

Aktuální otázky psychologického testování v ČR

Tomáš Urbánek

Psychologický ústav AV ČR, v.v.i.

tour@psu.cas.cz

Kvalita psychodiagnostických metod

- Nikoli: věc názoru
- Ale: hodnocení na základě argumentů a dat
- dána kvalitou standardizace – hlavně:
 - Validita
 - Reliabilita
 - Normy

Proč recenzovat diagnostické metody?

- Společenský kontext
 - zejména v kontextu PPP
- Zkušenosti z okolních zemí
 - UK, Španělsko, Nizozemí (EFPA)
 - USA, Burosův institut
- Ochrana profese
 - existence mnoha „psychodiagnostických metod“
- Ochrana klienta

Evropský kontext

- Evropská federace psychologických asociací (EFPA)
 - členství za ČR – Unie psychologických asociací ČR

www.efpa.eu

Board of Assessment

(dříve Standing Committee on Tests and Testing)

- zkvalitňování testových metod
- zlepšování praxe testování

- Evropské standardy uživatele testu...
- Evropské certifikáty pro používání ... testů...
- Evropský model recenze testu

Aktivity v ČR

- Pracovní skupina pro testy a testování při ČMPS
- Test fórum (www.testforum.cz)
 - internetový časopis věnovaný psychodiagnostice
- Víτάme zájemce o spolupráci

Evropský model recenze testu

- **Zdroje a předchůdci**

- Steering Committee on Test Standards
(Britská psychologická společnost)
- Komise pro testové záležitosti
(Holandská asociace psychologů)
- Španělský dotazník pro posuzování kvality testu
(Španělská psychologická asociace)

Recenze

- Obsahová stránka

 - soubor závazných kritérií

- Procesní stránka

 - jak recenze probíhá, kdo ji kontroluje

Recenzní řízení

- 4 osoby
 - ale možnost kumulace funkcí
- 2 recenzenti
- Odborný redaktor (consulting editor)
 - kontroluje recenzní řízení
- Vedoucí redaktor (senior editor)
 - zodpovídá za jednotnost používaných kritérií

Obsah recenze

Popisná část

1. Obecný popis nástroje
2. Klasifikace testu
3. Způsob měření a skórování
4. Počítačově generované zprávy
5. Nabídka, podmínky a náklady

Hodnotící část

1. Hodnocení testových materiálů
2. Hodnocení norem, reliability a validity
3. Kvalita počítačově generovaných zpráv
4. Závěrečné hodnocení

Klasifikace

- obsahová oblast (např. všeobecné schopnosti, numerické schopnosti, paměť, motivace, hodnoty atd.)
- oblasti použití (klinická, forenzní, práce a zaměstnání, sport atd.)
- způsob použití (podmínky standardizace a validizace)
- populace
- počty (škál, položek)
- formát položek

Klasifikace

- způsob administrace
 - interaktivní, papír-tužka, počítačový
- způsob odpovídání
- čas potřebný na administraci
 - tzn. příprava, administrace, skórování, analýza, zpětná vazba
- existence různých forem nástroje
 - paralelní formy
 - počítačová / internetová forma

Měření a skórování testu

- procedura skórování
 - jednoduché/komplexní ruční skórování,
počítačové
- získané skóry
 - věcná interpretace
- typy použitých škál
 - percentily nebo standardní skóry apod.

Počítačově generované zprávy

- dostupnost, forma a obsah
- média
 - text, grafika nebo obojí
- komplexnost zprávy
- struktura zprávy
- citlivost vůči kontextu
- klinická/pravděpodobnostní

Počítačově generované zprávy

- modifikovatelnost
- stupeň dokončenosti
- transparence
- styl a tón
- zamýšlený příjemce
 - kvalifikovaný uživatel testu, kvalifikovaný uživatel systému, respondent testu, třetí strana
- nabídka vydavatele testu týkající se zpráv

Nabídka, podmínky a náklady

- Dokumentace, která je součástí testového balíku
- Metody publikování
- Pořizovací náklady
- Opakující se náklady
- Kvalifikace vyžadovaná pro uživatele testu

Hodnocení testových materiálů

- 4 typy zdrojů
 - Manuály
 - Veřejně přístupné informace (např. akademické nebo jiné publikace)
 - Zprávy držené vydavatelem, které nebyly publikovány
 - Materiály, které jsou důvěrné

Vysvětlení principů testu

- Teoretické základy konstruktů
- Procedura vývoje testu
- Důkladnost analýz položek a model analýzy položek
- Vysvětlení obsahové validity
- Souhrn relevantního výzkumu

Další hodnocené materiály

- Adekvátnost dokumentace dostupné uživateli
- Kvalita procedurálních instrukcí
- Kvalita testových materiálů samotných
 - mimo jiné kvalita lokální adaptace testu

Hodnocení norem, reliability a validity

- normy
 - lokální vhodnost, vhodnost pro aplikace, velikosti výběrových souborů, procedury výběru
- validita
 - hledisko konstruktové validity a detaily jejího zjišťování, validita na základě vztahu ke kritériu
- reliabilita
 - spíše více zdrojů údajů o reliabilitě

Hlavní závěr

- Jakýkoli nástroj s jedním nebo více posouzeními [chybí informace] nebo [neadekvátní] týkajícími se atributů považovaných za kritické pro bezpečné používání nástroje by neměl být považován za nástroj, který splňuje minimální standardy.

(doslovná citace Modelu recenze)

Nutné standardy metod

- Validita
- Reliabilita
- Normy

- Kvalita manuálů a příruček

Validita

- existence studií
- kvalita studií
- nejen manuál, ale i relevantní studie vznikající v akademickém prostředí

Reliabilita

- vnitřní konsistence
- stabilita v čase (test-retest)



- INTERVALY SPOLEHLIVOSTI

Interval spolehlivosti

- O hodnotě výsledku je třeba uvažovat jako o intervalu
- Standardní chyba měření

$$SEM = SD \times \text{odmocnina}(1 - \text{reliabilita})$$

- Interval

(hodnota – 1,96 × SEM; hodnota + 1,96 × SEM)

Příklady

- **SD = 15 (např. IQ)**

reliabilita = 0,6

SEM = 9,5

reliabilita = 0,75

SEM = 7,5

reliabilita = 0,9

SEM = 4,7

- **SD = 10 (např. T-skór)**

reliabilita = 0,75

SEM = 5

reliabilita = 0,90

SEM = 3,16

- **SD = 2,5 (steny)**

reliabilita = 0,75

SEM = 1,25

reliabilita = 0,90

SEM = 0,79

Příklady

- **IQ ($r = 0,75$)**

86 \rightarrow (71,3; 100,7)

100 \rightarrow (85,3; 114,7)

- **T-skór ($r = 0,90$)**

55 \rightarrow (48,8; 61,2)

79 \rightarrow (72,8; 85,2)

- **sten ($r = 0,75$)**

4 \rightarrow (1,6; 6,4)

8 \rightarrow (5,6; 10,4)

Normy

- „stáří“ norem
- reprezentativnost populací, u nichž byly normy vytvářeny
- detailnost norem