

Zadání seminární práce

Výběrová validita

Psychologické škálování

PSYb2590: Základy psychometrie | Seminář 1

24. 2. / 2. 3. 2020

Seminární práce 1: Téma práce

Téma práce: Návrh postupu tvorby, validizace a standardizace 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...).

Skupiny po 4 osobách (výjimky možné po schválení).

Průběžné odevzdání, průběžné konzultace, zpětná vazba výhradně ústně.

Nelze opravit, není minimální počet bodů.

Seminární práce 2: Struktura práce

Struktura odpovídá formátu IMRAD.

Teoretický úvod (max 2 NS): Podrobný popis měřeného konstruktů, použité teorie, související pojmy, účel metody, způsob konstrukce.

- Způsob tvorby položek, nepovinně včetně kognitivního interview.

Metoda (max 2 NS): 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.

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

- Případné vyřazení nefungujících položek a navržení upravené verze instrumentu.

Diskuze (max 2 NS): zhodnocení fungování metody a návrh dalších kroků.

- Včetně kvalitativního zhodnocení svých zkušeností.

Seminární práce 3: Organizace

Skupiny: 4 studenti (jiný počet jen výjimečně).

- V rámci seminárních skupin (napříč jen ve výjimečných případech).
- **Konzultanti:** Hynek, Standa, Adam

Zpětná vazba výhradně ústně.

- Konzultace možné (a doporučené) po domluvě kdykoli!

Termín 1 (8./15. 3.): Vytvoření témat v ISu

- Včetně přihlášení členů týmu.
- Stručný popis tématu.
- **Prezentace tématu** (max. 5 min.) na 2. semináři.

Termín 2 (10. 4.): Odevzdání A

- 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 3 (1./8. 5.): Odevzdání B

- Finální verze celé SP.
- Kompletní datová matice, skript, metoda.
- Prezentace na 6. semináři (4./11. 5.).
 - Ve výjimečných případech lze termín posunout a prezentovat předběžnou verzi.

Hodnocen je jen finální text a prezentace.

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 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.

Jak byly výsledky prezentovány?

Seminární práce 5: Technické

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

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

Prezentujte rozpis na semináři.

Rozpis bude schválen vyučujícím (podle seminárních skupin).

- 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.

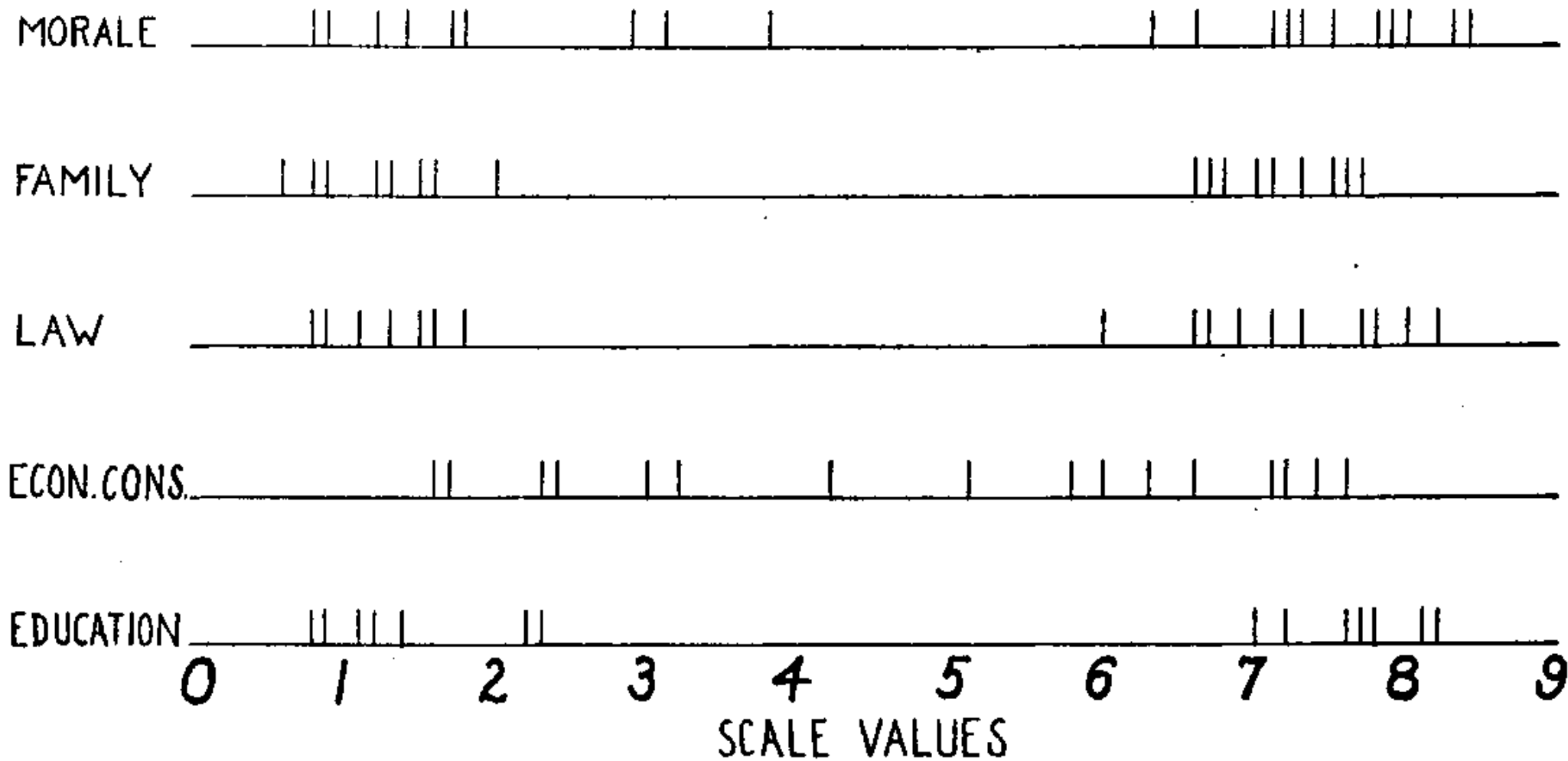


FIGURE 1

DISTRIBUTIONS OF THE SCALE VALUES ON THE MINNESOTA SURVEY OF OPINIONS

Obsahová validita

Odpovídá **obsah testu účelu** metody, resp. měřenému konstrukt?

- Jeden z mnoha způsobů uvažování o validitě testu (viz [FURR]).
- Odpovídá metoda právě měřenému konstrukt, a ne jinému?
- *Reprezentuje* metoda daný konstrukt v celé jeho šíři?
- Operacionalismus: Konstrukt je definován metodou. 😞

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

Důkazy obsahové validity:

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

Obsahová validita vs. „face validity“:

- Co se zdá metoda měřit při posouzení laikem? Odpovídá metoda konstrukt při posouzení laikem?
- **Zjevná validita**: shodu konstrukt 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, pravitel)
- 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

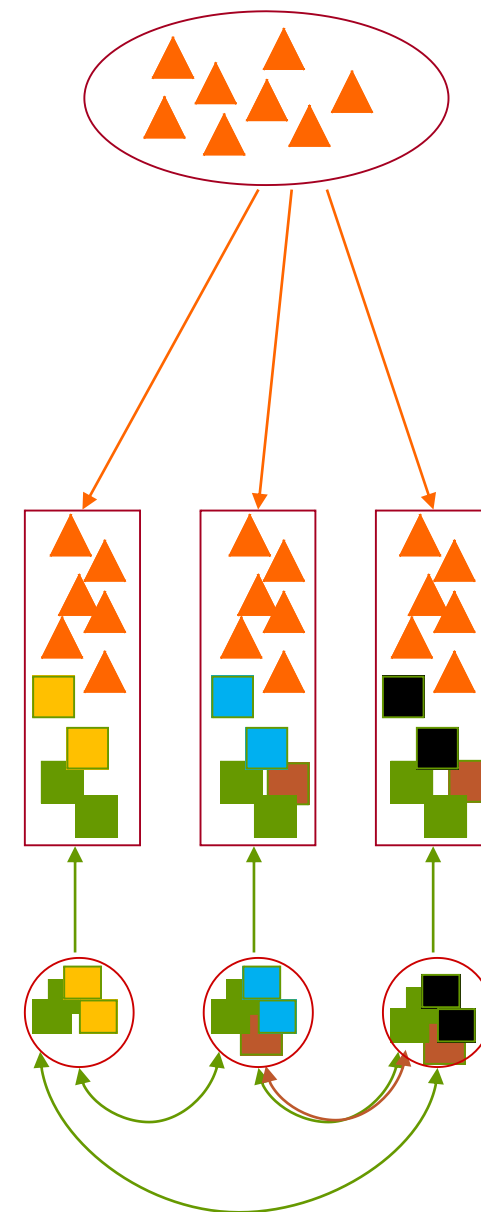
První je **společná** - týká se měřeného konstruktů.

- 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: podíl společného a celkového rozptylu: $\frac{\text{▲}}{\text{▲} + \text{■} + \text{■} + \text{■} + \text{■} + \text{■} + \text{■}}$

Specifický, ale náhodný rozptyl ■ ■ ■ snižuje pouze reliabilitu.

Specifický, ale nenáhodný rozptyl ■ ■ může reliabilitu nadhodnocovat.

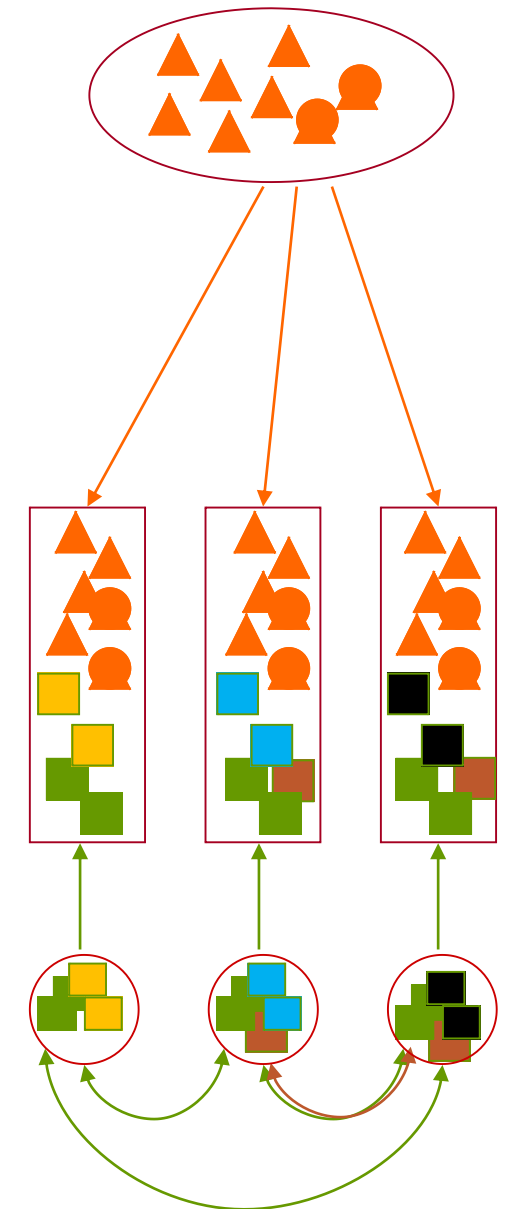
- Např. Cronbachovo alfa, split-half atp.

Nenáhodný 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)
- Máme představu o tom, jakými náhodnými vlivy může být odpověď zasažena

Položky dohromady v jedné metodě

- Měří všechny společný faktor
- Pokrývají **celou paletu možných projevů** (různých úrovní) konstruktů
- Specifické faktory nesdílí

Teorie faset:

Dekompozice obsahového univerza (=všech možných projevů konstruktů, 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...
- $(O - (T \Leftrightarrow S) - R)$



Obsahové univerzum.

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

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...



Dekompozice obsahového univerza: fasety modelového zadání – studijní styl

Zadání: „Tendence
k chování 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...

Atributů dále...
3×5×3 (atributů...)

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 na 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 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 = jádro položky = stimul.

Response = odpověď = kvantifikace = ...kde co.

Distraktory = možnosti k výběru (ne vždy).

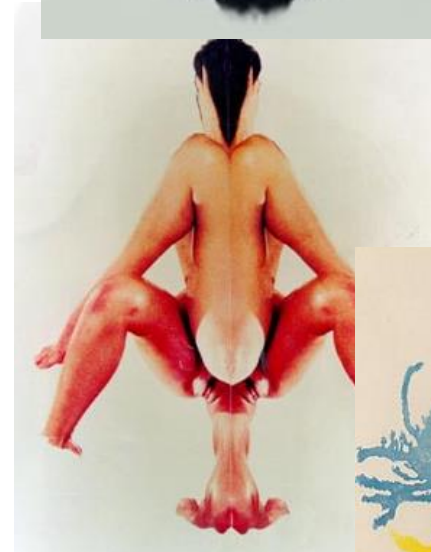
Jedna položka může implikovat více odpovědí.

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...

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.

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? Jde skutečně o „fyzikální“ výšku?

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?

Vytvořte příklad položky ke každé fasetě.

Jaký konstrukt má dotazník měřit?
Jde skutečně o „fyzikální“ výšku?

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?

Vytvořte příklad položky ke každé fasetě.