A PARADIGM FOR DEVELOPING BETTER MEASURES OF MARKETING CONSTRUCTS

Gilber A. Churchill (1979)

Introduced by Azra Dedic

in the course of “Measurement in Business Research”

Introduction

- Measurements are “rules for assigning numbers to objects to represent qualities of attributes”.
- What is measured? ATTRIBUTES of objects. NOT objects themselves.
- What is the goal? To have measures that are

RELIABLE and VALID

Construct

- Construct, e.g. customer satisfaction
- True level of satisfaction (True score) denoted Xt
- Observed score X0, rarely similar to Xt due to differences in stable characteristics, transient personal factors, situational factors etc.

Validity and Reliability

X0 = Xt + Xs + Xr, where

- Xs – systematic source of error
- Xr – random source of error

Validity: X0 = Xt

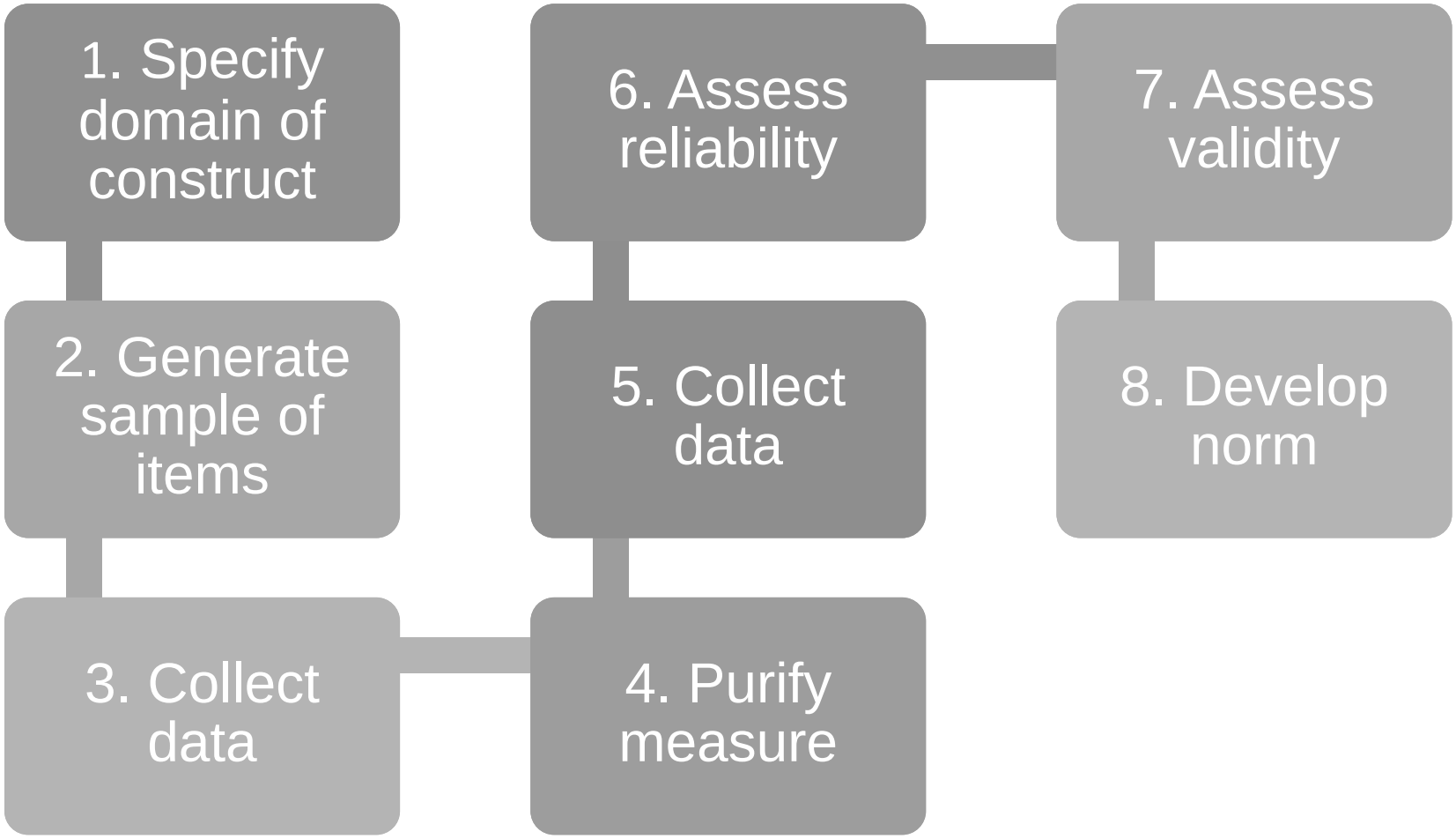
Perfect reliability: Xr = 0

- Validity => Reliability
- Reliability is necessary but not sufficient for Validity

Validity and Reliability (2)

- Objective: find X0 that approximate Xt
- Measures are inferences, their “goodness” is supported by the evidence, that is based on reliability or validity index
- Reliability forms: split-half, test-retest etc.
- Validity forms: face, content, predictive, concurrent, pragmatic, construct, convergent, discriminant.

Procedure for developing measures



Specify domain of the construct

- Exactly defining what is included in the definition and what is excluded
- Consulting the literature
- Widely varying definitions should be avoided
- Example: to measure customer satisfaction
  - Measure both expectations at the time of purchase and reactions at some time after the purchase
    - Expectations: cost, durability, quality, operating performance, aesthetic features, sales assistance, advertising, availability of competitor’s alternatives,

Generate sample of items

- Literature searches
- Experience surveys
- Insight-stimulating examples
- Critical incidents and focus groups

Purify the measure

- Domain sampling model: purpose of any particular measurement is to estimate the score that would be obtained if all the items in the domain were used
- In practice use of SAMPLE of items
- Measurement error due to inadequate sampling
- Correlation matrix of the items in the domain
  - Average correlation in the matrix
  - Dispersion of the correlation about the average
- Assumption: all items, “if they belong to the domain of the concept, have an equal amount of common core”

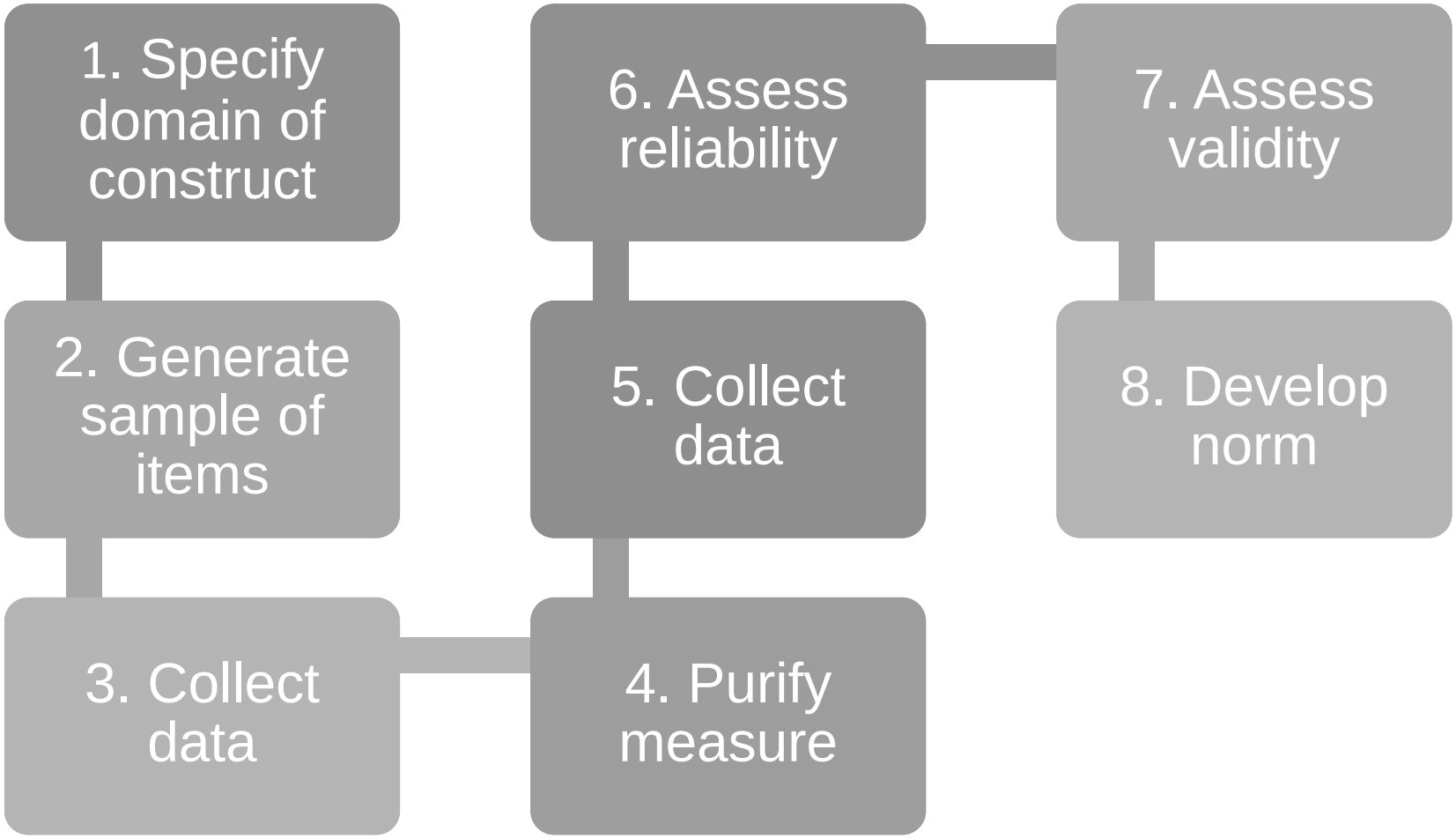
Coefficient Alpha

- Measure of internal consistency of a set of items
- Low coefficient alpha indicates that the sample of items badly describes the construct which motivated the measure
- Procedure by low alpha: some items should be eliminated.
  - Calculate correlation of each item with total score
  - Plot the correlations by decreasing order of magnitude
  - Items with correlations near zero should be eliminated
  - Items of substantial drop in the item-to-total correlations also deleted
- Mistake to do split-half reliability

Purify the measure (2)

- Desirable outcome: high coefficient alpha, dimensions agree with the conceptualized. Then, additional testing with a new sample of data.
- Second outcome: Factor analysis suggests the overlapping dimensions. Items with pure loadings on the new factor are retained, new alpha calculated.
- Non-desirable outcome: alpha coefficient is low and restructuring of items forming each dimension is unproductive. Loop back to 1. and 2.

Procedure for developing measures



Assess reliability with new data

- Source of error within a test or measure is the sampling of items.
- Coefficient alpha is the basic statistic for determining the reliability of a measure based on internal consistency, but it does not estimate errors external to the instrument.
- Collect additional data to rule our the chance possibility of previous findings
- Do not use test-retest reliability

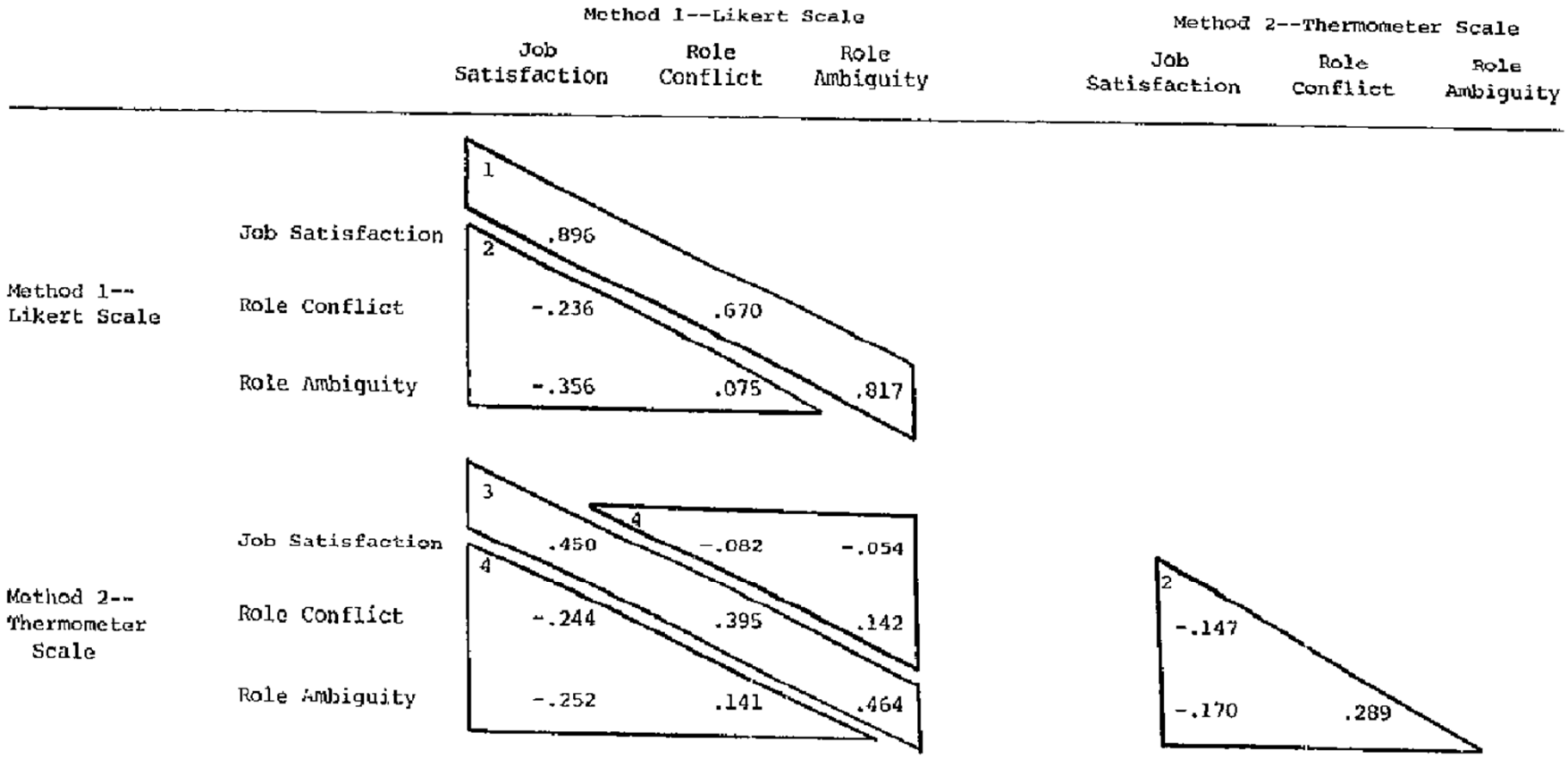
Assess Construct Validity

- Face or content valid measure has an appropriate sample
- To establish construct validity
  - Determine the extent to which the measure correlates with other measures designed to measure the same thing
  - Determine whether the measure behaves as expected

Correlations with Other Measures

- Any construct or trait should be measurable by at least two different methods
- Convergent validity – extent to which it correlates highly with other methods designed to measure the same construct
- Discriminant validity – the extent to which a measure a novel
- Multitrait-multimethod matrix: methods and traits generating it should be as independent as possible

Multitrait-multimethod matrix



		Method 1--Likert Scale			Method 2--Thermometer Scale		
		Job Satisfaction	Role Conflict	Role Ambiguity	Job Satisfaction	Role Conflict	Role Ambiguity
Method 1-- Likert Scale	Job Satisfaction	(1) .896			(2)		
	Role Conflict	-.236	.670				
	Role Ambiguity	-.356	.075	.817			
Method 2-- Thermometer Scale	Job Satisfaction	(3) .450	-.082	-.054	(4)		
	Role Conflict	-.244	.395	.142	-.147		
	Role Ambiguity	-.252	.141	.464	-.170	.289	

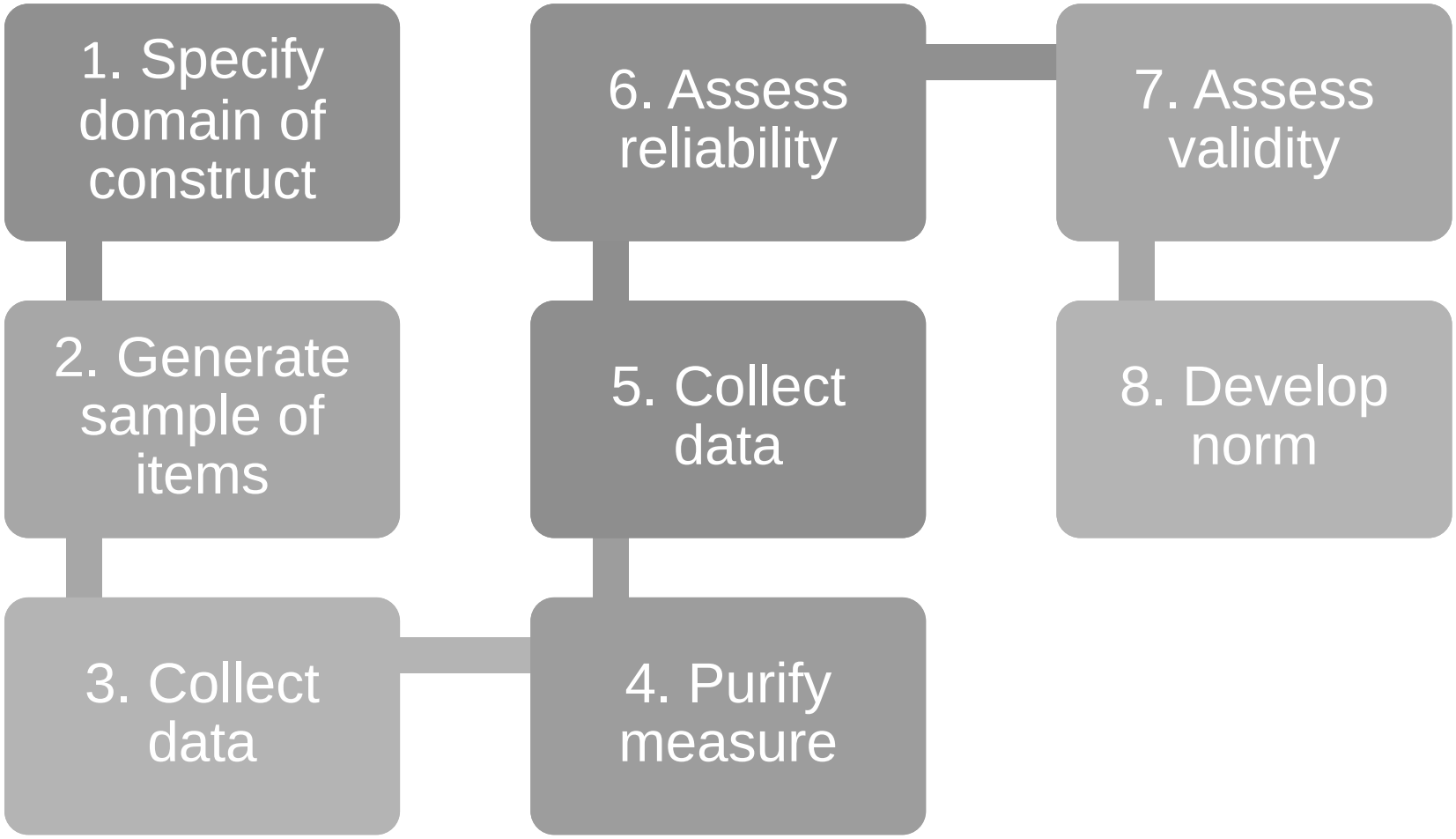
Table 1 of .450, .395 and .464 are all significant at the .01 level.

relation coefficient such as the coefficient of con-

Does the measure behave as expected?

- Internal consistency is insufficient condition for construct validity
- Assess whether scale correctly predicts criterion measure (criterion validity)
  - The constructs job satisfaction (A) and likelihood of quitting the job (B) are related.
  - The scale X provides a measure of A.
  - Y provides a measure of B.
  - X and Y correlate positively.
- Establish the validity by relating the measure to a number of other constructs and not only one

Procedure for developing measures



Developing Norms

- Assessing the position of the individual on the characteristic is to compare the person's score with the scores achieved by other people
- Norm quality depends on both the number of cases on which the average is based and their representativeness