

METODICHE DI VALIDAZIONE DEI QUESTIONARI E “ROBIN HOOD EFFECT”

Cristina Mazzi

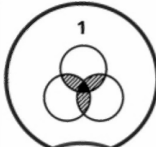
Unità per la Ricerca Clinica

ITEM DEVELOPMENT

SCALE DEVELOPMENT

SCALE EVALUATION

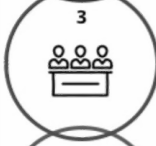
ITEM DEVELOPMENT



IDENTIFICATION OF DOMAIN AND ITEM GENERATION



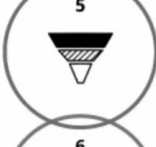
CONTENT VALIDITY



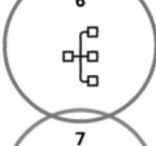
PRE-TESTING OF QUESTIONS



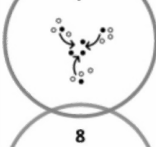
SAMPLING AND SURVEY ADMINISTRATION



ITEM REDUCTION



EXTRACTION OF FACTORS



TESTS OF DIMENSIONALITY



TESTS OF RELIABILITY



TESTS OF VALIDITY

SCALE DEVELOPMENT

SCALE EVALUATION

DOMAIN IDENTIFICATION

TO SPECIFY THE BOUNDARIES OF THE DOMAIN AND FACILITATE ITEM GENERATION.

ITEM GENERATION:

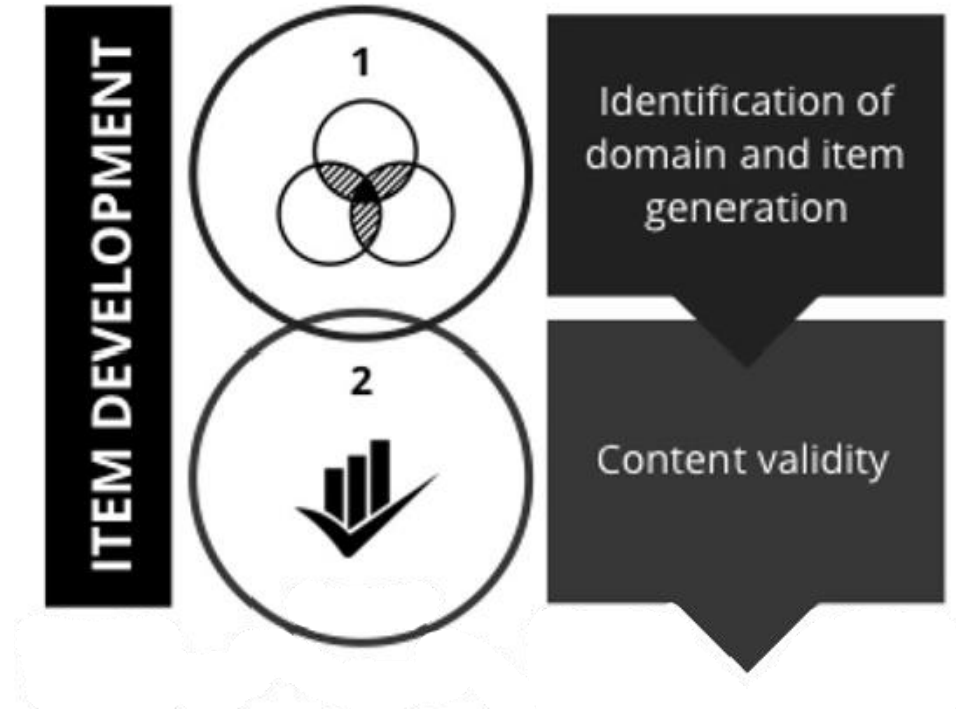
TO IDENTIFY APPROPRIATE QUESTIONS THAT FIT THE IDENTIFIED DOMAIN.

EVALUATION BY EXPERTS

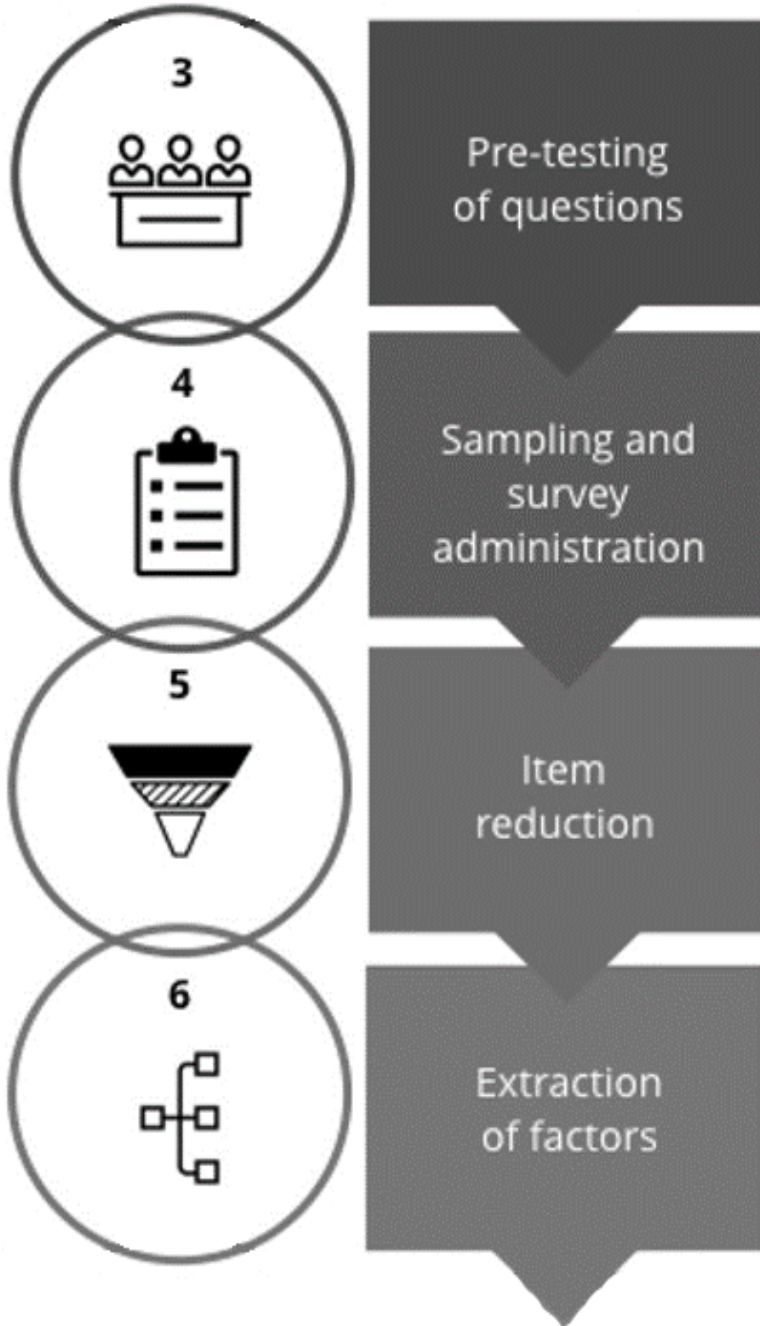
TO EVALUATE EACH OF THE ITEMS CONSTITUTING THE DOMAIN FOR CONTENT RELEVANCE, REPRESENTATIVENESS, AND TECHNICAL QUALITY.

EVALUATION BY TARGET POPULATION

TO EVALUATE EACH ITEM CONSTITUTING THE DOMAIN FOR REPRESENTATIVENESS OF ACTUAL EXPERIENCE FROM TARGET POPULATION.



SCALE DEVELOPMENT



COGNITIVE INTERVIEWS

TO ASSESS THE EXTENT TO WHICH QUESTIONS REFLECT THE DOMAIN OF INTEREST AND THAT ANSWERS PRODUCE VALID MEASUREMENTS.

SURVEY ADMINISTRATION, SAMPLE SIZE, TYPE OF DATA TO USE

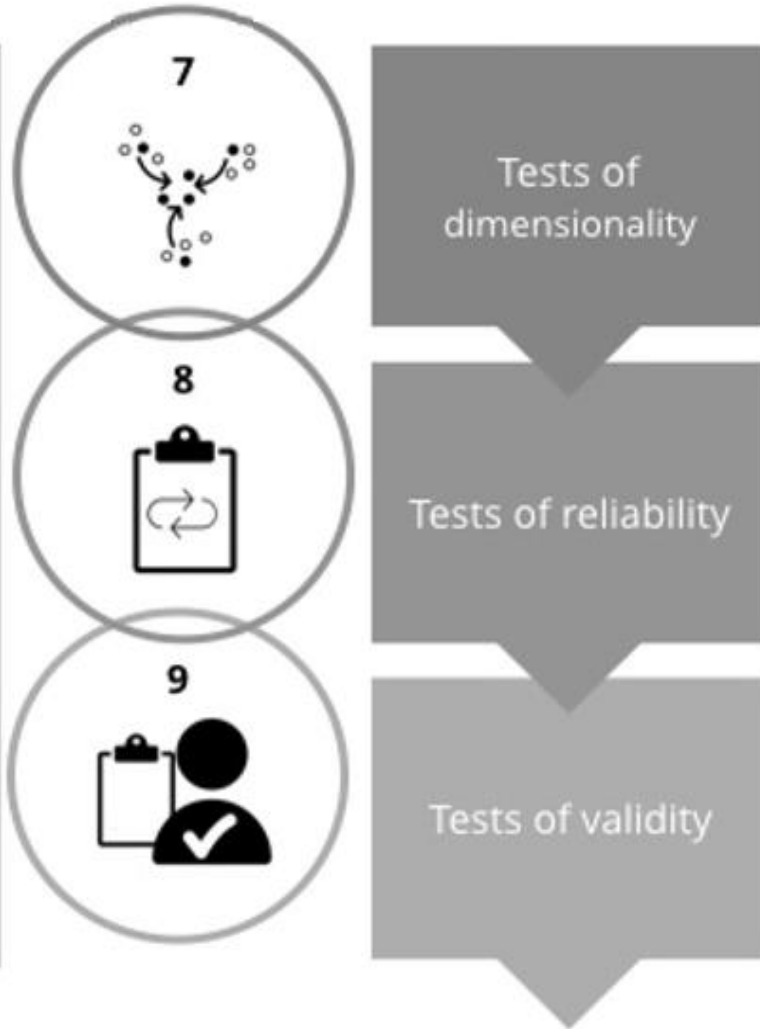
TO COLLECT DATA WITH MINIMUM MEASUREMENT ERRORS, TO ENSURE THE AVAILABILITY OF SUFFICIENT DATA FOR SCALE DEVELOPMENT, TO ENSURE THE AVAILABILITY OF DATA FOR SCALE DEVELOPMENT AND VALIDATION.

ITEM **DIFFICULTY INDEX**, ITEM **DISCRIMINATION TEST**, **INTER-ITEM** AND **ITEM-TOTAL CORRELATIONS**, DELETING OR IMPUTING **MISSING** CASES.

FACTOR ANALYSIS

TO DETERMINE THE OPTIMAL NUMBER OF FACTORS OR DOMAINS THAT FIT A SET OF ITEMS.

SCALE EVALUATION



TEST DIMENSIONALITY

TO VALIDATE WHETHER THE PREVIOUS HYPOTHETICAL STRUCTURE FITS THE ITEMS IN NEW GROUP, CONFIRMATORY FACTOR ANALYSIS.

CALCULATE RELIABILITY STATISTICS

TO ASSESS THE INTERNAL CONSISTENCY OF THE SCALE. I.E., THE DEGREE TO WHICH THE SET OF ITEMS IN THE SCALE CO-VARY, RELATIVE TO THEIR SUM SCORE.

TEST-RETEST RELIABILITY

TO ASSESS THE DEGREE TO WHICH THE PARTICIPANT'S PERFORMANCE IS REPEATABLE; I.E., HOW CONSISTENT THEIR SCORES ARE ACROSS TIME.

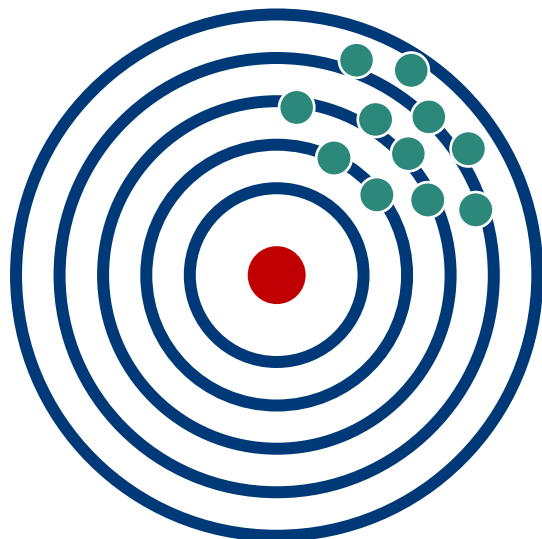
CRITERION VALIDITY

PREDICTIVE AND CONCURRENT VALIDITY.

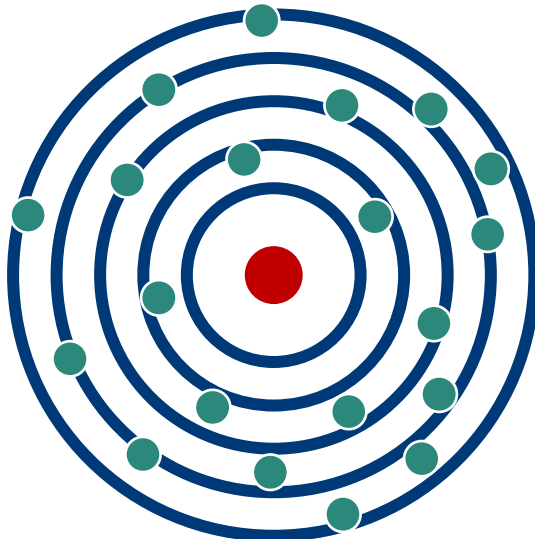
CONSTRUCT VALIDITY

CONVERGENT VALIDITY, DISCRIMINANT VALIDITY, DIFFERENTIATION BY "KNOWN GROUPS".

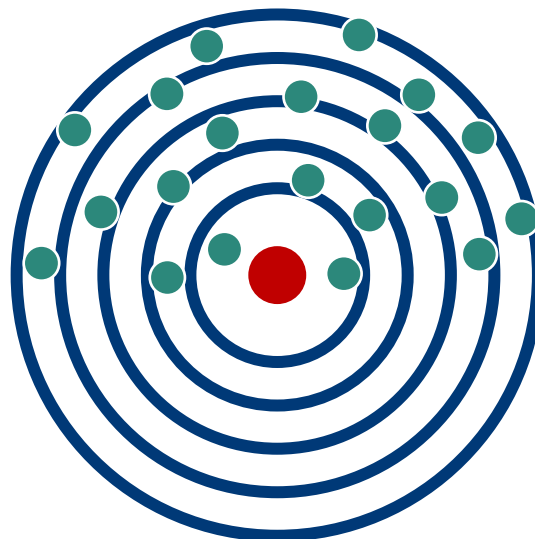
ABILITY TO DETECT CHANGE



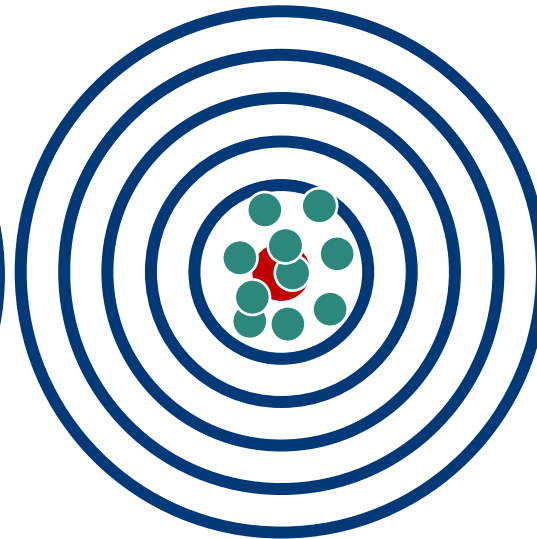
**RELIABLE
NOT VALID**



**NOT RELIABLE
VALID**



**NOT RELIABLE
NOR VALID**



**BOTH RELIABLE
AND VALID**

THE MINIMAL INTERESTING DIFFERENCE (M.I.D.)

- IT'S EASILY UNDERSTOOD BY CLINICIANS AS A KEY CONCEPT IN THE INTERPRETABILITY OF PRO SCORES;
- WILL INFORM JUDGMENTS ABOUT THE SUCCESS-FULNESS OF AN INTERVENTION;
- AN INDIVIDUAL PATIENT ACHIEVING THE SCORE EQUAL OR GREATER THAN THE MID MIGHT BE CONSIDERED A BENEFICIARY OF THE INTERVENTION, WHAT WOULD LEAD TO THE DEFINITION OF A **RESPONDER**.

MID FROM ANCHOR-BASED METHODS

- THE **RESPONDER** DEFINITION MAY BE DETERMINED EMPIRICALLY USING ANCHOR-BASED METHODS.
- ANCHOR-BASED METHODS EXPLORE THE ASSOCIATIONS BETWEEN THE TARGETED CONCEPT OF THE PRO INSTRUMENT AND THE CONCEPT MEASURED BY THE ANCHORS.
- ANCHORS ARE WHAT THE PATIENT **SAYS*** WOULD BE THE SMALLEST DIFFERENCE THAT “IS PERCEIVED TO BE BENEFICIAL”.

* *I.E. PATIENT RATINGS OF CHANGE IN DIFFERENT SITUATIONS*

MID FROM DISTRIBUTION-BASED METHODS

- DISTRIBUTION-BASED METHODS INTERPRET RESULTS IN TERMS OF THE RELATION BETWEEN THE MAGNITUDE OF EFFECT AND SOME MEASURE OF VARIABILITY IN RESULTS;
- THE MOST WIDELY USED DISTRIBUTION-BASED METHOD IS THE BETWEEN-PERSON STANDARD DEVIATION, OFTEN REFERRED TO AS **EFFECT SIZE**.

TO SUM UP

FOR SCALE TO BE USED AS EVIDENCE OF TREATMENT EFFECT, THE INSTRUMENTS MUST BE PROPERLY **DEVELOPED** AND **VALIDATED**. THE PROCESS IS COMPLICATED AND CAN VARY DEPENDING ON THE STATUS OF THE SCALE, THE TARGET POPULATION AND THE RESOURCES AVAILABLE.

VALID INSTRUMENTS ARE **RESPONSIVE** TO CHANGES IN HEALTH STATUS. RESPONSIVENESS IS ESSENTIAL TO DETERMINE THE **MID** AND TO SUPPORT THE USE OF PATIENT-REPORTED OUTCOMES AS MEASURABLE ENDPOINTS.

LETTURE DI RIFERIMENTO

- Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health*. 2018 Jun 11;6:149. doi: 10.3389/fpubh.2018.00149. PMID: 29942800; PMCID: PMC6004510.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008 Feb;61(2):102-9. doi: 10.1016/j.jclinepi.2007.03.012. Epub 2007 Aug 3. PMID: 18177782.

Grazie per l'attenzione.