvorbu položek a sběr dat.
- Celý Úvod + část kapitoly Metoda.
 - Tvorba vzorku, odhad jeho velikosti, způsoby ověření validity/reliability, očekávaný počet položek.
 - Návrh tvorby položek a ukázkové položky.

Termín C (17. 4.): Schválení sběru dat

- Hotová metoda, položky, informovaný souhlas.
- Ideálně hotová metoda připravená ke sběru dat.

Termín D (8. 5.): Hotová práce

- Finální verze celé SP.
- Ve formátu APA manuskriptu.
- Kompletní datová matice, skript, metoda.

Hodnocen je jen finální text.

Seminární práce 4: Hodnotící kritéria

Je konstrukt **dobře operacionalizovaný**, aby bylo možné hodnotit kvalitu jeho měření?

- Je použita adekvátní teorie pro teoretické zakotvení konstruktů?

Má navržený **design měřicí metody** potenciál k měření zvoleného konstruktů?

Je adekvátní **návrh způsobu tvorby položek**, jsou **výsledné položky** kvalitní vzhledem k účelu?

Jsou postupy **ověření validity a reliability** vhodné pro daný konstrukt a design metody?

Je **návrh vzorku** vhodný pro pilotní ověření dané metody?

- (Nehodnotíme, pokud se sběr dat nepodaří realizovat zcela podle představ, důležitý je záměr.)

Jsou **statistické analýzy** korektně provedené? Jsou výsledky dobře reportované?

Je text na dostatečné **formální úrovni**?

- Reportování statistik, odborný jazyk. Přehlednost textu. Stylistika a pravopis.

Jsou **výsledky** dobře reportované? Jsou dostatečně **diskutované** v návaznosti na teorii?

Jaká byla **prezentace výsledků** na semináři?

Seminární práce 5: Technické

Založte si rozpis v IS MU (viz interaktivní osnova).

- Můžete do něj rovnou přidat kolegy (jako studenti).
- Rozpis pojmenujte názvem týmu nebo něčím, co vystihne vaši metodu.
- Do anotace dejte pár slov o plánované metodě.

Rozpis bude schválen vyučujícím.

- Ke každému rozpisu se automaticky vytvoří odevzdávárna.
- Do ní budete vkládat obě části seminární práce a související soubory.

Rozpisy můžete využít i jako „burzu“ témat.

- Můžete využít i diskuzi v IS MU.

Seminární práce 5: Technické

Formát textu odpovídá APA manuskriptu výzkumného článku.

- Formátem, rozsahem i strukturou.
- Dodáme šablonu odpovídající APA formátu.
 - Tj. obsahuje úvodní stranu, abstrakt, running-head a další náležitosti.
- Tabulky a ilustrace nedávejte nakonec, ale **na správné místo přímo do textu!**

Na titulní straně v poznámce autorské podíly.

- Nehrají roli při hodnocení, prostě nás to jen (anonymně) zajímá.
- Mohou tedy být „anonymizované“ 😊

Součástí finálního odevzdání jsou:

- Text včetně znění položek (např. v příloze).
- Data, případně též analytický skript.
- Odkaz na dotazník (případně jeho kopie).
- Vzor informovaného souhlasu (pokud byl použit).

V odevzdáárně tedy na konci semestru bude:

- První verze textu.
- Odkaz na dotazník/metoda.
- Druhá verze textu, obsahuje i přílohy.
- „Supplemental materials“, jsou-li (soubor s daty, skriptem...).

FAQ: Co bývá nejčastější chyba?

Podcenění teoretického ukotvení konstruktů

Z textu by mělo jasně plynout, jaká je podstata konstruktů, jak se má manifestovat a/nebo s jakými konstrukty by měl souviset / nesouviset

Slabý úvod pak negativně ovlivní i tvorbu položek (není jasné, jak je tvořit), design studie (není jasné, co vlastně chceme zjistit) a diskusi (výsledky se těžko interpretují)

Technické chyby v analýze jsou mnohem menší problém než špatná nebo chybějící teorie

FAQ: Co bývá druhá nejčastější chyba?

Netransparentní postup tvorby položek

Mělo by být jasné, odkud přesně se vaše položky vzaly

Ideálně bych doporučil konkrétně popsat fasetový model a každou položku přiřadit k některé fasetě

O fasetách více shortly

FAQ: Musí metoda fungovat?

Důležitá je smysluplnost celého postupu

Pokud jsou jednotlivé kroky dobře zdůvodněné, tak je úplně jedno, jestli metoda nakonec „něco měří“

Podstatné je tedy řemeslo, ne výsledek

FAQ: Kolik lidí musíme otravovat?

Ideálním postupem je provedení power analýzy:

- Např. v G*Power: <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>

Power analýza odhaduje optimální velikost vzorku (mj.) za předpokladu očekávané velikosti efektu, který chceme pozorovat (např. Cohenovo $d > 0,30$)

Ideálně byste tedy měli vybrat analýzy, které plánujete provádět, a pro tu, která má největší nároky na vzorek, provést power analýzu

V tomto kurzu typicky celý postup zjednodušíme tak, že doporučujeme provést power analýzu pro korelaci o velikosti odpovídající vašemu designu

Alternativně lze použít rules-of-thumb pro velikost vzorku ve faktorové analýze

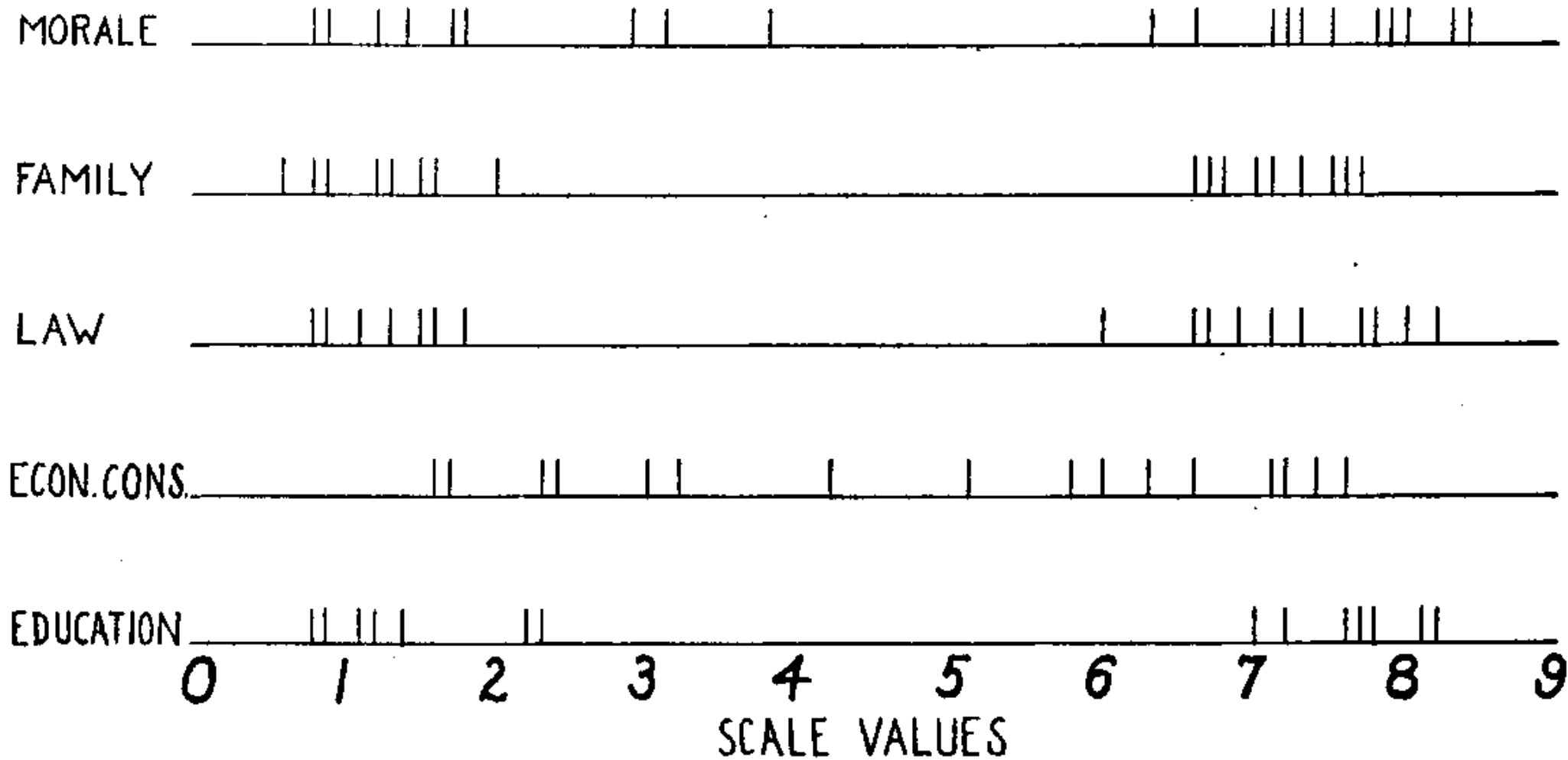


FIGURE 1

DISTRIBUTIONS OF THE SCALE VALUES ON THE MINNESOTA SURVEY OF OPINIONS

Obsahová validita v jednom slidu

1. Položky cílí na měřený konstrukt (např. fluidní inteligenci podle C-H-C)
2. Aby byl vztah položka-konstrukt čistý, musí být splněny dvě podmínky:
 - V položkách **nic nepřebývá** (např. kromě fluidní inteligence neměříme i pracovní tempo)
 - V položkách **nic nechybí** (např. z fluidní intelligence dle C-H-C složené z indukce, dedukce a kvantitativního usuzování neměříme jen dedukci)
3. Této vlastnosti metody říkáme **obsahová validita**

Obsahová validita

racionální posuzování validity

Odpovídá **obsah testu** měřenému konstruktů a účelu měření?

- Můžeme odpovědi na položky (v zamýšlené populaci) považovat za projevy, znaky, symptomy různých úrovní měřeného konstruktů?
- Zachycují položky všechny důležité druhy projevů (symptomy, znaky..)?
- Je vliv příbuzných, podobných či souvisejících konstruktů minimální?

Někdy též **výběrová validita**: jsou položky dobře *vybrané* z univerza možných položek?

Důkazy obsahové validity:

- Posouzení metody/položek experty.
- Racionální argumentace: Design tvorby položek. Soulad s teorií v plné šíři.

Obsahová validita vs. „face validity“:

- **Zjevná validita**: shodu konstruktů s metodou posoudí i laik správně.
- **Zdánlivá validita**: metoda ve skutečnosti měří něco jiného, než si laik myslí.

Položka – indikátor konstruktů

Pozorovatelný projev vysoké/nízké úrovně měřeného konstruktů

- Pozorovatelný druhými – pozorování chování (i pomocí přístrojů, pravitků)
- Pozorovatelný subjektem – výpovědi, odpovědi na otázky

Když je projev pozorován, je to **jednoznačně** známkou přítomnosti určité úrovně měřeného konstruktů

Protože to „jednoznačně“ v psychologii zřídka platí, snažíme se měřit **opakovaně** (tím víc, čím méně jsme si jednoznačností jistí).

- více různými položkami
- v různých časech, podmínkách (omezeno stabilitou samotného konstruktů)

Validita položek

Každá položka obsahuje dvě hlavní rozptylové složky

- =příčiny rozdílů mezi lidmi v pozorování/odpovědích

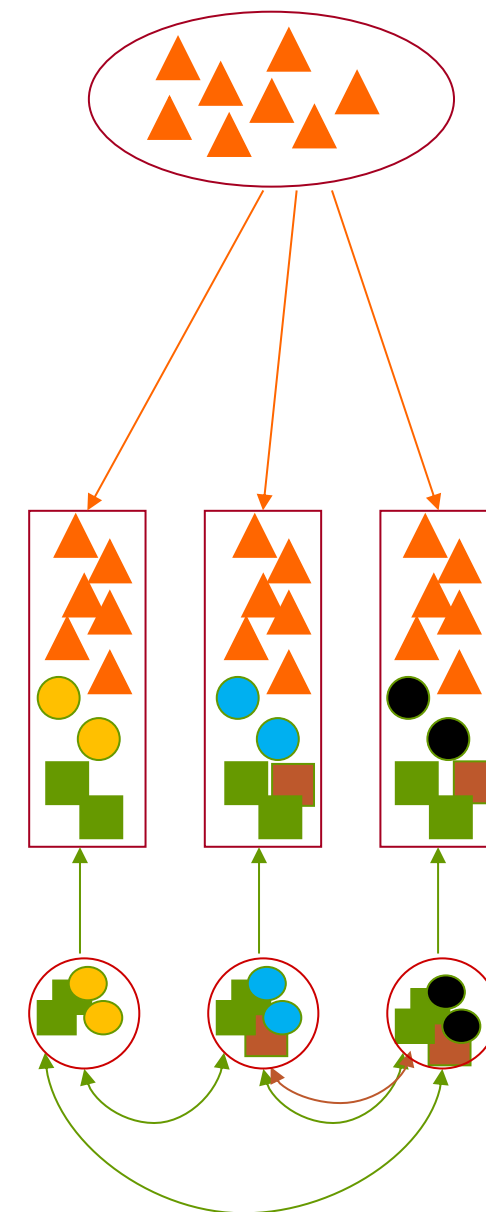
První je **společná** - týká se měřeného konstruktů.

- Pozorování/Odpověď je **způsobena** měřenou latentní proměnnou. ▲

Druhá složka je **specifická** pro danou položku.

- Náhodné vlivy na pozorování/odpovědi v daném čase. ●
- Další systematické vlivy/důvody na pozorování/odpovědi ■
- Specifické faktory

Položky by se měly „překrývat“ ve společném rozptylu ▲ a dobře pokrývat celé významové spektrum měřeného konstruktů a zároveň co nejméně „překrývat“ ve specifických složkách ■■



Validita položek

Reliabilita testu: podíl společného a celkového r.: $\frac{\text{▲}}{\text{▲} + \text{●} + \text{●} + \text{●} + \text{■} + \text{■}}$

Specifický, ale nesdílený rozptyl ●●● snižuje pouze reliabilitu.

Specifický, ale sdílený rozptyl ■■ může reliabilitu nadhodnocovat.

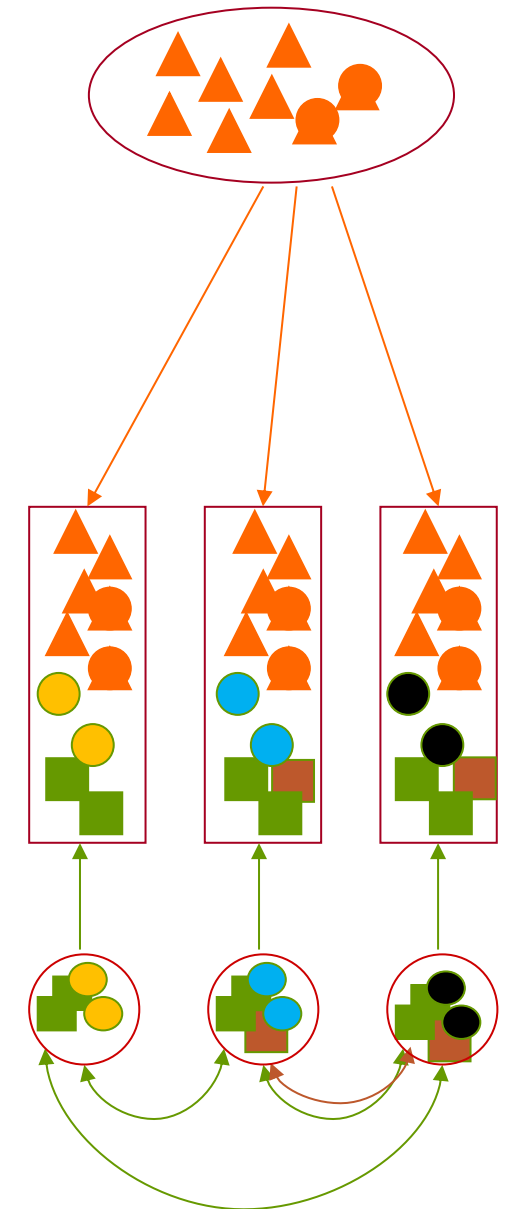
- Např. Cronbachovo alfa, split-half atp. Nelze dost dobře odlišit od ▲.

Sdílený specifický rozptyl ■■ navíc snižuje validitu.

- Systematické zkreslení, které nelze oddělit od společného rozptylu ▲.

Společný rozptyl (konstrukt) může být tvořen více fasetami: ▲ a ●.

Validitu snižuje i to, pokud část společného rozptylu ● není v položkách obsažena.



Tvorba položek

Dobré položky

- Jsme si jistí společným faktorem – *Opravdu měřený konstrukt způsobuje tohle?*
- Známe i specifické faktory (může jich být hodně) a nejlépe je i kontrolujeme.
- Máme představu o tom, jakými náhodnými vlivy může být odpověď zasažena.

Položky dohromady v jedné metodě.

- Měří všechny **společný faktor** (nebo známe jejich „faktorovou strukturu“).
- Pokrývají celou paletu možných projevů (a různých úrovní) konstruktů.
- Specifické faktory nesdílí.
- Nechybí jim žádný z „aspektů“ konstruktů.

Teorie faset:

Dekompozice obsahového univerza (= všech možných projevů konstruktu, teorie)

Teorie faset (Guttman; Shye & al.)

- Základem **dekompozice univerza** určité teorie.
- Mapping definitions – zobrazující definice.
- Série definic obsahu nějakého chování podle tendence, příležitosti a odpovědi...



- **Obsahové univerzum.**
 - Množina všech možných položek a „vážených významů“ daného konstruktu.
- **Položková banka.**
 - Množina reálně vytvořených položek.
- **Pilotní verze metody.**
 - Položky vybrané z položkové banky pro ověření.
- **Další pilotní verze metody.**
- ...
- **Finální verze metody.**
 - Revize metody...

*Nezapomenout na
kognitivní pilotáž!*



Teorie faset:

Dekompozice obsahového univerza (= všech možných projevů konstruktu, teorie)

Teorie faset (Guttman; Shye & al.)

- Základem **dekompozice univerza** určité teorie.
- Mapping definitions – zobrazující definice.
- Série definic obsahu nějakého chování podle tendence, příležitosti a odpovědi...

Podoba a kvalita metody ovlivňují naše úvahy o měřeném konstruktu a tedy i obsahovém univerzu.

Obsahové univerzum.

- Množina všech možných položek a „vážených významů“ daného konstruktu.

Položková banka.

- Množina reálně vytvořených položek.

Pilotní verze metody.

- Položky vybrané z položkové banky pro ověření.

Další pilotní verze metody.

...

Finální verze metody.

- Revize metody...

Nezapomenout na kognitivní pilotáž!

Dekompozice obsahového univerza: fasety modelového zadání – studijní styl

Zadání: „Tendence
k úzkosti ve výkonové
situaci“

Situace:

- Typy výkonových situací
- Zkouška
- Seminárka
- Průběžné studium

Čas:

- dlouhodobost/krátkodobost
- před – v průběhu – po

Složka postoje:

- Kognitivní
- Afektivní
- Behaviorální

Chování

- Odklon
- Různé
- Prokrastinace
- Řešení úkolu
- Perseverace...

Afektivní

- Útlum / agitace
- Konkrétní emoce...

Kognitivní

- Dispozice
- Realizace
- Znalost
- Rychlost...

Ale dále...
3×5×3 (atributu...)

Toto je pouze příklad, fasety je třeba vždy volit adekvátně měřenému atributu.

Teorie faset: Dekompozice obsahového univerza

Pomyslná matice $3 \times 5 \times 3 (\times 5 \times 2 \times 4 \dots)$ kombinací formulací položek.

- Odtud pojem „multifasetový konstrukt“
 - Při přípravě na zkoušku se zaměřuji především na získání obecné orientace v látce.
 - Před zkouškou trpím silnými pocity úzkosti.
 - Během zkoušení mívám problémy s vybavováním.
 - Po zkoušce se obvykle odreaguji ve společnosti.
- Položka je tedy tvořena vždy výběrem prvků z příslušné kombinace faset.
 - Některé kombinace mohou být „nepoužity“.
- Fasety jsou více či méně univerzální – např. složky postoje jsou pro psychologii obecné, ale další fasety závisí na konkrétní teorii, jsou z teorie odvozené.

Více faset neimplikuje více dimenzí/faktorů.

Struktura položky

Položka jako „kognitivní experiment“.

- **Item stem** = podnět = kořen/jádro položky = stimulus.
- **Response** = odpověď = kvantifikace = konkrétní pozorování.
- **Distraktory** = možnosti k výběru (nikoli nutně).
- **Skóre** = bodová hodnota odpovědi.

Jedna položka může implikovat více odpovědí.

Jeden kořen se může dále dělit.

- Instrukce vs. podnětový materiál.

Vztah odpovědi a skóru nemusí být přímý.

- Vážené odpovědi, interpretované odpovědi...
- Projektivní metody (ROR, TAT, MMPI...).

Odpověď na položku není totožná se skórem položky!

- Přiřazení čísel jevům...



Typy a formáty položek 1

Konstrukt vždy implikuje formát položek (a formát ovlivňuje měřený konstrukt).

Z hlediska typů metody:

- *Testy schopností*
- *Osobnostní dotazníky*
- *Silové, rychlostí vs. kapacitní testy*
- *Výkonové vs. introspektivní*
- *Objektivní testy*
- *Projektivní testy*
- V principu: existuje „správná“ či „diagnostická“ odpověď?

Z hlediska formátu odpovědi:

- Implikuje použité statistiky.
- Položky s volnou kvantifikací (čas, x-bodové škály) lze obvykle zpracovat pod CTT s předpokladem normálního rozdělení.
- Položky se „správnou“, „diagnostickou“ odpovědí je nutno v konečném zpracování brát jako dichotomické.
- Dichotomizace ex-post omezuje variabilitu – rozlišovací účinnost!

Z hlediska úrovně měření:

- Nominální – ordinální – intervalové – poměrové
- (Nemusí být shodné s úrovní výsledné škály.)

Typy a formáty položek 2

Doplňování odpovědi: $7+2=...$; Co mají společného obraz a socha?; Za dvacet let budu...

- Odpovědi je nutné skórovat: dichotomické, škála, vícenásobná volba...

Posuzovací škály, někdy též uspořádané kategorie, „stejně se jeví intervaly“ atd.

- Nikdy–zřídka–občas–často–vždy. Souhlasím–spíše souhlasím–spíše nesouhlasím–nesouhlasím.

Vícenásobná volba (multiple-choice): Více než 3 varianty s 1 správnou odpovědí a X distraktory.

Dichotomická položka: Výběr 2 variant (ano/ne, souhlasím/nesouhlasím).

Nucená volba (ipsativní): Jste spíše: tolerantní, nebo náročný?

Mnohonásobná volba: Vlastně série dichotomií...

Spojování (párování).

Řazení: Seřadte výroky podle výstižnosti: Jsem energický... nezávislý ... důvěřivý...

Pozorovací schémata, měření reakčního času, velmi mnoho dalších formátů.

Typické „škály“ v psychologii

Vizuální analogová škála (Hayes and Paterson, 1921)

Thurstonova škála (1928): 3 typy, např. „metoda stejně se jevících intervalů“.

Likertova škála (1932): Metoda sigma vs. zjednodušená metoda

Guttmanova škála (40. léta)

- Rozšíření původní Bogardovy (1925) škály sociální distance.
- Původně deterministický model později rozšířen na stochastický model, základ IRT.

Osgoodův semantický diferenciál (1957)

Doporučujeme kap. 5: Price, L. R. (2016). Psychometric Methods: Theory into Practice. New York: Guilford Press.

Terminologické shrnutí

Atribut: „elementární měřitelný zástupce reálného světa“^a. Měřitelná a měřená vlastnost.

Konstrukt: atribut postulovaný teorií s určitými charakteristickými vlastnostmi a vztahy.

- Může, ale nemusí být vícedimenzionální.

Faktor: Myšlená latentní proměnná, která kauzálně ovlivňuje... něco (třeba odpovědi v dotazníku).

- Úzce spojená s faktorovou analýzou jako nástrojem pro identifikaci různých faktorů v mnohodimenzionálních datech. Může jít ale i o statistický artefakt bez konkrétního významu.

Dimenze: Synonymum faktoru, matematictější terminologie.

Faseta: V tomto kontextu myšlená charakteristika konstruktů, která může, ale nemusí být faktorem.

- Faseta může být sycena i jiným „specifickým faktorem“, který není součástí konstruktů/jiným konstruktem.

Škála/subškála: Vyhodnocovaná součást diagnostické metody, opatřená vlastním skóre a instrumentálně spjatá s touto metodou.

- Pokud není zároveň odhadem určitého faktoru, bude mít zřejmě nízkou validitu.

Viz [diskuzi na Facebooku](#) 😊

^a <https://cs.wikipedia.org/wiki/Atribut>

Seminární úkol: dotazník výšky

Jste ve světě, kde nefungují pravítka.

Potřebujete měřit výšku lidí a jste odkázáni jen na psychodiagnostické metody.

- 4–5členné „výzkumné týmy“.

Jaký konstrukt má dotazník měřit?

Jaké jsou možné fasety obsahového univerza takového konstruktů?

Jaké konstrukty budou do měření intervenovat? Jak jejich vliv minimalizovat?

Jaký formát položek je pro měření vhodný?

K několika fasetám vytvořte několik položek.