



DATA GATHERING METHOD

Dr. Sevil Hakimi
Msm. PhD.

THE NECESSITY OF INSTRUMENTS DEVELOPMENT

- Good researches in health sciences depends on **good measurement**. The foundation of all rigorous research designs is the use of measurement tools that are psychometrically sound



- Who is gather the data?
- In which condition?
- *Associations of gestational weight gain with maternal body mass index, waist circumference, and blood pressure measured 16 y after pregnancy: the Avon Longitudinal Study of Parents and Children (ALSPAC)*



- Six *trained research midwives* abstracted data from obstetric medical records. No between-midwife variation in mean values of abstracted data and repeat data entry checks demonstrated error rates consistently ,



Weight and height were measured while the subjects were *wearing light clothing and no shoes*. Weight was measured to the nearest *0.1 kg* by using Tanita scales (Tanita Europe BV, Amsterdam, Netherlands). Height was measured to the nearest *0.1 cm* by using a Harpenden stadiometer (Holtain Ltd, Crymych, United Kingdom). WC was measured to the nearest *1 mm at the midpoint between the lower ribs and the pelvic bone with a flexible tape*. *Seated* BP was measured by using a Dinamap 9301 Vital Signs Monitor (Morton Medical, London, United Kingdom). Two readings of SBP and DBP were recorded, *and the mean is used here*.



In a quantitative study any number of strategies can be adopted when collecting data and these can *include interviews, questionnaires, attitude scales or observational tools. Questionnaires are the most commonly used data gathering instruments and* consist mainly of closed questions with a choice of fixed answers.



*When data was gathered
by questionnaire*



```
graph TD; A[When data was gathered by questionnaire] --> B[Is it standard questionnaire? Or....]; B --> C[Is it developed by a researcher?];
```

*Is it standard
questionnaire? Or....*

*Is it developed by a
researcher?*



- If a *previously designed* instrument is selected the researcher should clearly establish that chosen instrument is the *most appropriate*. Previously designed instruments are often in the form of standardized tests or scales that have been developed for the purpose of measuring a range of views, perceptions, attitudes, opinions or even abilities.



When we should develop a questionnaire (culture, sensitive issue, religion)

Advantages: culture sensitive

Disadvantages: Time and money consuming



ITEM DEVELOPMENT???

- ✓ Statements should be *simple and as short* as possible
- ✓ The language used should be *familiar* to target respondents
- ✓ Items should address only a *single issue, double barreled* items should be not be used
- ✓ Leading questions should be avoided as they may bias responses
- ✓ The issue of *negatively worded, reversely scored items* had stimulated much discussion.
- ✓ Proponents: It may reduce response set bias
- ✓ Others: It may intersperse psychometric properties



WHAT IS THE VALIDITY?

- validity is defined as the ability of the instrument to measure the attributes of the construct under study. It Reflects the real meaning.



TYPES OF VALIDITY

- 1. Face validity
- 2. Content validity
- 3. Pragmatic (criterion) validity
 - A. Concurrent validity
 - B. Predictive validity
- 4. Construct validity
 - A. Convergent validity
 - B. Discriminant validity



FACE VALIDITY

- Face validity means that the instrument looks, on the face of it, as if it measures the construct of interest. It is the easiest way to claim support for construct validity and, as a result, is frequently reported in the literature. Face validity is a subjective assessment, so it is the weakest form of validity.
- Assessment of face validity in papers



- It is not a form of validity in the true sense of indicating that the tool measures the construct of interest; however, it does provide insight into how potential participants might interpret and respond to the items. Investigators seek experts or lay people to review the instrument for grammar, syntax, organization, appropriateness, and confirmation that it appears to flow logically.



HOW CAN ASSESS THE FACE VALIDITY

- Qualitative face validity
- Quantitative face validity (Lay experts)



MEASUREMENT OF ITEM SCORE

- In this method, the impact of each item is determined from the proportion of patients who identified it as important, and the mean importance score attributed to the item (impact score = frequency × importance).
 - Those items associated with an impact score 1.5 (which corresponds to a mean frequency of 50% and a mean importance of 3 on the 5-point Likert scale) were retained for further analysis.
- 

An inventory for assessment of female workers' health promotion behavior based on the integrated model of planned behavior and self-efficacy.

- *Face validity*

To assess face validity, experts and 10 female workers assessed the clarity and fluency of statements and revised some of the items.



LINGUISTIC VALIDATION

- The aim of Linguistic Validation of a questionnaire in a specific language is to produce a translated version in a foreign language which is conceptually equivalent to the original version, as well as clear and easy to understand



THE LINGUISTIC VALIDATION SHOULD CONSIST IN AT LEAST 3 STEPS

- Forward translation (includes the production reconciliation version)
- Backward translation
- Patient testing



FORWARD TRANSLATION

- Production of 2 forward versions: each of translators will produce an *independent forward translation (relevant)* of the original items and instructions and response choices.
- Production a pooled version, version 1: both translators must discuss the translation and agree on the reconciliation version.



- In the case of interpretation problems of the original questionnaire the author can be contacted and according to his explanation of the particular problem the first version in the target language might be modified, leading to the production of a second intermediary forward version



BACKWARD TRANSLATION STEP

- It is the translation of the first version of the questionnaire into source language.
- Native speaker of the source language bilingual in target language.



- The translator will translate the first version of the questionnaire produced in phase 1 back into the source language. He should no access to the original version of the questionnaire.



COMPARISON OF THE BACKWARD VERSION

WITH THE ORIGINAL VERSION

- Done by the local project manager during the meeting with the backward translator in order to detect any misunderstanding, mistranslation or inaccuracies in the intermediary forward version of the questionnaire.
- Production of a report in English on the issues which were discussed item- by- item and how the final decision were made (including English equivalents if items and target language expressions discussed).



PATIENT TESTING

- To test the translated questionnaire on the patients to determine whether the language used is simple and appropriate.



- The second version of questionnaire (obtained after 2 phase) *has to be tested on a panel of a minimum of 5 patients with the condition investigated in the questionnaire. The patients should all be native speakers of the target language.*
- The comprehension test should be perform through face to face interviews during which the interviewer should inquire weather the patient had any difficulty in understanding the questionnaire



- And check the patient`s interpretation of all items. In case of any problem the interviewer may propose or test alternatives of translations (if this problem had been anticipated), or ask the person to propose alternatives.
- A report on the interviews should be produced: it should outline the number of subjects, their age, The time it took to complete the questionnaire, the difficulties encounter, the solution suggested, and retained and how the third version of the questionnaire was produced.



- It should be mentioned whether the questionnaire is self administer ? (critique)

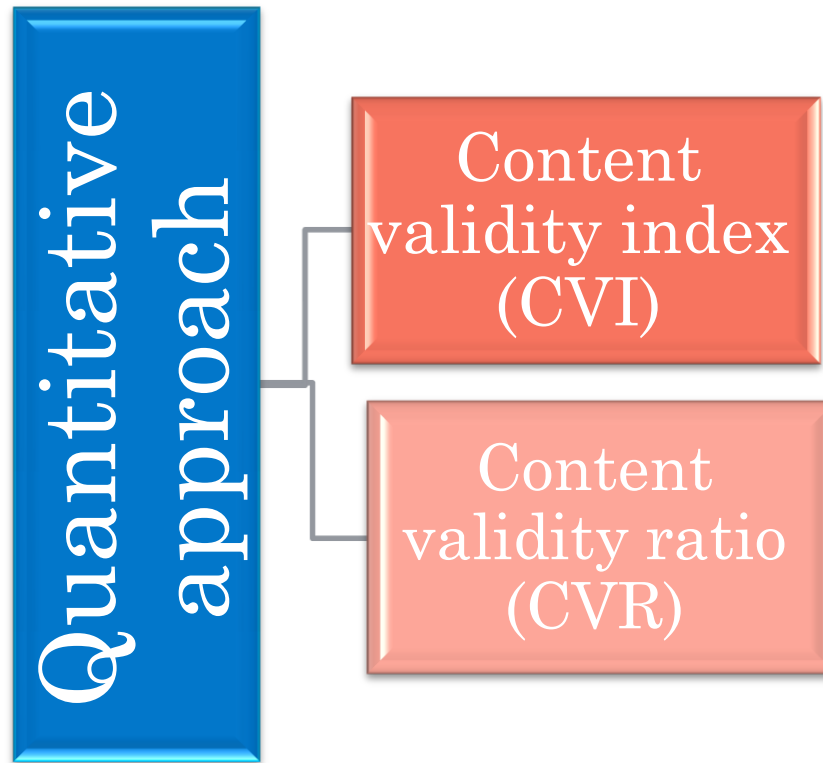


CONTENT VALIDITY

Content validity refers to the degree to which the content of the items reflects the content domain of interest (APA, 1954)

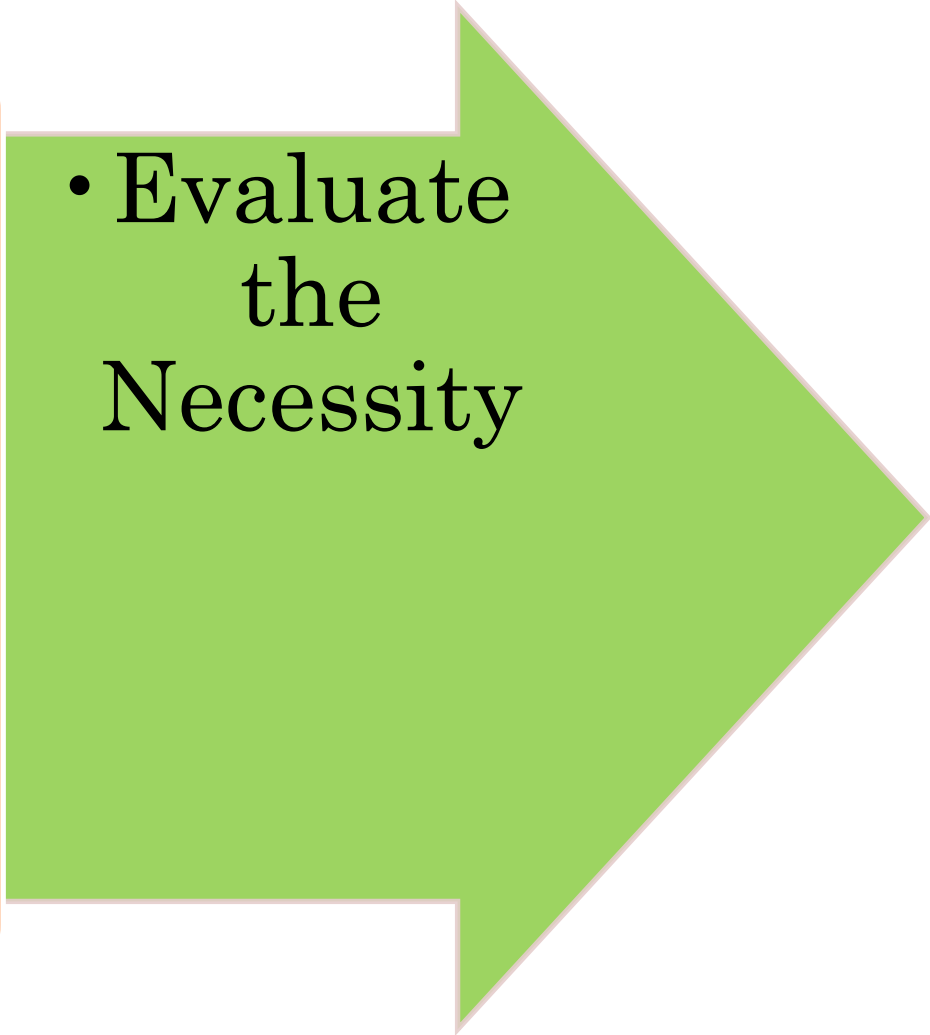
Is the content about what we say the test is about?







CVR

- 
- Evaluate the Necessity



CVI=0.79



Relevancy •



Clarity •



Simplicity •



Table 1. Two Methods of Calculating the Content Validity Ratio (CVR) and the Content Validity Index (CVI)

Author	Lawshe (1975)	Lynn (1986)
Rating scale	Scale used for rating items: 013 Not necessaryUsefulEssential	Scale used for rating items: 1234 IrrelevantExtremely relevant
Calculations	To calculate the CVR (a score for individual scale items): $CVR = \frac{n e - N/2}{N/2}$ Note. n e=The number of experts who rated an item as “essential.” N=the total number of experts. Example: 8 of 10 experts rated an item as essential. The CVR would be (8 – 5/5)=.60. The CVI is the mean CVR for all retained items.	The CVI for each scale item is the proportion of experts who rate the item as a 3 or 4 on a 4-point scale. Example: 4 of 6 content experts rated an item as relevant (3 or 4), the CVI would be: 4/6=.67. This item would not meet the .83 level of endorsement required to establish content validity using a panel of 6 experts at the .05 level of significance. Therefore, it would be dropped. The CVI for the entire scale is the proportion of the total number of items deemed content valid. Example: 77 of 80 items were deemed content valid. The CVI would be: 77/80=.96
Acceptable range	Depends on the number of reviewers	Depends on the number of reviewers

MINIMUM VALUE OF CVR (P=0.05)

No of panelists	Main value
5	0.99
6	0.99
7	0.99
8	0.75
9	0.78
10	0.62
11	0.59
12	0.56
15	0.49
20	0.42
25	0.37
30	0.33
35	0.31
40	0.29



- Number of experts
- Who is accountable for final decision? principle investigator or others?



RELIABILITY

- When someone says you are a 'reliable' person, what do they really mean?
- Are you a reliable person? ☺



STABILITY RELIABILITY

Test-retest ○

SAME TEST – DIFFERENT TIMES ○

- Testing phenomenon at two different times;
- The degree to which the two measurements of “Same Ting,” using same measure, are related to one another



Intraclass correlation coefficient

همبستگی درون رده ای

How can we estimate?

A repeatability study required to help establish and quantify reproducibility, and thus provide an indication of the '**test-retest**' **reliability of a measurement.**

The measurements could be from two people (or two types of equipment), or the same person on two, or more, occasions.



- Prevalence, Risk Factors, and Impact on Health Related Quality of Life of Overactive Bladder in China
- greater symptom severity. Granted authorization from Prof. Yukio Homma, Chinese Urological Association (CUA) translated OABSS into Chinese following the “Cross-cultural adaptation of health-related quality of life measures,” and assessed its test–retest reliability and correlation with other OAB rating tools among Chinese OAB patients. It proved that the Chinese version of OABSS had good test–retest reliability (Internal correlation coefficient 0.5902–0.9274),





1 : Participant

1

	Participant	Therapist1	Therapist2	Therapist1_2	var	var	var	var	var
1	1	17.13	18.78	16.78					
2	2	16.08	17.42	16.31					
3	3	10.91	10.73	10.60					
4	4	14.96	15.65	14.70					
5	5	13.00	11.52	12.63					
6	6								
7	7								
8	8								
9	9								
10	10								
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									
27									
28									

Reliability Analysis

Items:

- Participant
- Therapist 1 second readin...
- Therapist 1 first reading [T...
- Therapist 2 reading [Thera...

Statistics...

Reliability Analysis: Statistics

Descriptives for

- ☒ Item
- ☐ Scale
- ☐ Scale if item deleted

Inter-Item

- ☐ Correlations
- ☐ Covariances

Summaries

- ☐ Means
- ☐ Variances
- ☐ Covariances
- ☐ Correlations
- ☐ Hotelling's T-square
- ☒ Intraclass correlation coefficient

ANOVA Table

- ☒ None
- ☐ F test
- ☐ Friedman chi-square
- ☐ Cochran chi-square
- ☐ Tukey's test of additivity

Model: Two-Way Mixed

Type: Consistency

Confidence interval: 95 %

Test value: 0

Continue Cancel Help

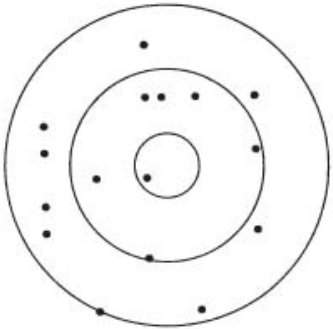
Data View

Variable View

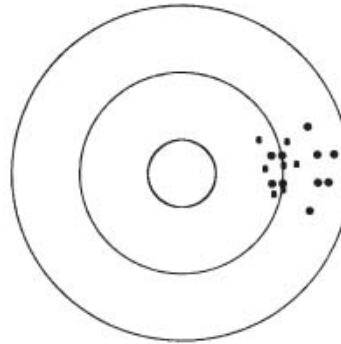
- Interrater
- Kappa cohen



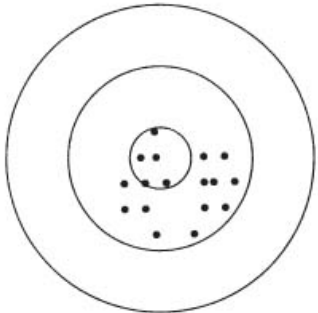
Think in terms of
'the purpose of tests'
and the *'consistency'*
with which the
purpose is
fulfilled/met



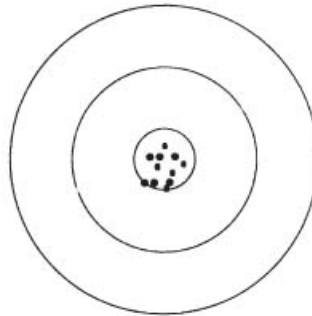
**Neither Valid
nor Reliable**



**Reliable
but not
Valid**



**Fairly Valid
but
not very
Reliable**



**Valid &
Reliable**

- Second hand data
- Qualitative data
- Credibility: Maximum of variance, member checking
- Transferability: Writing the methodology by detail
- Confirmability: peer checking

