Developing Valid & Reliable Scales

An intro to the mind reader’s toolbox

Heather Wojton, Ph.D.

Institute for Defense Analyses

April 2019

Surveys are a form of psychological measurement

Nearly **everyone** in industrialized countries is affected by psychological measurement at some point in their lives.

- Standardized knowledge and intelligence tests in education
- Personality tests in the hiring process
- Political polls
- Death penalty

The Department of Defense engages in psychological measurement to:

- Place military personnel into specialties
- Evaluate the mental health of military personnel
- Evaluate the quality of human-system interaction
- Identify factors that affect crime rates on military bases

You **must** understand the properties that affect psychological measurement to develop quality surveys

Objectives

1. Identify psychological measurement's goals and challenges
2. Understand basic measurement concepts and how they apply to psychological measurement
3. Understand scale development basics
4. Understand the importance of reliability and validity testing scales, factors that affect reliability and validity, and how to conduct reliability and validity testing

Psychologists use instruments to measure behavior

Because they are interested in the behavior…

Example: facial expressions or error rates

OR to assess unobservable psychological attributes

Example: workload or memory

Method for assessing psychological attributes:

1. Identify a behavior believed to represents a specific psychological attribute, state, or process

2. Measure the behavior

3. Interpret the measurement in terms of the underlying psychological attribute, state, or process

Surveys sample behavior that is sensitive to the underlying psychological attribute

How do you measure working memory?

link	rule	horizon
win	slim	timetable
add	opportunity	elephant
cup	platinum	cathedral
list	livelihood	computer
knot	overestimate	mouse
spade	regiment	pencil
watch	government	elevator

Notice: we made an inference from an observable behavior to an unobservable psychological attribute

To be valid, the behavior must be **theoretically linked** to the psychological attribute.

Behavior must be sampled systematically

Typically, samples of behavior are collected to:

1. Compare the behavior of 2 or more people at the same point in time
2. Compare the behavior of the same people at different points in time
3. Compare the behavior of people under different conditions

Psychometrics evaluates the attributes of surveys

Concerned with 3 types of information

1. The type of data generated by the measurement instrument
   - Surveys, for example, generate scores or ratings
2. The **reliability** of the data
3. The **validity** of the data

Psychological measurement is challenging

Psychological phenomenon are complex

Participant reactivity & experimenter bias

Score sensitivity

□ No

□ Yes

VS.

Strongly Disagree				Strongly Agree
1	2	3	4	5

Course Objectives

1. Identify psychological measurement’s goals and challenges
2. Understand basic measurement concepts and how they apply to psychological measurement
3. Understand scale development basics
4. Understand the importance of reliability and validity testing scales, factors that affect reliability and validity, and how to conduct reliability and validity testing

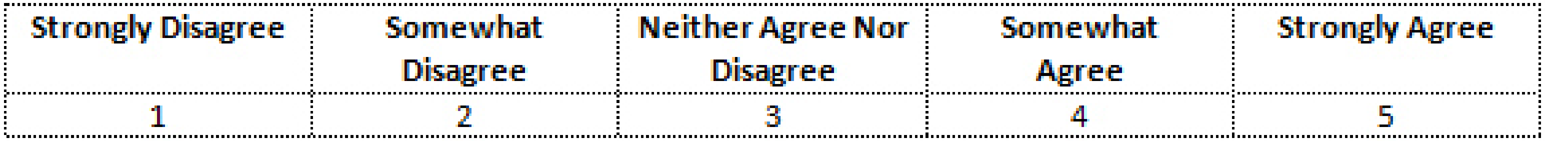
Measurement is the assignment of numerals to objects or events according to rules

Basic Measurement Concepts

Scaling is the process by which numbers are assigned to represent the quantities of psychological attributes

To appreciate the concept of scaling you must understand:

1. The meaning of numerals
2. How numerals can be used to represent psychological attributes
3. Problems associated with trying to connect numerals and psychological attributes



Strongly Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Strongly Agree
1	2	3	4	5

Numerical Properties

The properties of *identity*, *order*, and *quantity* reflect key differences in how numbers represent psychological attributes

Identity: numerals serve strictly as labels of categories, reflecting differences in kind rather than amount

Property: Conveys information about the relative amount of an attribute that people possess

Quantity: Numerals act as real numbers, reflect the actual amount of an attribute people possess

Levels of Measurement

The properties of numbers are closely related to the levels of measurement proposed by Stevens (1946)

Stevens’s levels of measurement are a set of rules that link the properties of numbers to particular types of observations

Levels of Measurement

Property of Numbers

	Nominal	Ordinal	Interval	Ratio
Identity	X	X	X	X
Order		X	X	X
Quantity			X	X
Rational Zero				X
Example	*Sex*	*Education*	*Memory*	*Behavior*

Scales should possess the property of quantity

Units of measurement are continuous, standardized quantities

- — The number 1 defines the size of a basic unit on the scale
- — Each number represents a count of basic units

Requires that units of measurement be clearly defined

- — In physical measurement, these units are readily apparent

  Example: measurement units are a tape marked off in inches or centimeters

- — In psychological measurement, the units are often less obvious

  Example: measurement units are responses to a series of survey questions

Methods exist to determine the degree to which measurement units on a scale reflect true psychological units

Course Objectives

1. Identify psychological measurement's goals and challenges
2. Understand basic measurement concepts and how they apply to psychological measurement
3. Understand scale development basics
4. Understand the importance of reliability and validity testing scales, factors that affect reliability and validity, and how to conduct reliability and validity testing

The primary goals of scale development are to:

1. Construct a scale with measurement units that are clearly defined and represent the appropriate psychological attribute

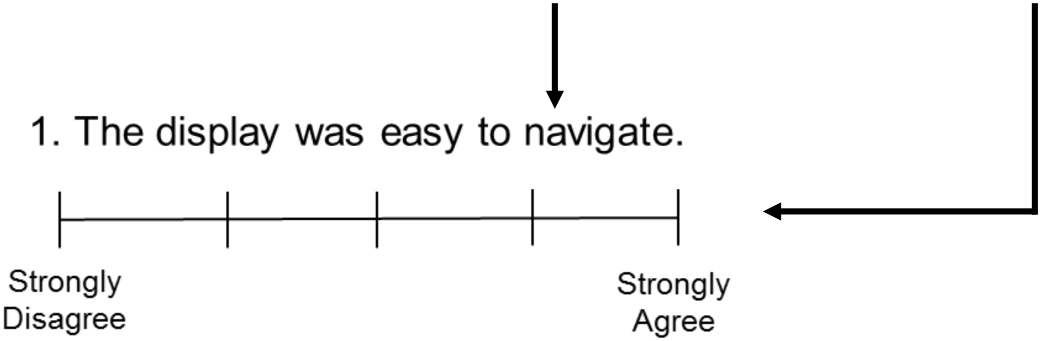
2. Construct a scale whose measurement units closely match the “true” psychological units for the attribute of interest

3. Construct a scale that possesses the property of quantity

Surveys measure people’s attitudes, opinions, feelings

Comprised of a series of questions

Questions consist of 2 parts, the item and response option



1. The display was easy to navigate.

A scale is a set of questions designed to measure the same psychological attribute (thought or feeling)

Not all surveys are scales!

Composite Scores

Typically, responses to items from multi-item scales are combined in some way to create a total score

— Can be summed or averaged

There are three primary advantages to composite scores

1. Better representation of the psychological attribute’s complexity
2. Estimates are more reliable (more about that later)
3. Composite scores clearly possess the property of quantity

Dimensionality Example

Consider how well each personality trait describes you:

1. Talkative
2. Assertive
3. Imaginative
4. Creative
5. Outgoing
6. Intellectual

Group these traits into clusters based upon similarity

How many clusters or “dimensions” did you create?

1 2 3

Cluster 1

Talkative

Assertive

Outgoing

“Extraversion”

Cluster 2

Imaginative

Creative

Intellectual

“Openness to Experience”

How many clusters or “dimensions” did you create?

Cluster 1	Cluster 2	Cluster 3
Talkative	Imaginative	Intellectual
Assertive	Creative	
Outgoing		

Dimensionality is a fundamental question in scale development, evaluation, and use.

Composite scores should reflect a single attribute

In general, when we measure an attribute of a person, we intend to measure a single attribute

- Example: adding measures of personality and memory together produces a total score that is meaningless

However, a scale may include multiple dimensions

- Example: you might have a workload scale that measures 3 dimensions of workload including mental, physical, and temporal.

Scale Dimensionality

Ask yourself 3 questions about as you develop a scale:

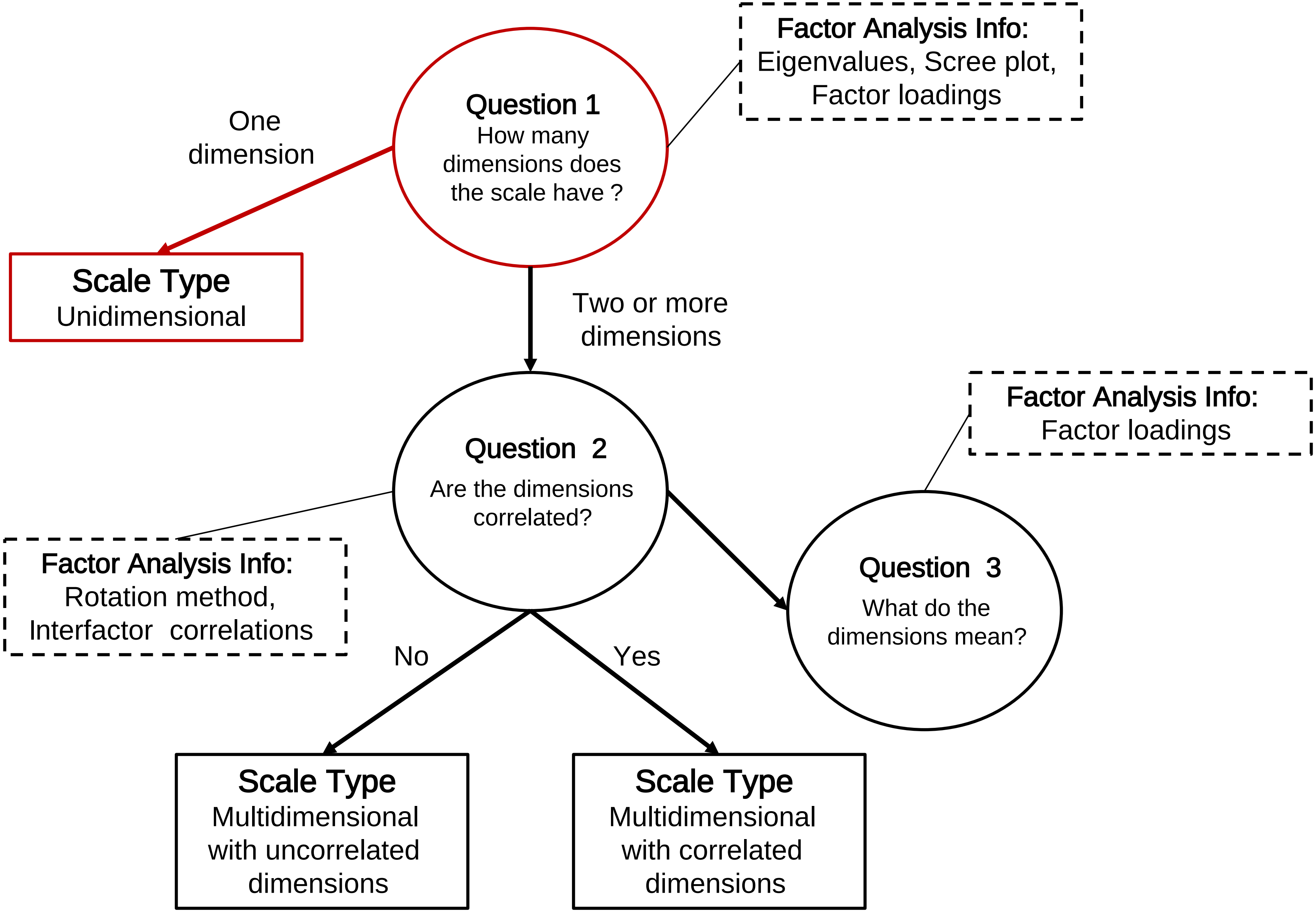
1. How many dimensions are reflected in the scale?
2. If the scale has multiple dimensions, are they correlated with each other?
3. If the scale has multiple dimensions, what are they?

Answers determine how scales are scored and interpreted

— Including if it’s appropriate to compute a “total” scale score for multi - dimensional scales

Factor analysis is useful for determining the number of dimensions in a scale and how they are correlated

Types of Scales

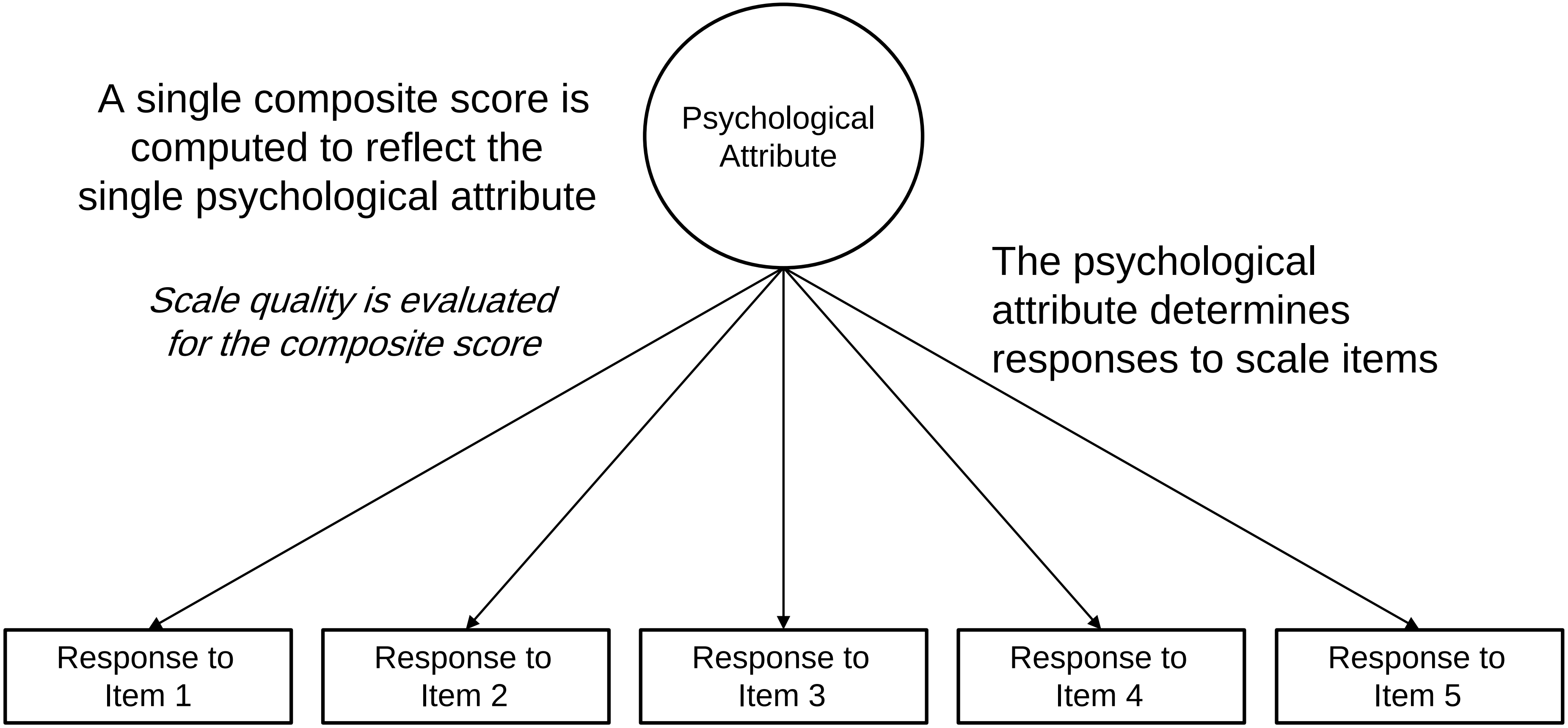

Factor Analysis Info:
Eigenvalues, Scree plot, Factor loadings
Question 1
How many dimensions does the scale have ?
One dimension
Scale Type
Unidimensional
Two or more dimensions
Factor Analysis Info:
Factor loadings
Question 2
Are the dimensions correlated?
Factor Analysis Info:
Rotation method, Interfactor correlations
Question 3
What do the dimensions mean?
No
Yes
Scale Type
Multidimensional with uncorrelated dimensions
Scale Type
Multidimensional with correlated dimensions

Furr & Bacharach (2014)

Unidimensional scales reflect a single psychological dimension

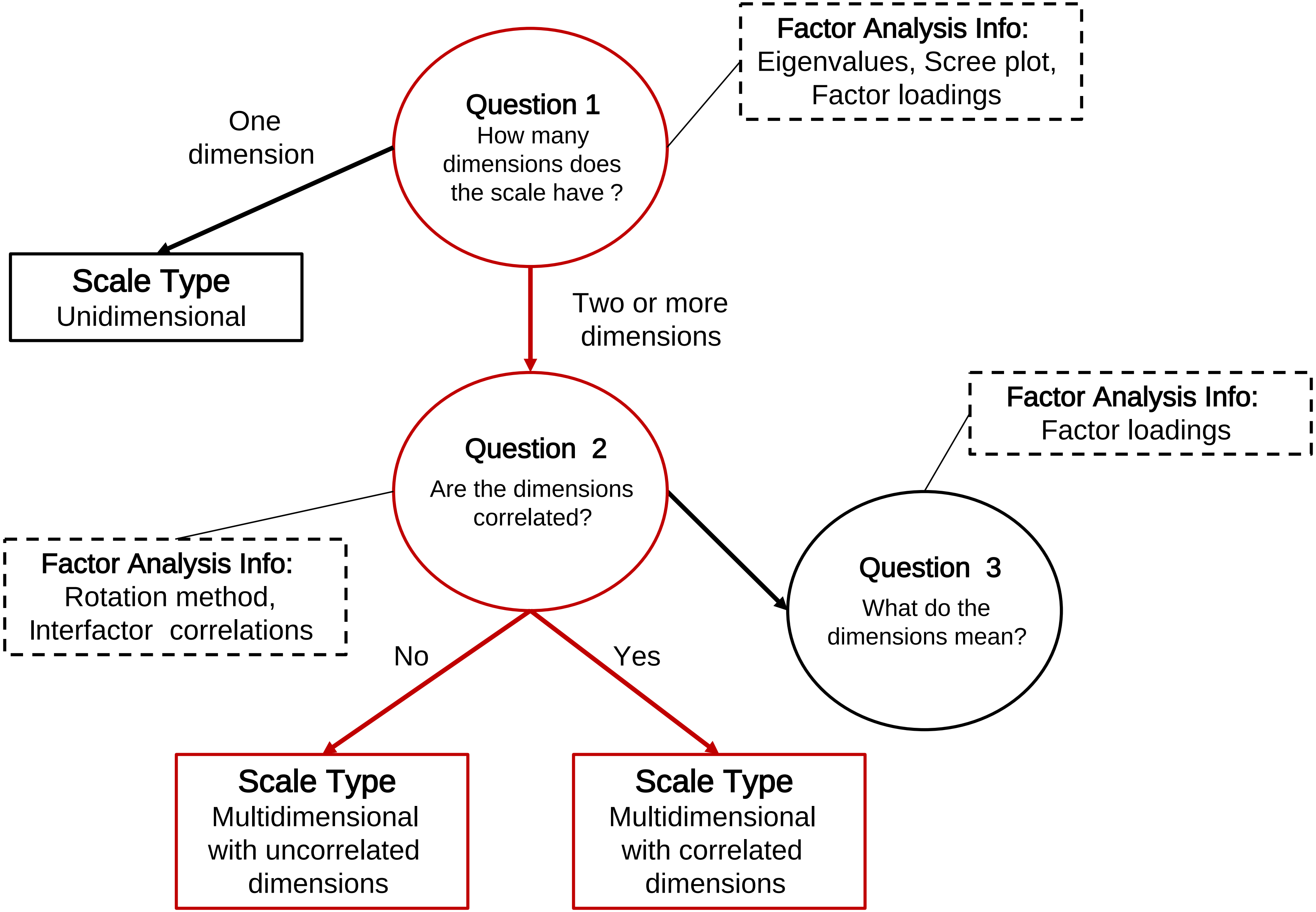
A single composite score is computed to reflect the single psychological attribute

Scale quality is evaluated for the composite score



The psychological attribute determines responses to scale items

Types of Scales



Question 1
How many dimensions does the scale have ?

Factor Analysis Info:
Eigenvalues, Scree plot, Factor loadings

One dimension

Scale Type
Unidimensional

Two or more dimensions

Question 2
Are the dimensions correlated?

Factor Analysis Info:
Rotation method, Interfactor correlations

No

Scale Type
Multidimensional with uncorrelated dimensions

Yes

Scale Type
Multidimensional with correlated dimensions

Question 3
What do the dimensions mean?

Factor Analysis Info:
Factor loadings

Furr & Bacharach (2014)

Multidimensional scales reflect two or more psychological attributes

Dimensions can be correlated or uncorrelated

- Scales with correlated dimensions are called *scales with higher-order factors*
- Scales with uncorrelated dimension are called *scales with uncorrelated dimensions* (surprise!)

Scales with Higher-Order Factors

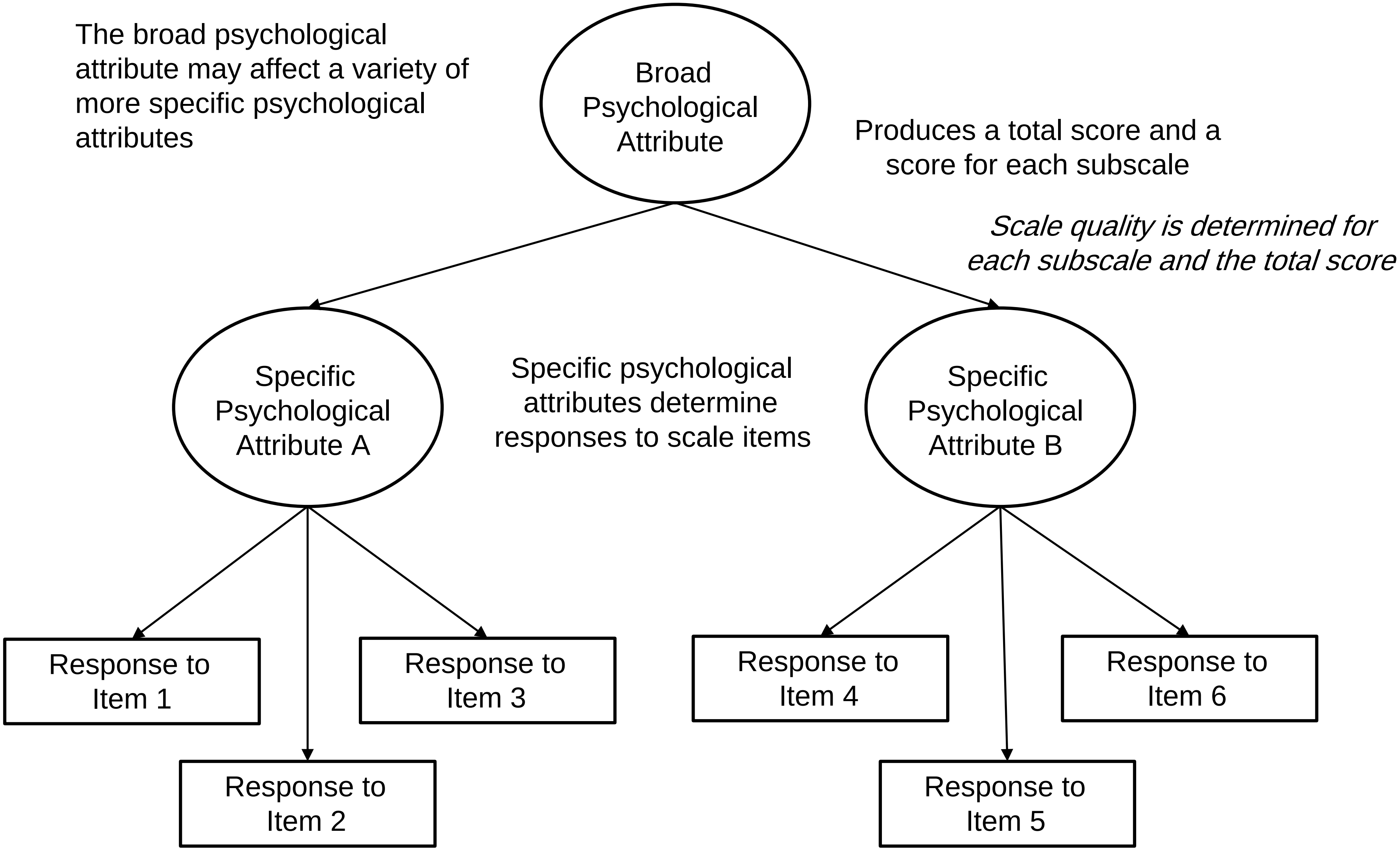
These scales include clusters of items that assess different psychological attributes called subscales

Each subscale reflects a different aspect of a broader psychological attribute and is itself unidimensional

Because the subscales (dimensions) are correlated, these scales can produce various scores

- — A score for each subscale
- — A total score, combined across subscales

Scales with Higher-Order Factors



The broad psychological attribute may affect a variety of more specific psychological attributes

Produces a total score and a score for each subscale

Scale quality is determined for each subscale and the total score

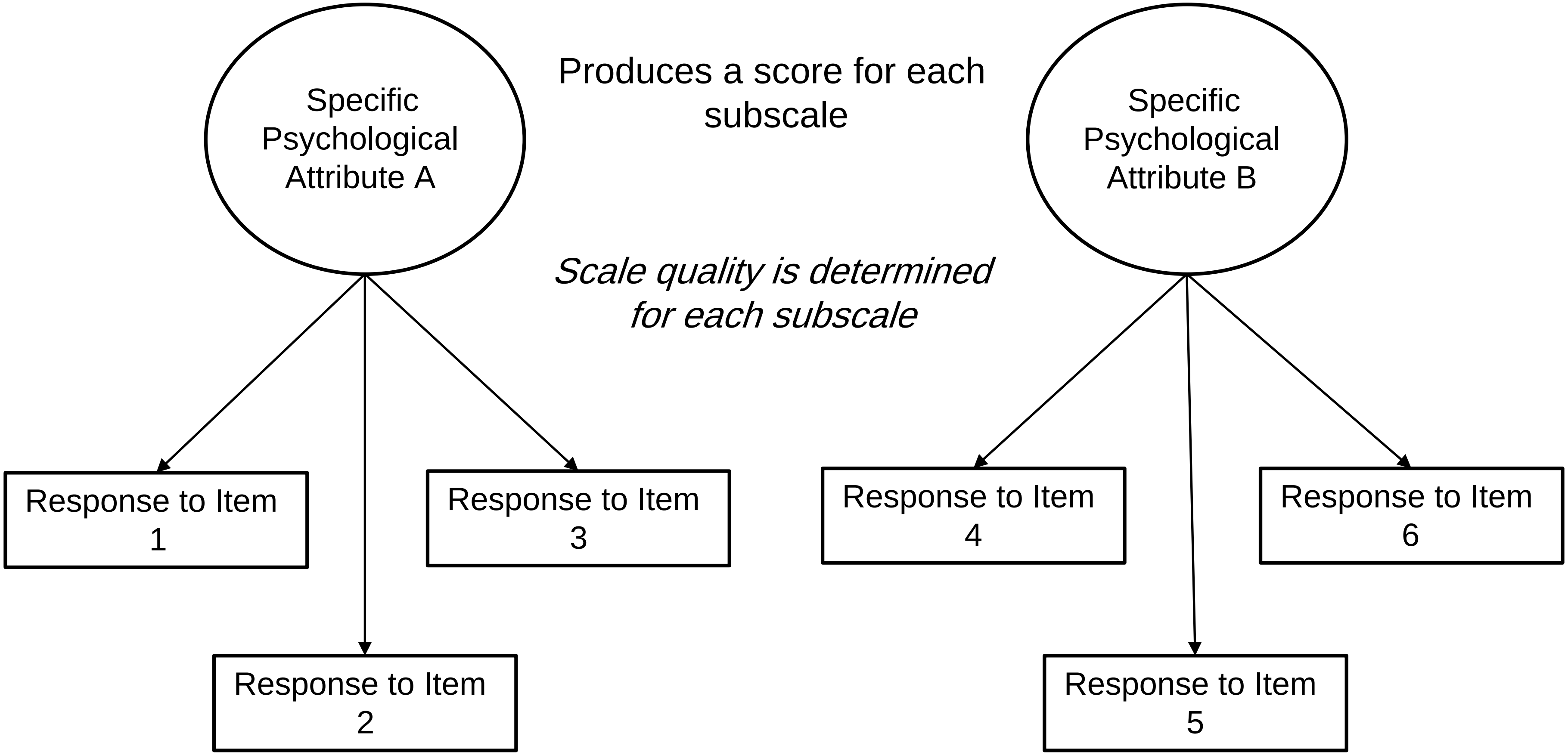
Specific psychological attributes determine responses to scale items

Scales with Uncorrelated Dimensions

Similar to scales with higher -order factors except the subscales are not linked by a broader psychological attribute.

In essence, these scales are a set of unrelated unidimensional scales that are presented with their items mixed together

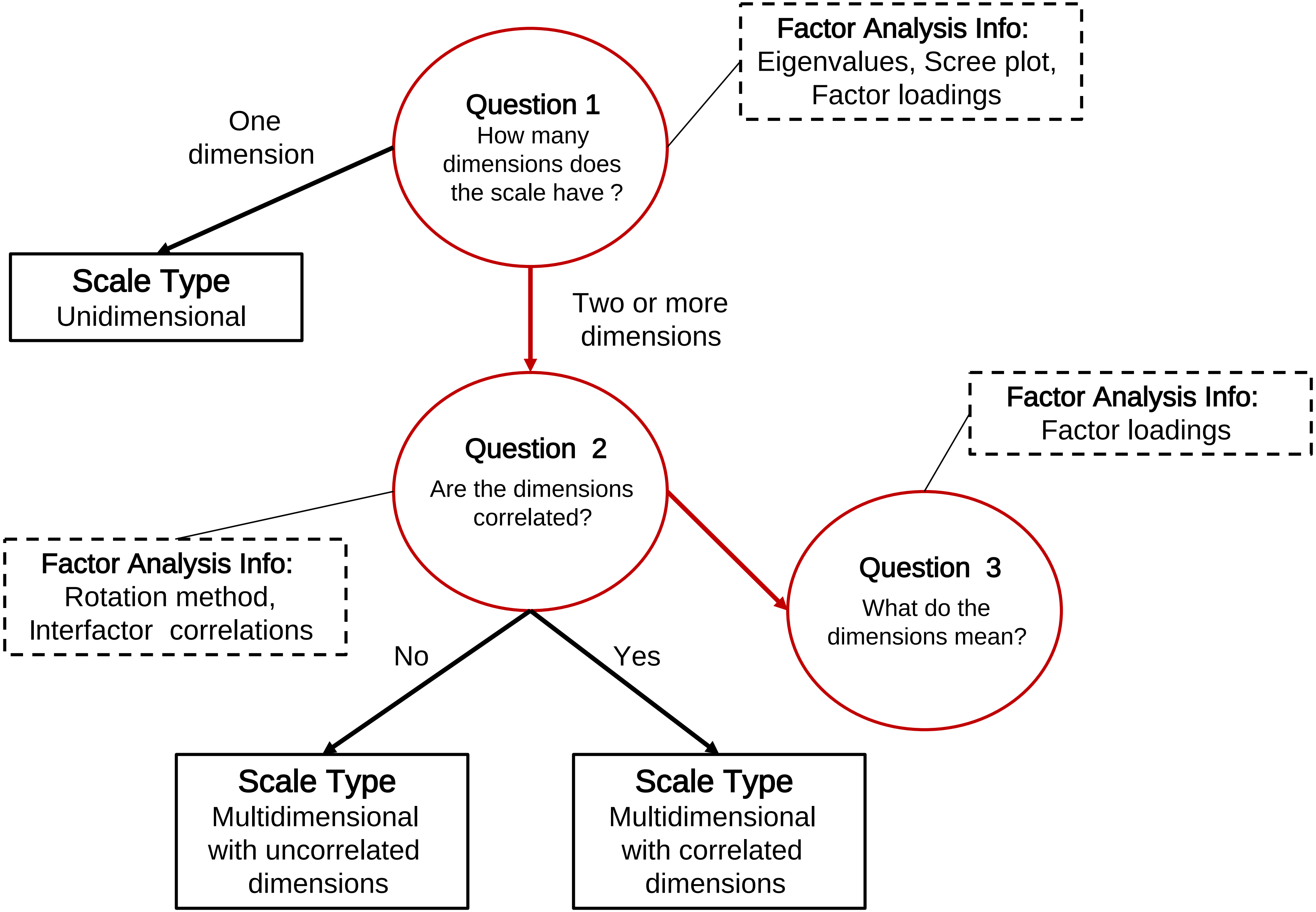
Scales with Uncorrelated Dimensions



Produces a score for each subscale

Scale quality is determined for each subscale

Types of Scales



Question 1
How many dimensions does the scale have ?

Factor Analysis Info:
Eigenvalues, Scree plot, Factor loadings

One dimension

Scale Type
Unidimensional

Two or more dimensions

Question 2
Are the dimensions correlated?

Factor Analysis Info:
Rotation method, Interfactor correlations

No

Scale Type
Multidimensional with uncorrelated dimensions

Yes

Scale Type
Multidimensional with correlated dimensions

Question 3
What do the dimensions mean?

Factor Analysis Info:
Factor loadings

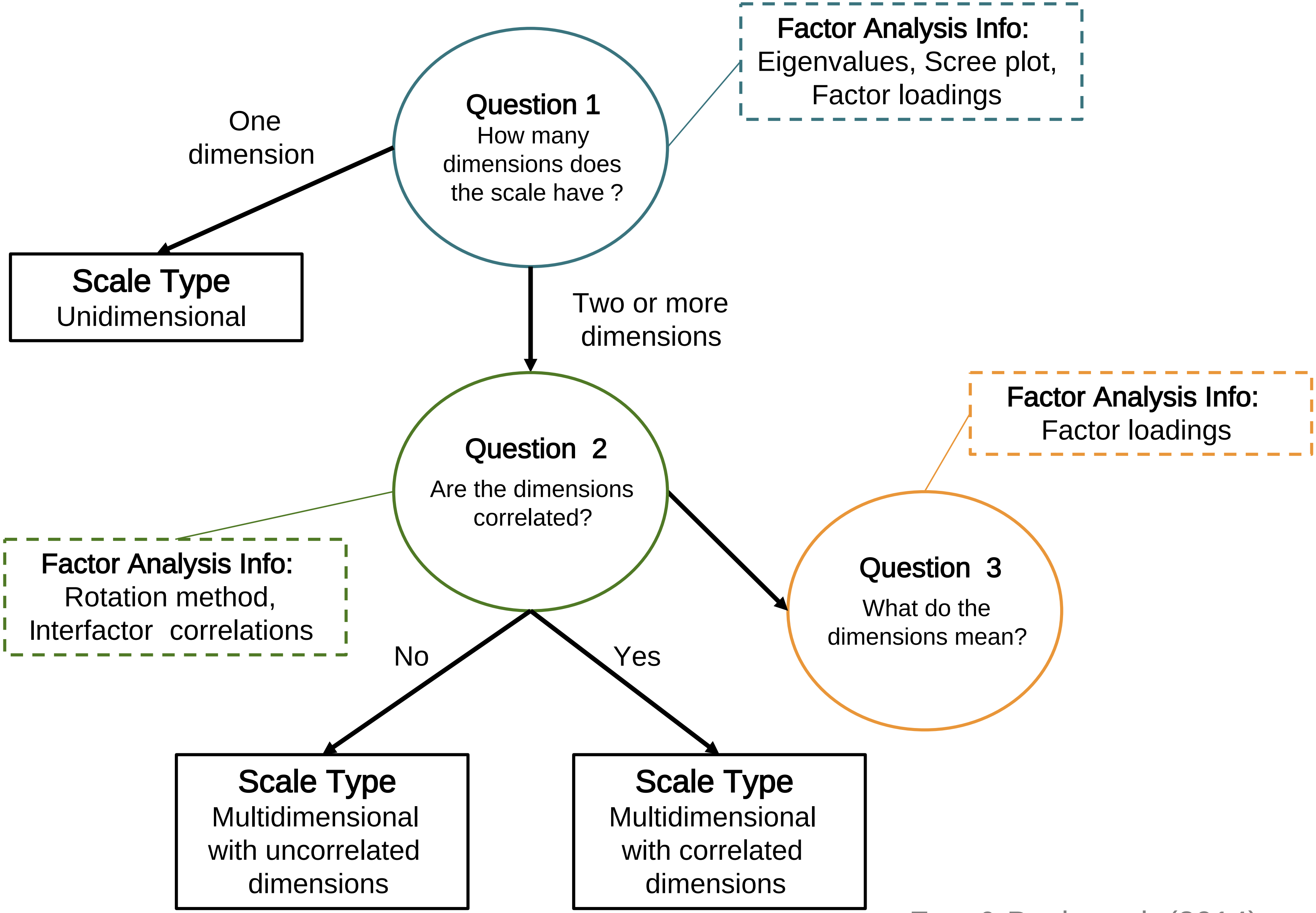
Furr & Bacharach (2014)

Psychological Meaning of Scale Dimensions

Research verifies the psychological attribute that is represented by each dimension

Factor analysis is a fundamental tool used to answer core questions about scale dimensionality

Types of Scales



Question 1
How many dimensions does the scale have ?
Factor Analysis Info:
Eigenvalues, Scree plot, Factor loadings
One dimension
Scale Type
Unidimensional
Two or more dimensions
Question 2
Are the dimensions correlated?
Factor Analysis Info:
Factor loadings
Factor Analysis Info:
Rotation method, Interfactor correlations
No
Yes
Question 3
What do the dimensions mean?
Scale Type
Multidimensional with uncorrelated dimensions
Scale Type
Multidimensional with correlated dimensions

Furr & Bacharach (2014)

Factor analysis is the most common method for evaluating scale dimensionality

Other statistical methods (e.g., cluster analysis, multidimensional scaling) are also available

Two types of factor analysis

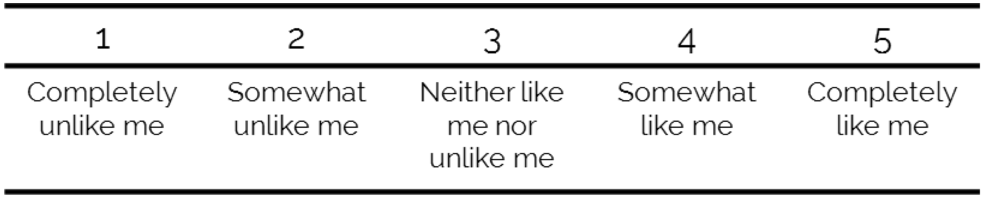
- — **Exploratory Factor Analysis (EFA)**
- — Confirmatory Factor Analysis (CFA)

This method grounds clusters of items in empirical data rather than idiosyncratic interpretations

Factor Analysis Example

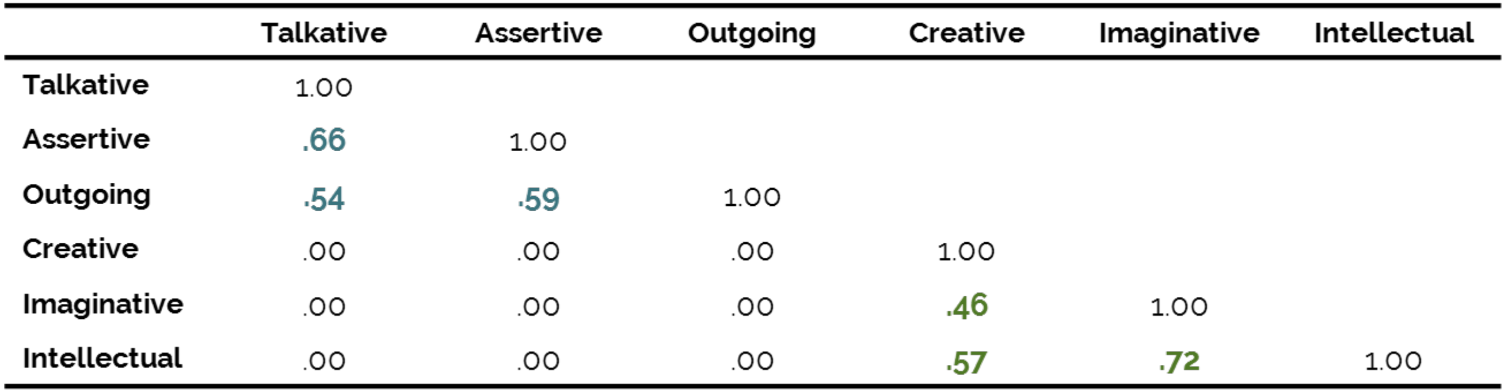
Imagine 100 soldiers rated how well 6 traits described them on the scale provided:

1. Talkative
2. Assertive
3. Imaginative
4. Creative
5. Outgoing
6. Intellectual



1	2	3	4	5
Completely unlike me	Somewhat unlike me	Neither like me nor unlike me	Somewhat like me	Completely like me

We can compute the correlations among the six items to help us identify and interpret the dimensions reflected



	Talkative	Assertive	Outgoing	Creative	Imaginative	Intellectual
Talkative	1.00					
Assertive	.66	1.00				
Outgoing	.54	.59	1.00			
Creative	.00	.00	.00	1.00		
Imaginative	.00	.00	.00	.46	1.00	
Intellectual	.00	.00	.00	.57	.72	1.00

Factor Analysis Example

Examining correlations is a very basic factor analysis

- Not typically possible with real data because there are more items and the correlational structure is less obvious

EFA simplifies this process

- Often an iterative process
- Results of one step lead researchers to reevaluate prior steps

Conducting Exploratory Factor Analysis (Basics)

Input participants raw scores into a statistical software package

- Ratings should be reverse scored if necessary before conducting EFA

Step 1: Choose an extraction method

- Specific statistical technique implemented
- **Options:** principal axis factoring (PAF), maximum likelihood (ML), and principal components analysis (PCA)

Results are often similar. However, some experts recommend PAF over PCA (MacCallum & Strahan, 1999). ML is typically reserved for CFA

Conducting Exploratory Factor Analysis (Basics)

Step 2: Identify the number of factors and extract them

Researchers typically rely on *eigenvalues*

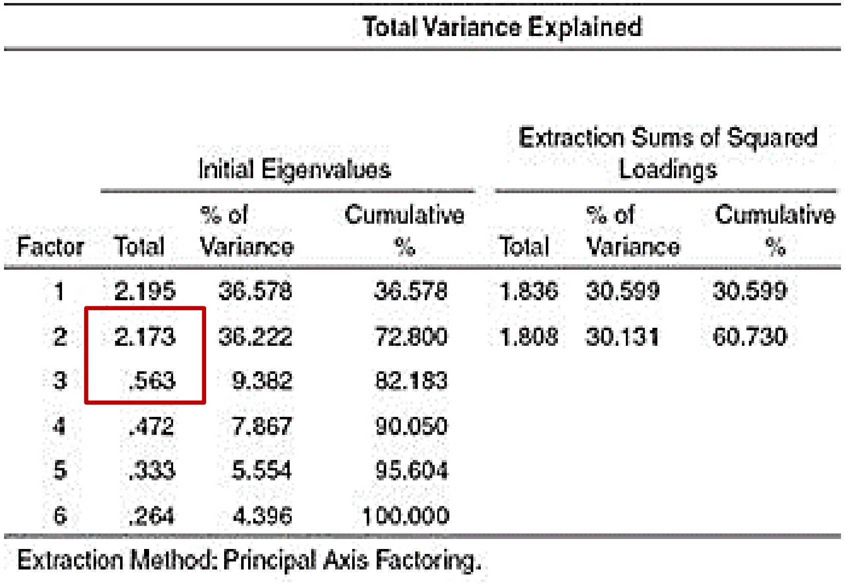
Eigenvalues are a special set of scalars associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic roots, characteristic values (Hoffman and Kunze 1971), proper values, or latent roots (Marcus and Minc 1988, p. 144)

Don't worry! You need to understand how eigenvalues are used not *necessarily* what they are…

How to Use Eigenvalues

Examine the relative sizes of the eigenvalues

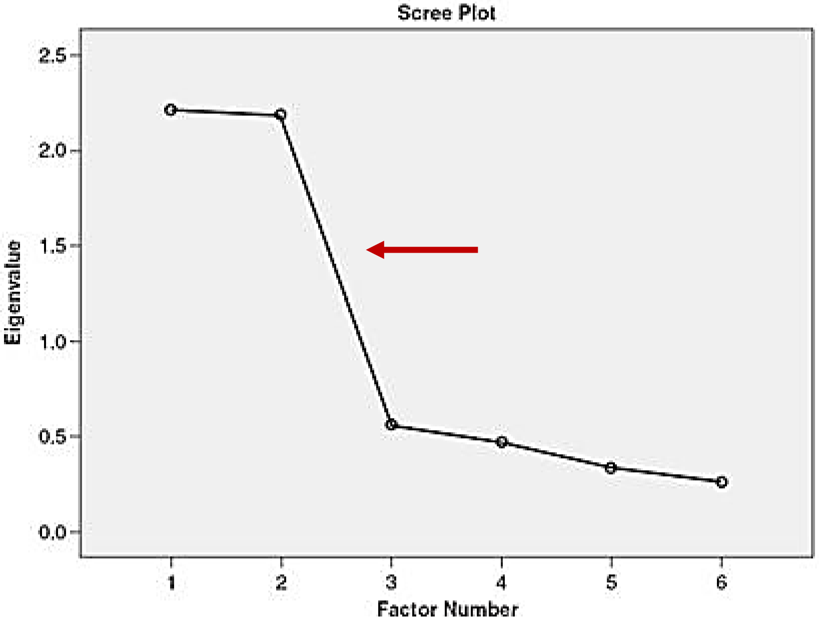
- Find point where all subsequent differences between values are relatively small
- The location of this point is indicative of the number of dimensions in the scale
- This same logic can be applied to scree plots



Total Variance Explained

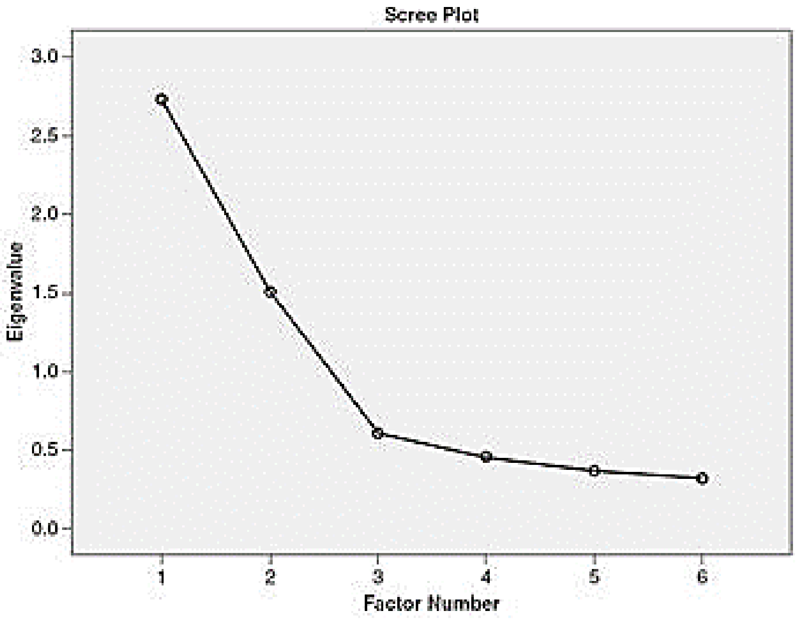
	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.195	36.578	36.578	1.836	30.599	30.599
2	2.173	36.222	72.800	1.808	30.131	60.730
3	.563	9.382	82.183			
4	.472	7.867	90.050			
5	.333	5.554	95.604			
6	.264	4.396	100.000			

Extraction Method: Principal Axis Factoring.



Caution

The appropriate number of factors is not always clear



1. Extract the number you think it is
2. Then examine the associations between items and that factor
3. Iterate if needed

If a clear number of dimensions cannot be determined, the scale likely needs to be revised

Conducting Exploratory Factor Analysis (Basic)

Step 3: Decide how you will rotate factors

The purpose of this step is to clarify the psychological meaning of the factors

Two types of rotations

— **Orthogonal:** generates uncorrelated factors (“Varimax”)

— **Oblique:** generates correlated or uncorrelated factors (“Promax”, “Direct Oblimin”)

Conducting Exploratory Factor Analysis (Basic)

Step 4: Examine Item-Factor Associations

These associations are determined using *factor loadings*

— Each item has a loading on each factor

Examine the loadings to identify which items are most strongly linked to each factor

— The similarities among items linked most strongly to a factor points to the factor's psychological meaning

Factor loadings range from -1 to 1

— Interpreted as correlations or standardized regression coefficients depending upon the rotation

Factor Loadings

Orthogonal rotations yield factor loadings that can be interpreted as correlations between each item and each factor

Oblique rotations yield 2 types of factor loadings

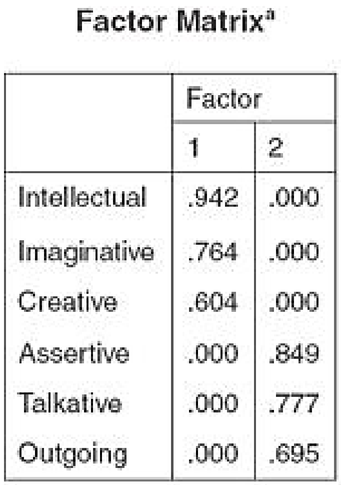
- — **Pattern coefficients:** item-factor association, controlling for the correlation between factors
- — **Structure coefficients:** simple item-factor correlations

Consider the size and direction of the loading

- — Loadings above .30 are “reasonable”, above .70 are “strong”
- — Interpret the direction like a correlation – a negative loading indicates that high scores on the item are associated with low scores on the underlying factor

Factor Loadings Example

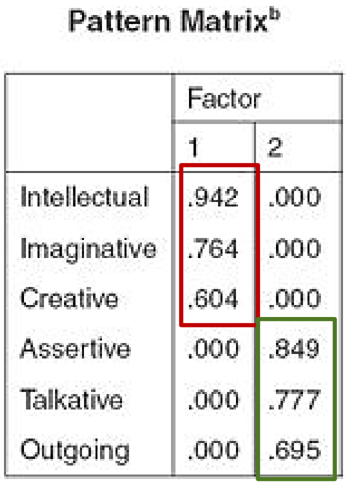
Imagine we chose an oblique rotation for the personality scale and obtained the following output



Factor Matrix[a]

	Factor	
	1	2
Intellectual	.942	.000
Imaginative	.764	.000
Creative	.604	.000
Assertive	.000	.849
Talkative	.000	.777
Outgoing	.000	.695

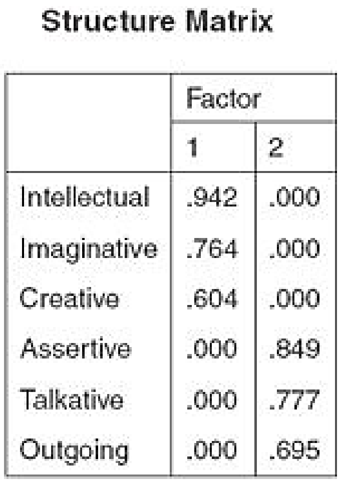
Loadings before rotation is applied



Pattern Matrix[b]

	Factor	
	1	2
Intellectual	.942	.000
Imaginative	.764	.000
Creative	.604	.000
Assertive	.000	.849
Talkative	.000	.777
Outgoing	.000	.695

Pattern Coefficients



Structure Matrix

	Factor	
	1	2
Intellectual	.942	.000
Imaginative	.764	.000
Creative	.604	.000
Assertive	.000	.849
Talkative	.000	.777
Outgoing	.000	.695

Structure Coefficients

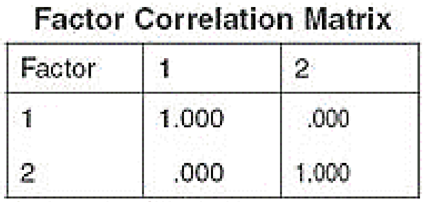
Results obtained from real data are rarely this tidy

Conducting Factor Analysis (Basic)

Step 5: Examine the association among factors

Oblique rotations allow factors to be correlated or uncorrelated

The degree of correlation among factors determines how to score the scale



Factor Correlation Matrix

Factor	1	2
1	1.000	.000
2	.000	1.000

How should this scale be scored?

We would create 2 subscales

One composite for each subscale

No total score!

In contrast to EFA, CFA specifies the scale structure a priori

Course Objectives

1. Identify psychological measurement’s goals and challenges
2. Understand basic measurement concepts and how they apply to psychological measurement
3. Understand scale development basics
4. Understand the importance of reliability and validity testing scales, factors that affect reliability and validity, and how to conduct reliability and validity testing

Scale Quality

We want our scales to be reliable and valid

Reliability: extent to which scale scores are a function of respondents' true psychological differences as opposed to measurement error

Validity: extent to which scale scores reflect what the scale is intended to measure

Statistical methods exist to evaluate reliability and validity

Methods for Establishing Reliability

There are three primary methods for estimating reliability

- — Alternate forms reliability
- — Test-retest reliability
- — Internal consistency reliability

Note: all of these methods are derived from the notion of parallel tests

Provide estimates of the proportion of observed score variance that is attributable to true score variance

$$\boldsymbol{Observed\ score\ = True\ score + Error}$$

Alternate Forms Reliability

1. Develop two *parallel* versions of a scale
   - Items must probe the same psychological attribute
   - Equivalent amount of error variance

$$\boldsymbol{Observed\ score} = \boldsymbol{True\ score} + \boldsymbol{Error}$$

2. Administer the two versions of the scale to the same group of people

3. Measure the correlation between scores on both versions

Potential Issues

We can never be **certain** that versions are parallel

Repeated testing may inflate correlations

Test-Retest Reliability

Avoids the parallel versions problem

1. Administer the same scale on two different occasions
   - Assume true scores are stable across the two occasions

$$\boldsymbol{Observed\ score\ = True\ score + Error}$$

2. Measure the correlation between scale scores

Potential Issues

Some psychological attributes are more stable than others

- For example, usability vs. workload

Test conditions must be as similar as possible

Internal Consistency Reliability

Doesn't require 2 versions of a scale or 2 test occasions!

Can only be used with multi-item scales

Method differs slightly depending upon data type (binary vs. continuous) and desired estimation procedure (use of item variances, inter-item correlations, inter-item correlations)

All procedures are a two-step process

1. Administer scale to a group of people
   — Each item is treated as different "versions" of a scale
2. Estimate consistency of items using an equation

Most common procedure is Cronbach's alpha

Cronbach’s alpha

1. Administer the scale to a group of people

2. Calculate the covariance between each pair of items
   — Covariance reflects the degree of association between 2 variables (items)
   — We hope to find that items in a scale positively covary

3. Sum the covariances
   — The larger the sum is, the more consistent the items are with each other

4. Submit the variance of scores on the complete test and the sum of the covariances into the following equation:

$$\propto = \boldsymbol{estimated\ R_{xx}} = \left(\frac{\boldsymbol{k}}{\boldsymbol{k-1}}\right)\left(\frac{\sum \boldsymbol{c_{ii}}}{\boldsymbol{s_x^2}}\right)$$

sum of covariances

# items

variance of total score

Cronbach’s alpha

Produces a score between 0 and 1

The closer the score is to 1, the greater the internal consistency

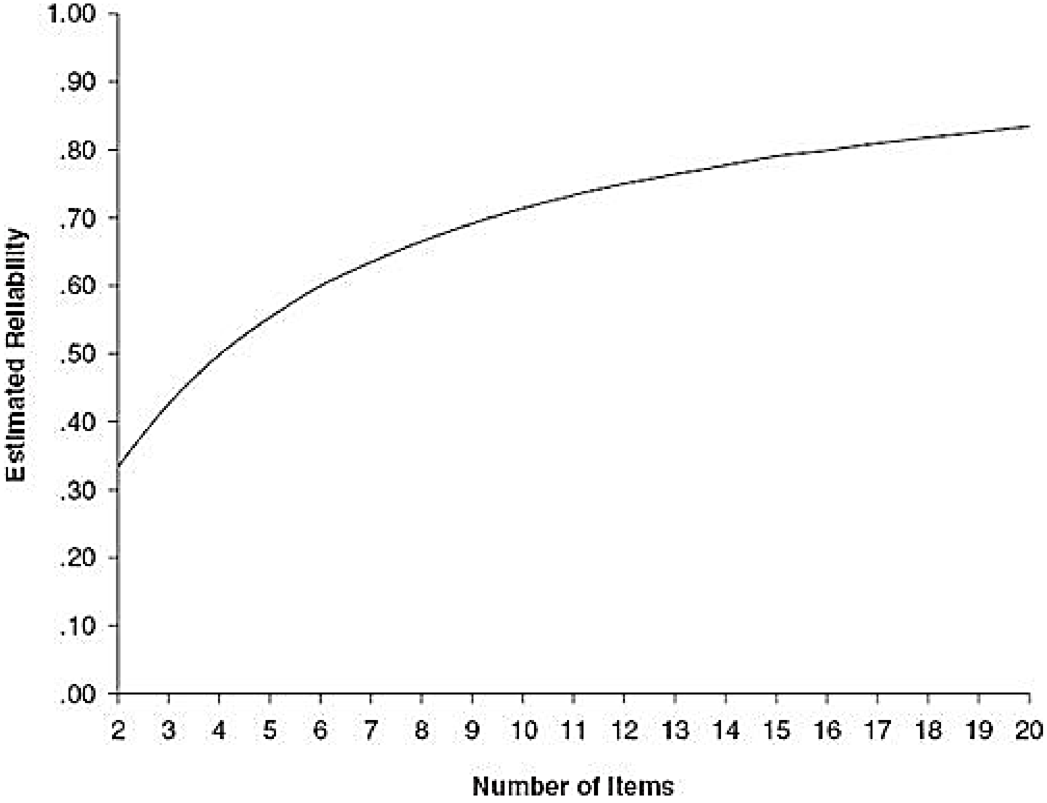
- $\alpha = 0.70$ is typically recognized as an “acceptable” level of internal consistency

Most statistical software also computes “Cronbach’s alpha if deleted”

- Useful for identifying items that degrade internal consistency

Evidence suggests that Cronbach’s alpha serves as a sort of “lower bound” on internal consistency

Note: Longer scales are more reliable



For a scale with an average inter -item correlation of 0.30

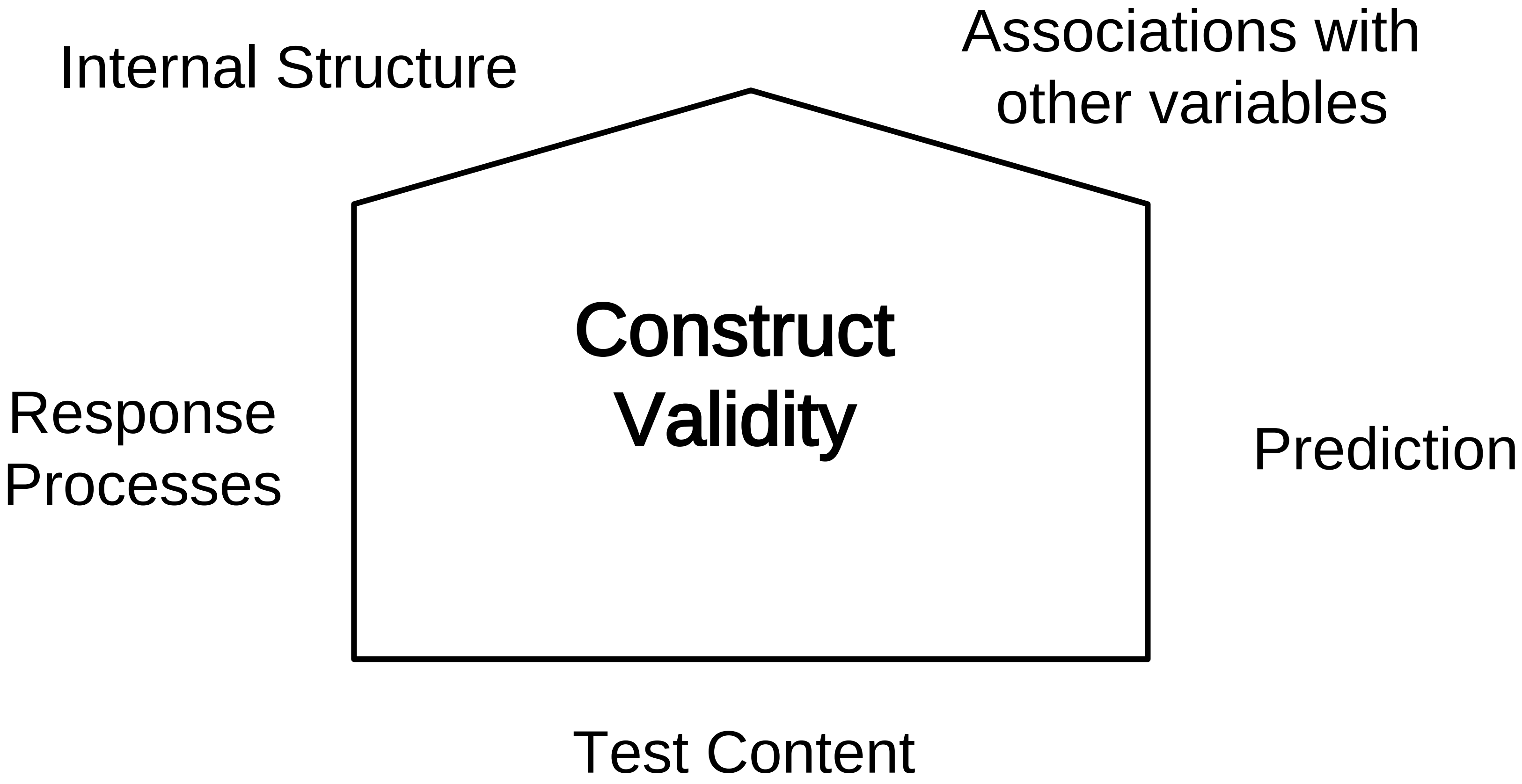
Conceptualizing Validity

Validity is a matter of degree

Validity is a based on *theory* and *evidence*

…NOT someone's "experience"

Information Needed to Establish Validity



Internal Structure

Associations with other variables

Construct Validity

Response Processes

Prediction

Test Content

Methods for Establishing Validity

The **internal structure** (number of dimensions) of the scale should match the theoretically based structure of the scale

— Factor analysis!

The **psychological processes** that respondents actually use when responding to scale items should match the processes they should use

— Qualitative procedure called *Cognitive Interviewing*

Demonstrate **associations** between the scale and measures of theoretically related psychological attributes

— Commonly called *Convergent Validity*

— Statistical Procedures: Correlation, Regression

Methods for Establishing Validity

Demonstrate that the scale **predicts** behaviors that it should theoretically be able to predict

- Commonly called *Predictive Validity*
- Statistical Procedures: Regression

Course Objectives

1. Identify psychological measurement’s goals and challenges
2. Understand basic measurement concepts and how they apply to psychological measurement
3. Understand scale development basics
4. Understand the importance of reliability and validity testing scales, factors that affect reliability and validity, and how to conduct reliability and validity testing

Questions?

U.S. ARMY EVALUATION CENTER

CASE STUDY: Developing the Operational Assessment of Training Scale

Shane Hall

Scale Development Process

Current Approach:

- Evaluate new equipment training (NET) for almost all Army acquisition systems during OT
- NET survey assessments vary in length and detail
- All evaluators are trying to answer the same question … Did soldiers receive high-quality training?

Can we rely on a single question to address the quality of training?

NO – concept is too broad

Develop a set of items!

1. Reviewed AEC’s database questions and selected a subset that were well-written and most applicable.
2. Developed items for each element of training.

Initial list of AEC-drafted items

Operator

1. The overall training enabled me to safely operate the system.
2. The overall length of training was appropriate.
3. The hands-on training was useful.
4. The training simulator/device provides realistic training.
5. The practical exercises were useful.
6. Classroom training was useful.
7. The pace of training was appropriate.
8. The content was organized and easy to follow.

Maintainer

1. The overall training enabled me to maintain the system.
2. The overall length of training was appropriate.
3. The hands-on training was useful.
4. The practical exercises were useful.
5. Classroom training was useful.
6. The pace of training was appropriate.
7. Training devices were useful.
8. The content was organized and easy to follow.

Initial list of AEC-drafted items

Operator

1. The overall training enabled me to safely operate the system.
2. The overall length of training was appropriate.
3. The hands-on training was useful.
4. The training simulator/device provides realistic training.
5. The practical exercises were useful.
6. Classroom training w
 of t

Issue: Failed to center items around what defines a quality training

Maintainer

4. es were useful.
5. Classr ining was useful.
6. The pace of training was appropriate.
7. Training devices were useful.
8. The content was organized and easy to follow.

Initial AEC-drafted items - Iteration 2

Instructor

1. The instruction was given at an acceptable pace.
2. The instructor adapted the training content to my needs.
3. The instructor has adequate knowledge of the system.
4. The instructor adequately answered questions.

Materials

1. The classroom materials were useful.
2. The classroom materials flowed in a logical sequence.
3. The classroom materials were an appropriate length.
4. The classroom materials were not too redundant.

Interactive Materials

1. The hands-on exercises were useful.
2. The practical exercises were operationally realistic.
3. The practical exercises provided full coverage of my mission.
4. The practical exercises were relevant to my job.

Confidence to use my system

1. I am confident I can perform the required tasks.
2. Training prepared me to transition to the new capability.
3. I do not required additional training.

Initial AEC-drafted items - Iteration 2

Instructor

1. The instruction was given at an acceptable pace.
2. The instructor adapted the training content to my needs.
3. The instructor has adequate knowledge of the system.
4. The instructor adequately answered questions.

1. The cla useful.
2. Th l sequenc

n.

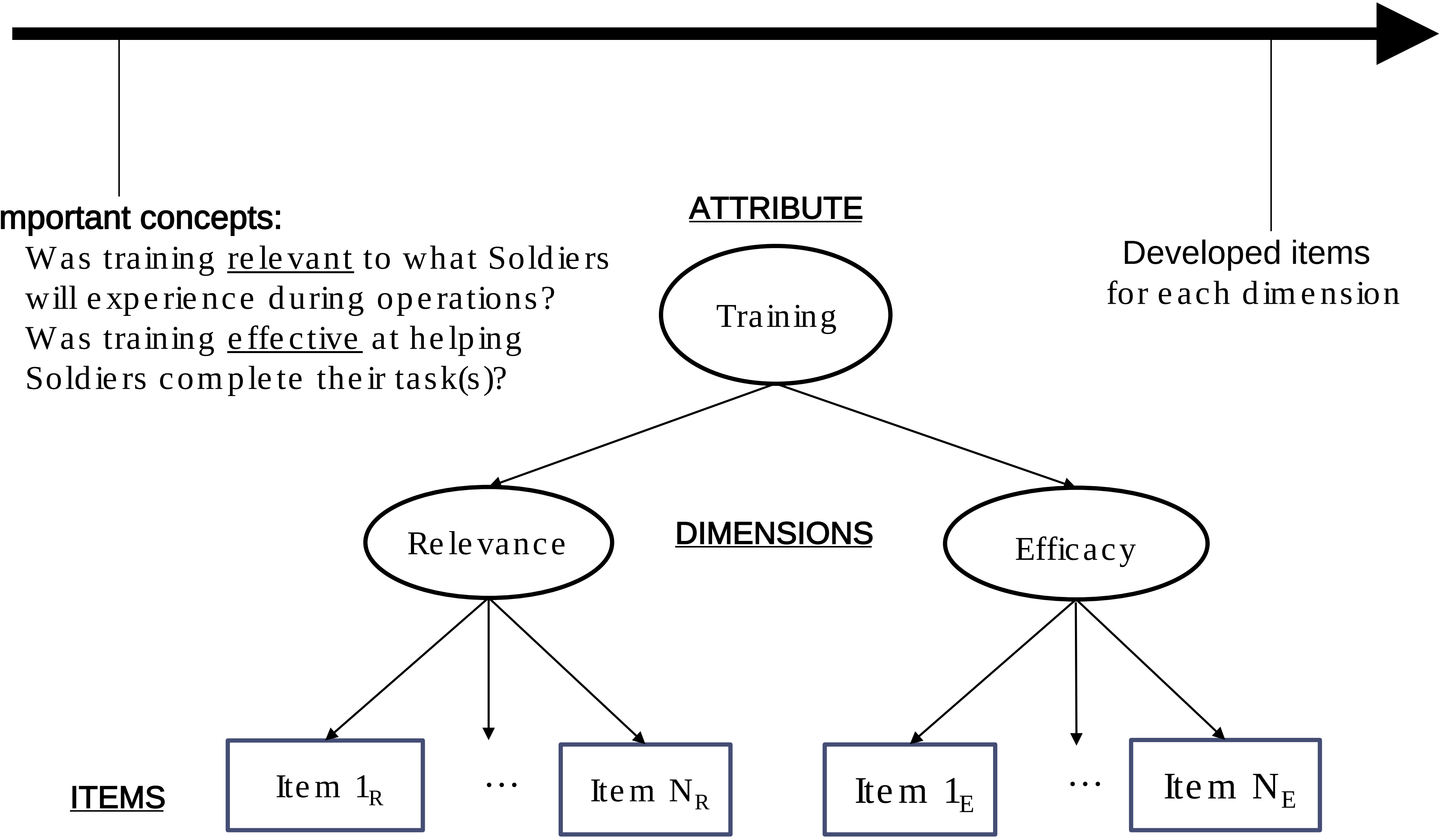
al exercises we

Issue: Dimensions developed are elements of training, not what defines a good training

Confidence to use my system

1. I am confident I can perform the required tasks.
2. Training prepared me to transition to the new capability.
3. I do not required additional training.

Scale Development Process

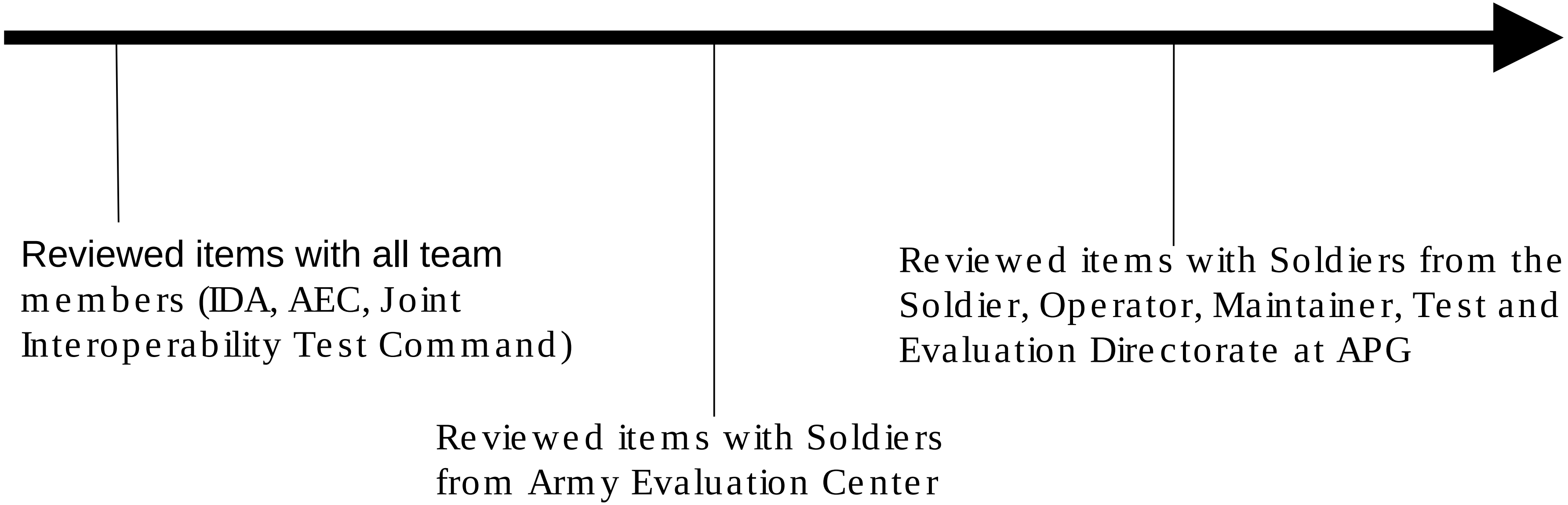


Important concepts:

- Was training relevant to what Soldiers will experience during operations?
- Was training effective at helping Soldiers complete their task(s)?

Developed items for each dimension

Scale Development Timeline



Reviewed items with all team members (IDA, AEC, Joint Interoperability Test Command)

Reviewed items with Soldiers from Army Evaluation Center

Reviewed items with Soldiers from the Soldier, Operator, Maintainer, Test and Evaluation Directorate at APG

In-depth review by multiple stakeholder groups

List of Final Items

Relevance

1. I can see myself using what I learned in training during real operations.
2. All of the information covered was relevant to how I interact with the system.
3. Training accurately portrayed operations in the field.
4. Training did not cover important ways I interact with the system.
5. Training adequately covered all important ways I interact with the system.
6. I would not make changes to the course content.
7. The course covered topics I don’t think should have been covered.
8. The training had a lot of information that wasn’t relevant to me.
9. The course’s level of difficulty was appropriate for someone in my position.

Efficacy

1. I’d be confident using the system during real operations without additional training.
2. I’d want additional training before using the system during real operations.
3. The training improved my understanding of how to interact with the system.
4. The training prepared me to properly interact with the system.
5. Training prepared me to solve common problems.
6. The training prepared me to easily use the system to accomplish my mission.

What’s Next

Piloting scale at the following OTs:

- Joint Air-to-Ground Missile (JAGM)
- Apache
- Next Generation Squad Weapon (NGSW)
- Joint Planning and Execution System (JPES)

Data from test events will determine:

- proper scale dimensionality
- reliability and validity