Concepts in questionnaire technique

Alice Ørts Hansen

Associate professor
Department of Ortopaedic
Odense University Hospital and Svendborg Hospital
Department of Clinical Research, University of Southern Denmark, Odense, Denmark.

Agenda

Time	Contents	Pages	Teacher
9.00– 10.00	Introduction to questionnaires - conceptualisation	7-13, 33-35	AØH
10.00 – 11.00	Exercise: Conceptualisation		AØH
11.00 – 12.30	Concepts in questionnaire technique • Observed vs. latent variables • Conceptual frameworks • Reflective and formative models • Scales vs. indexes COSMIN taxonomy and requirements to questionnaire validation	13-17, 35-37 Articles on www	AØH
12.30 – 13.00	*Lunch*		
13.00 – 13.45	How to measure the construct – operationalisation • Preparing items and answer categories	35-37 41-50	AØH
13.45 – 14.30	Exercise: Preparing items and answer categories		AØH
14.30 – 15.00	Structure and design of a questionnaire study Q & A about own study		AØH

* www = see www.clinimetrics.sdu.dk

Types of variables

Observed variables - ´directly´ measurable variables

- What is your weight (kg)?
- Have you been admitted to a hospital?
- In the last week have often did you feel stressed?
- Grip strength

Latent variables - ´indirectly´ measurable variables

- Can be measured inderectly, measuring observable charcteristics related to the non-observable construct
- Abstract concepts ex. experienced stress

Latent variables and factors

Almost all latent variables are multifactorial

- I.e. more factors are necessary to describe a certain variable

Each factor can be described with

- A global item or
- An instrument consisting of more items (a multi-item-scale)

Examples of latent variables

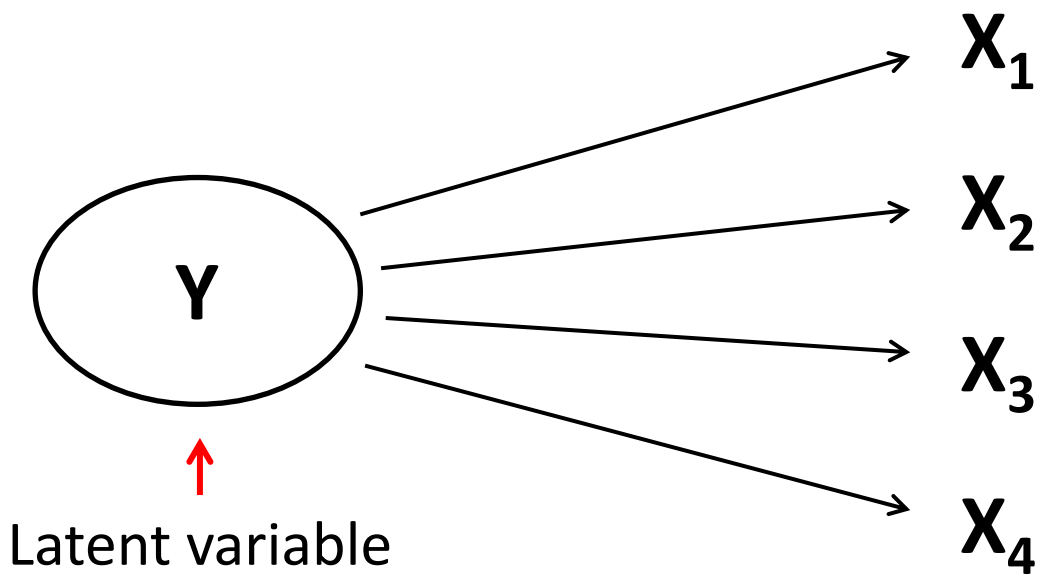
Intelligence: Measured through various cognitive tests assessing reasoning, problem-solving, memory, etc.

Depression: Inferred from symptoms such as sadness, loss of interest, fatigue, changes in sleep patterns, and others on a depression inventory scale.

Socioeconomic Status: Estimated through variables like income, education level, and occupation.

Job Satisfaction: Assessed through employee responses to questions about work environment, management support, and job role satisfaction.

Latent variable Y and items $X_{1\text{-}4}$



An example


Happy
Satisfaction with work
Fitness
X_1
X_2
X_3
Good friends
Leisure time

Exercise

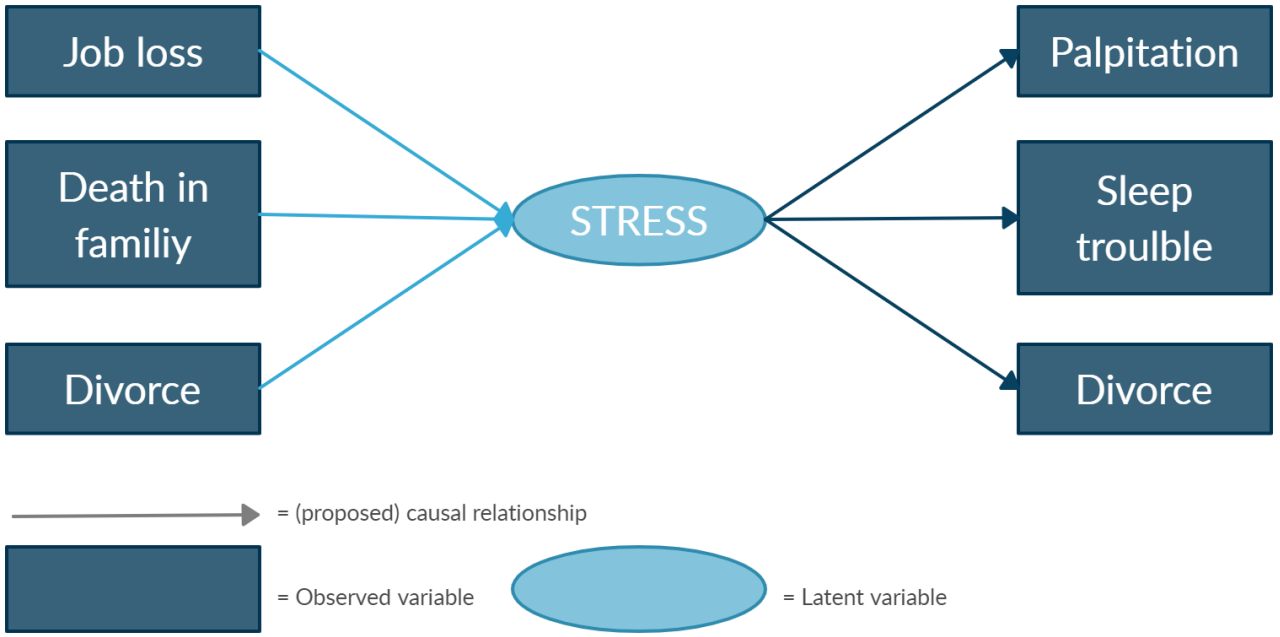
Give examples af questions that measure social network

- Netværkets størrelse
- Kontaktfrekvens
- Oplevet social støtte
- Oplevet ensomhed
- Relationernes kvalitet

Conceptual frameworks

Definition:

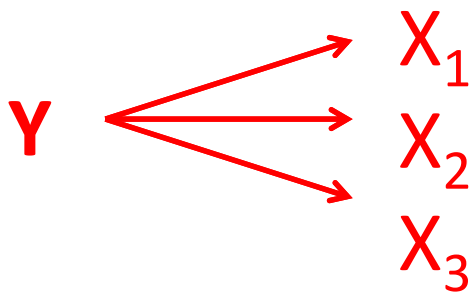
– ”A model representing the relationships between the items and the construct (could be a latent variable) to be measured”



Reflective vs. formative models

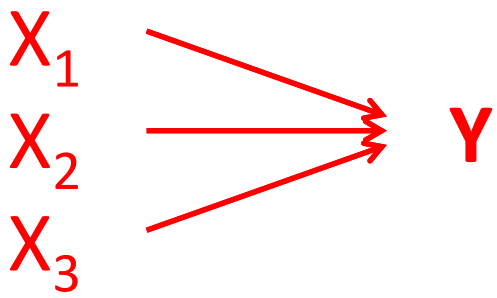
Reflective model

If the construct Y is changed, then *all* items will change too



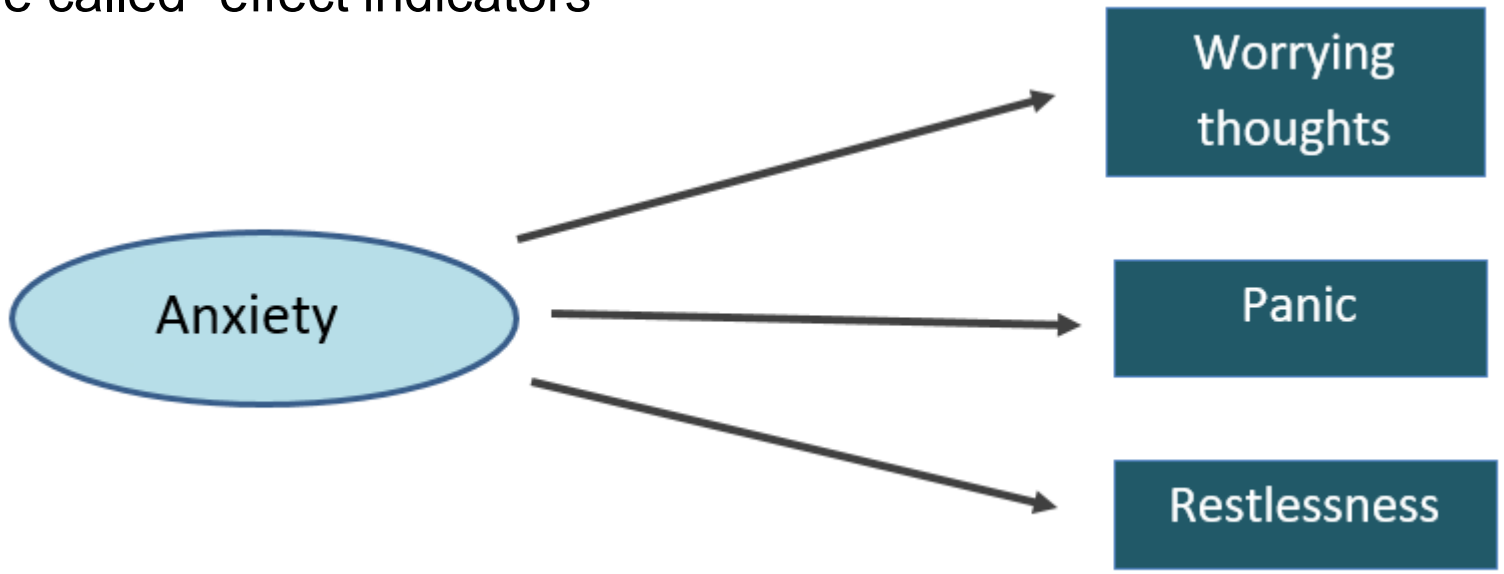
Formative model

If the construct Y is changed, then *all* items will <u>not</u> necessarily change



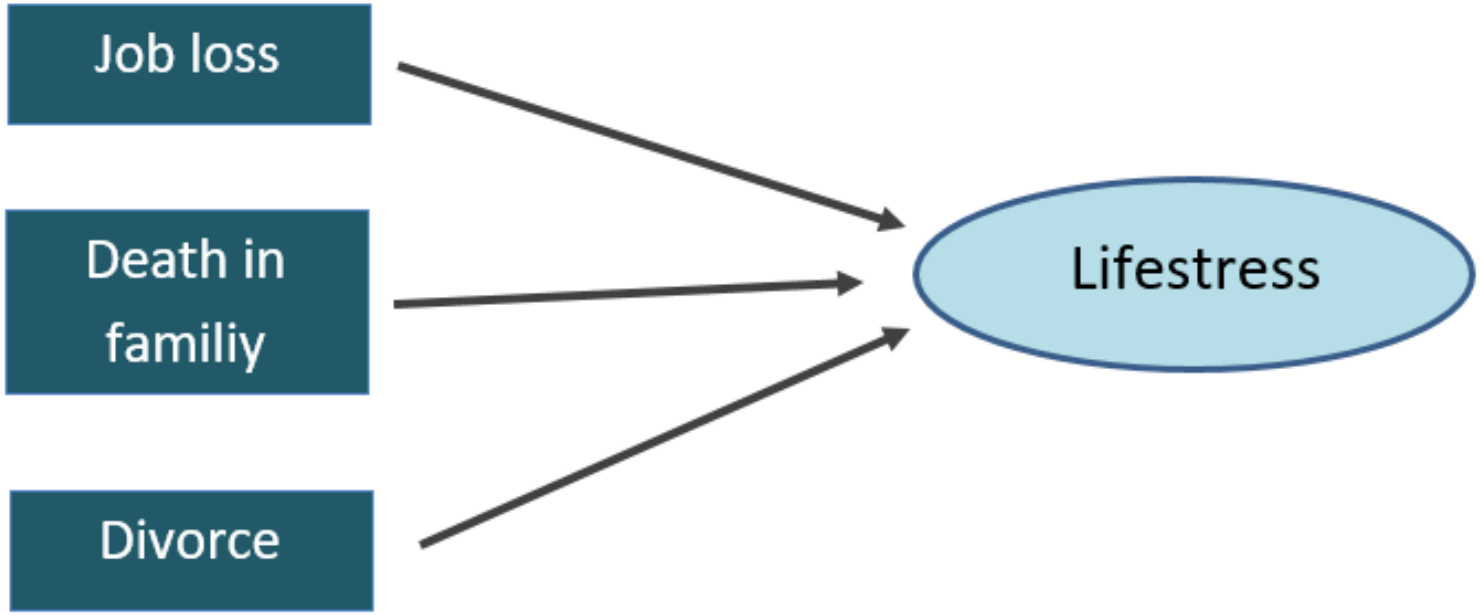
Reflective model

- The construct is reflected by the items
- Hence items correlate -> internal consistency is important
- All items change when the construct changes
- Items are manifestations of the construct
- Items are called ”effect indicators”



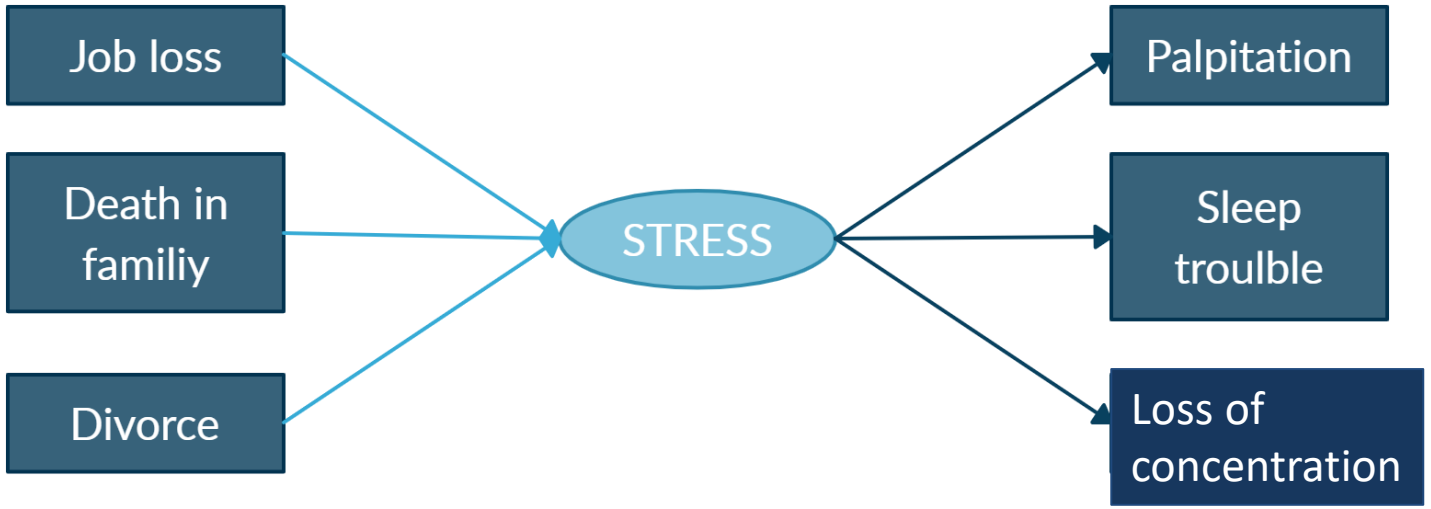
Formative model

- The items form/cause the construct
- Each item contributes to a part of the construct
- Together the items form the whole construct
- Internal consistency not important
- Items are called ”causal indicators”



Reflective vs. formative model

- Do we expect the items to change when the construct changes?
  - Yes → Reflective model
  - No → Formative model
- **The distinction is not always clear cut**



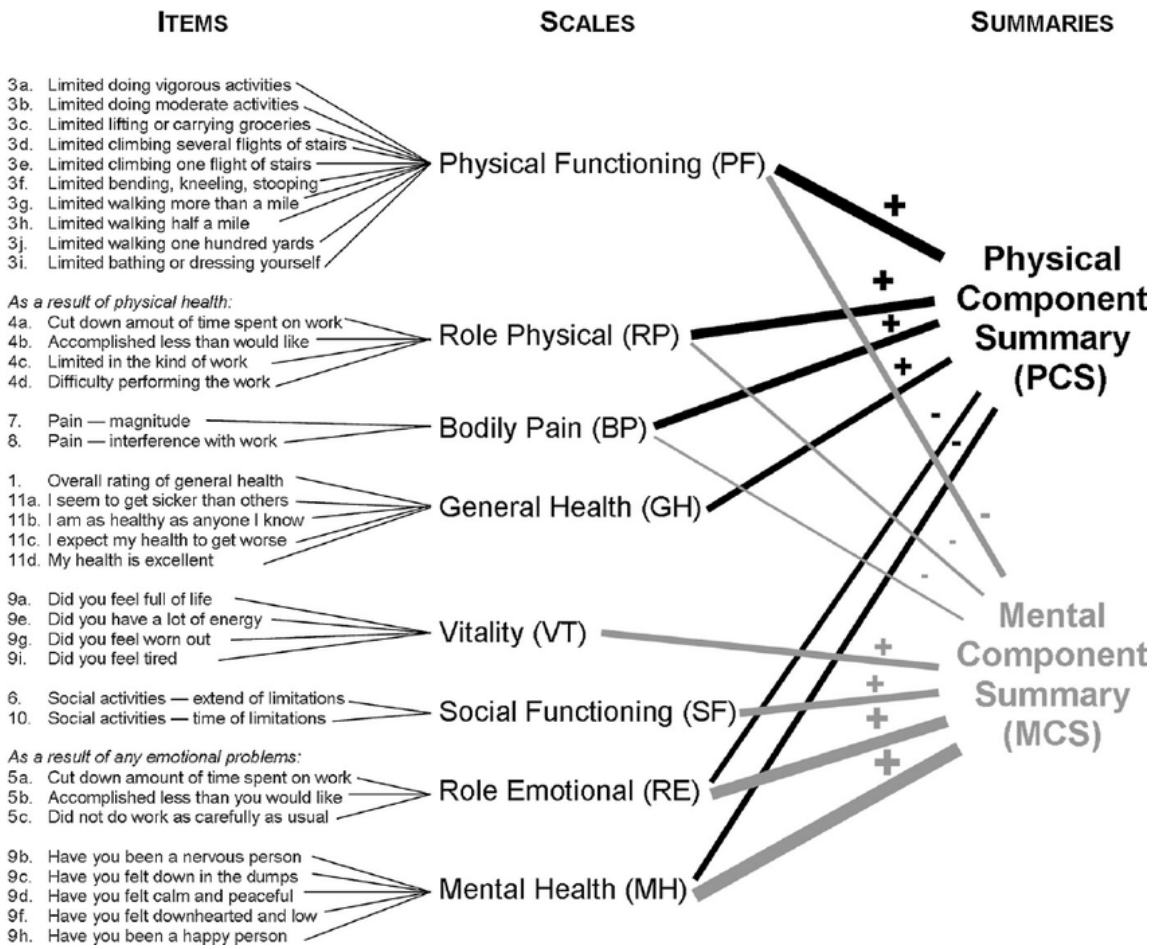
Reflective vs. formative model

Reflective	Formative
A change in latent variable/construct will change the effect indicators	A change in latent variable/ construct will NOT change the effect indicators
Is often uni-dimensional	Is multi-dimensional
Modern Item Response Theory (IRT) can be used	Modern Item Response Theory (IRT) is biased
Factor analysis and Chronbachs alpha makes sence	Factor analysis and Chronbachs alpha makes no sence
Measures the latent variable	Measures e.g. a confounder

Scale, index and profile

Multi-item instruments	Uni or multi-dimensional	Scores
Scale	*Unidimensional*: set of items measuring one dimension	Sum-score based on a reflective model or latent constructs
Index	*Multidimensional*: set of items measuring different dimensions	Sum-score based on a formative model or observable constructs
Profile	*Multidimensional*	One score per dimension

SF36 – Scale, index or profile?

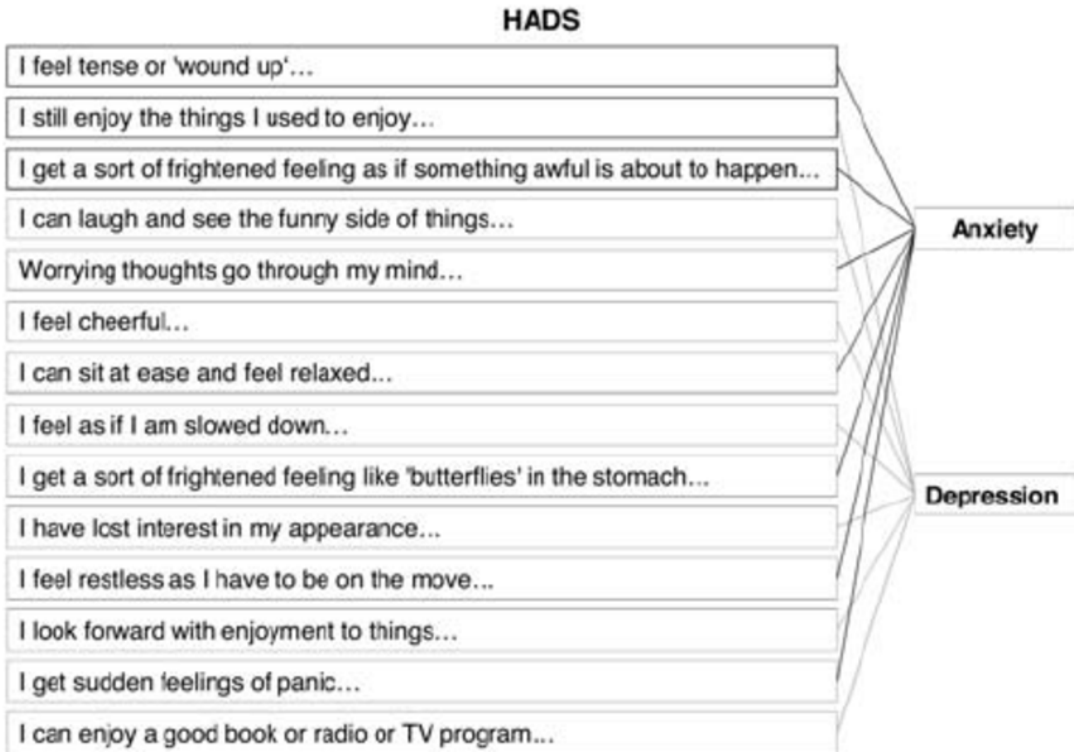


APGAR-score - Scale, index or profile?

Apgar Scoring System

Indicator		0 Points	1 Point	2 Points
A	Activity (muscle tone)	Absent	Flexed arms and legs	Active
P	Pulse	Absent	Below 100 bpm	Over 100 bpm
G	Grimace (reflex irritability)	Floppy	Minimal response to stimulation	Prompt response to stimulation
A	Appearance (skin color)	Blue; pale	Pink body, Blue extremities	Pink
R	Respiration	Absent	Slow and irregular	Vigorous cry

Hospital Anxiety and Depression Scale - Scale, index or profile?

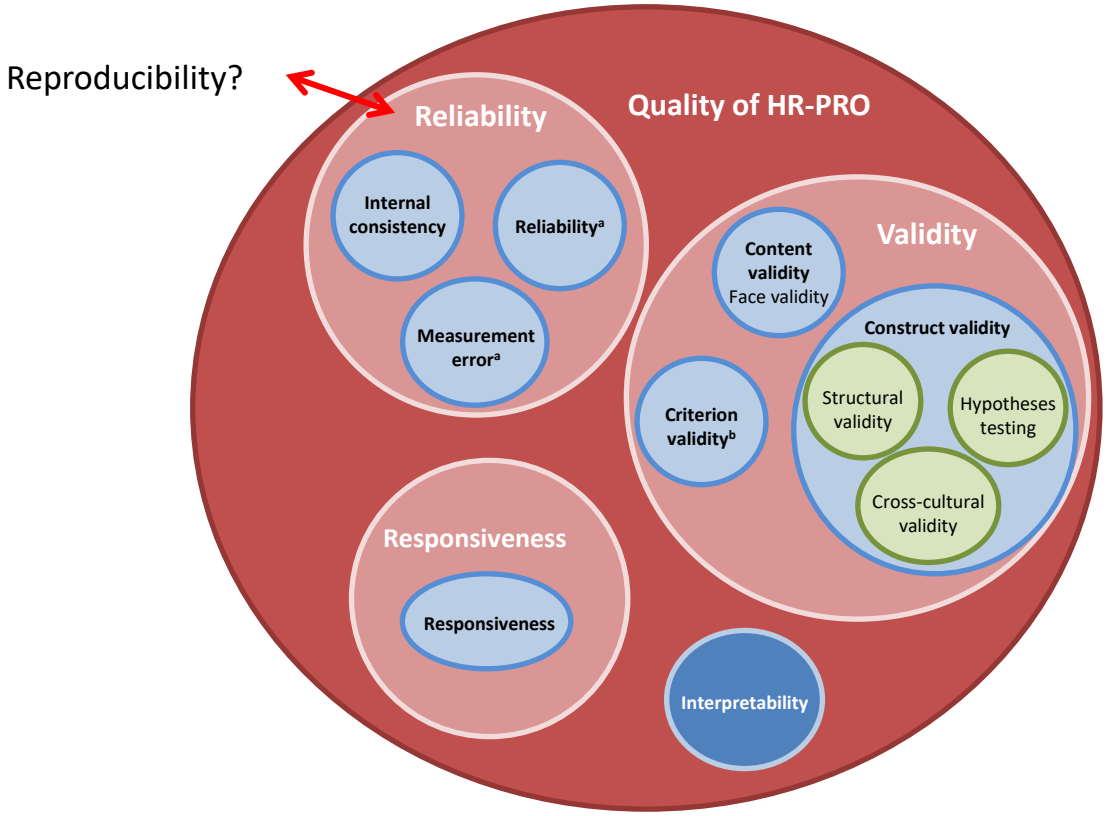


Agenda

Time	Contents	Pages	Teacher
9.00– 10.00	Introduction to questionnaires - conceptualisation	7-13, 33-35	AØH
10.00 – 11.00	Exercise: Conceptualisation		AØH
11.00 – 12.30	Concepts in questionnaire technique • Observed vs. latent variables • Conceptual frameworks • Reflective and formative models • Scales vs. indexes	13-17, 35-37	AØH
	COSMIN taxonomy and requirements to questionnaire validation	Articles on www	
12.30 – 13.00	*Lunch*		
13.00 – 13.45	How to measure the construct – operationalisation • Preparing items and answer categories	35-37 41-50	AØH
13.45 – 14.30	Exercise: Preparing items and answer categories		AØH
14.30 – 15.00	Structure and design of a questionnaire study Q & A about own study		AØH

* www = see www.clinimetrics.sdu.dk

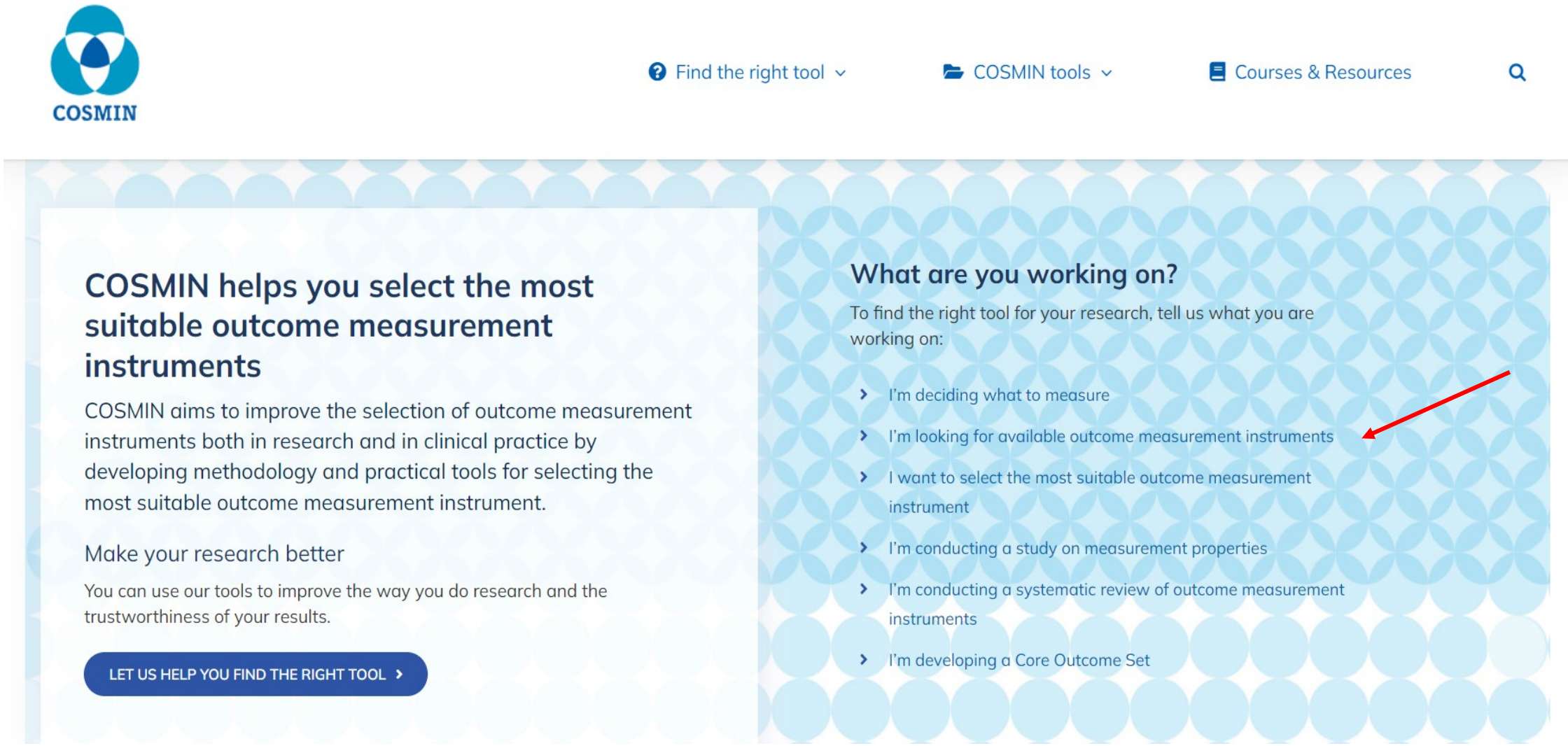
The COSMIN taxonomy

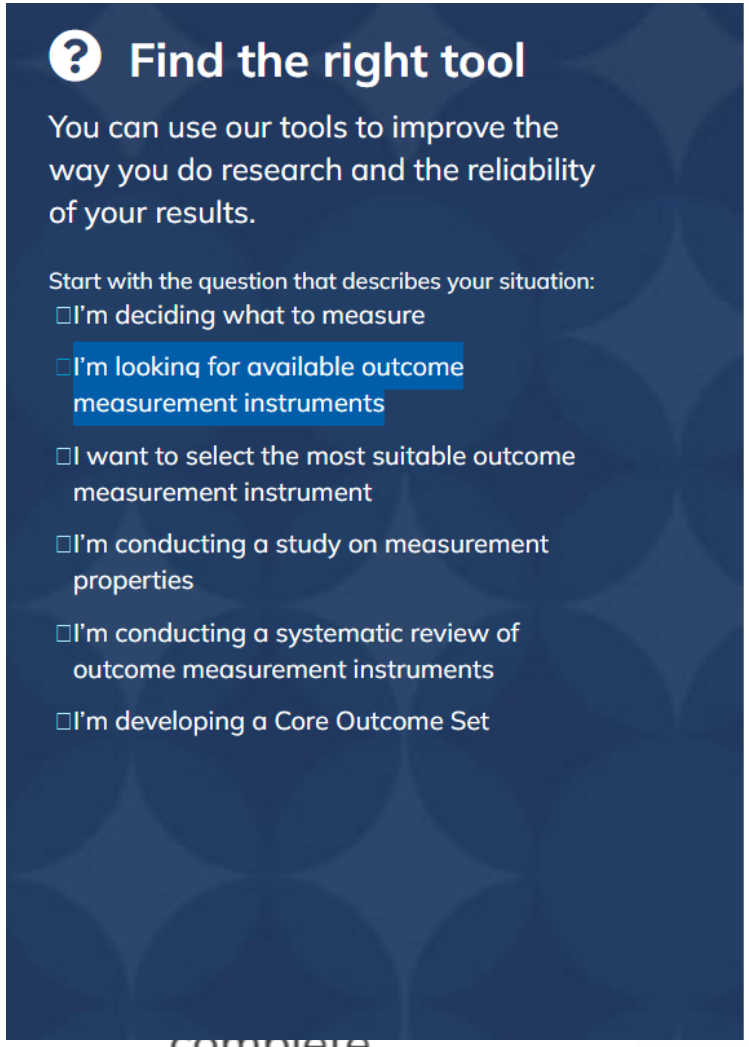


[a](test–retest, inter-rater, intra-rater)

[b](concurrent validity, predictive validity)

www.cosmin.nl





complete.

I'm looking for available outcome measurement instruments

You are searching for overviews of which outcome measurement instruments are available to measure a specific construct; e.g. as input information when deciding on what to measure and how to measure it.

Or you are searching for studies on the quality (i.e. measurement properties, such as reliability and validity) of a specific instrument, and need some help finding the appropriate studies. Different ways to find your information are available.

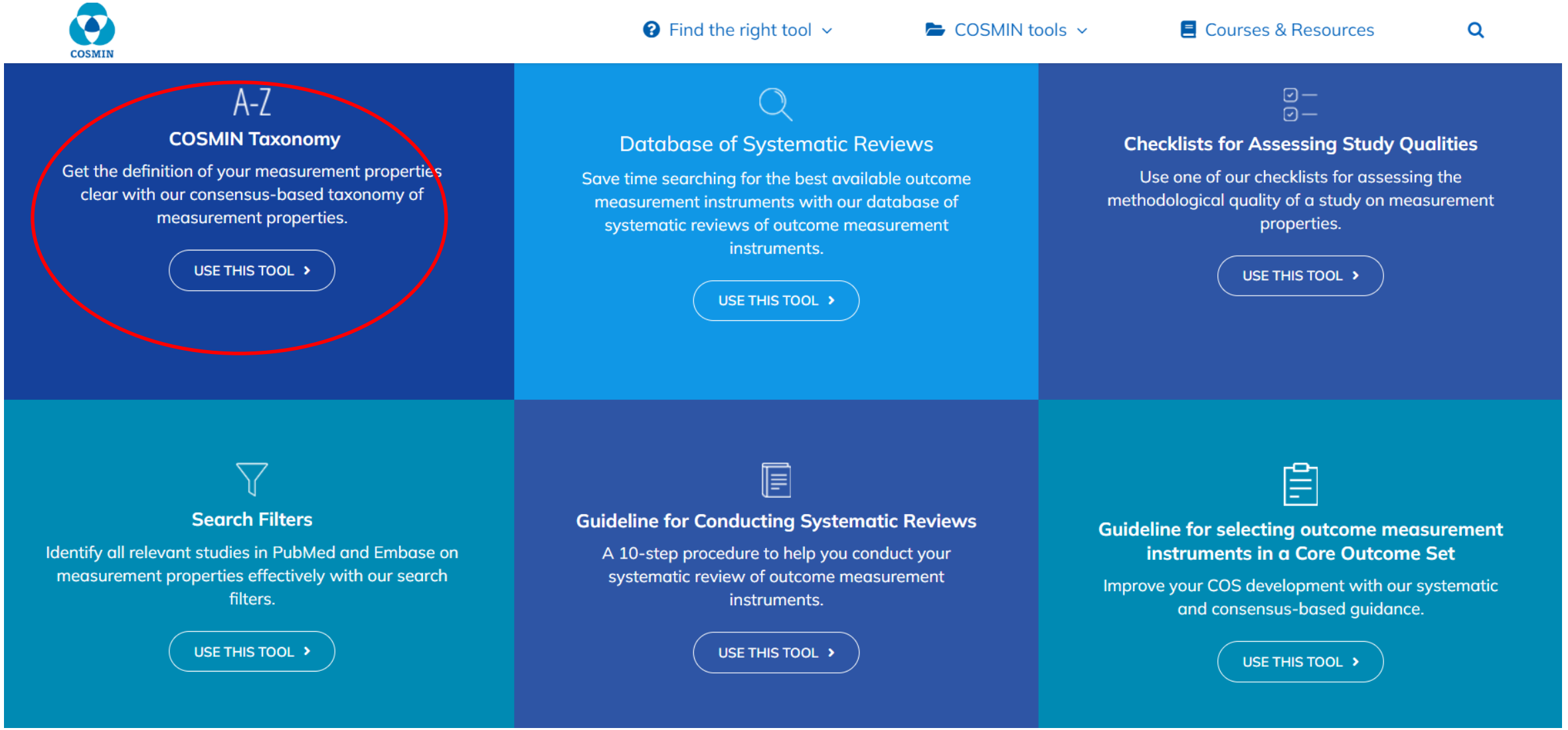
Different ways to find your information are available:

1 Systematic reviews of outcome measurement instruments provide overviews of available instruments and their quality for measuring specific constructs in specific populations. Have a look in the COSMIN Database for Systematic Reviews to find out whether the review you need is available.

2 Core Outcome Sets (COS), including a recommendation on the outcome measurement instruments that should be used to measure the core outcomes, are great resources when searching for outcome measurement instruments. COSMIN recommends to measure and report the outcomes and outcome measurement instruments upon which consensus was reached in a specific field in your trial.

To find out if a Core Outcome Set (COS) for your specific field has been developed, have a look in the COS database that is maintained by the COMET initiative at www.comet-initiative.org. For clinical practice purposes, Standard Sets developed by ICHOM can be accessed, available at www.ichom.org.

3 PROMIS® (Patient-Reported Outcomes Measurement Information System) is a set of efficient, reliable, and valid patient-reported outcome measurement instruments for measuring aspects of physical, mental, and social health, that can universally be applied to adults and children with or without (chronic) diseases. This enhances standardization of measurements across populations. PROMIS instruments are developed based on modern psychometric techniques (Item Response Theory) and can be used for Computerized Adaptive Testing, whereby patients get more relevant and less questions to complete.

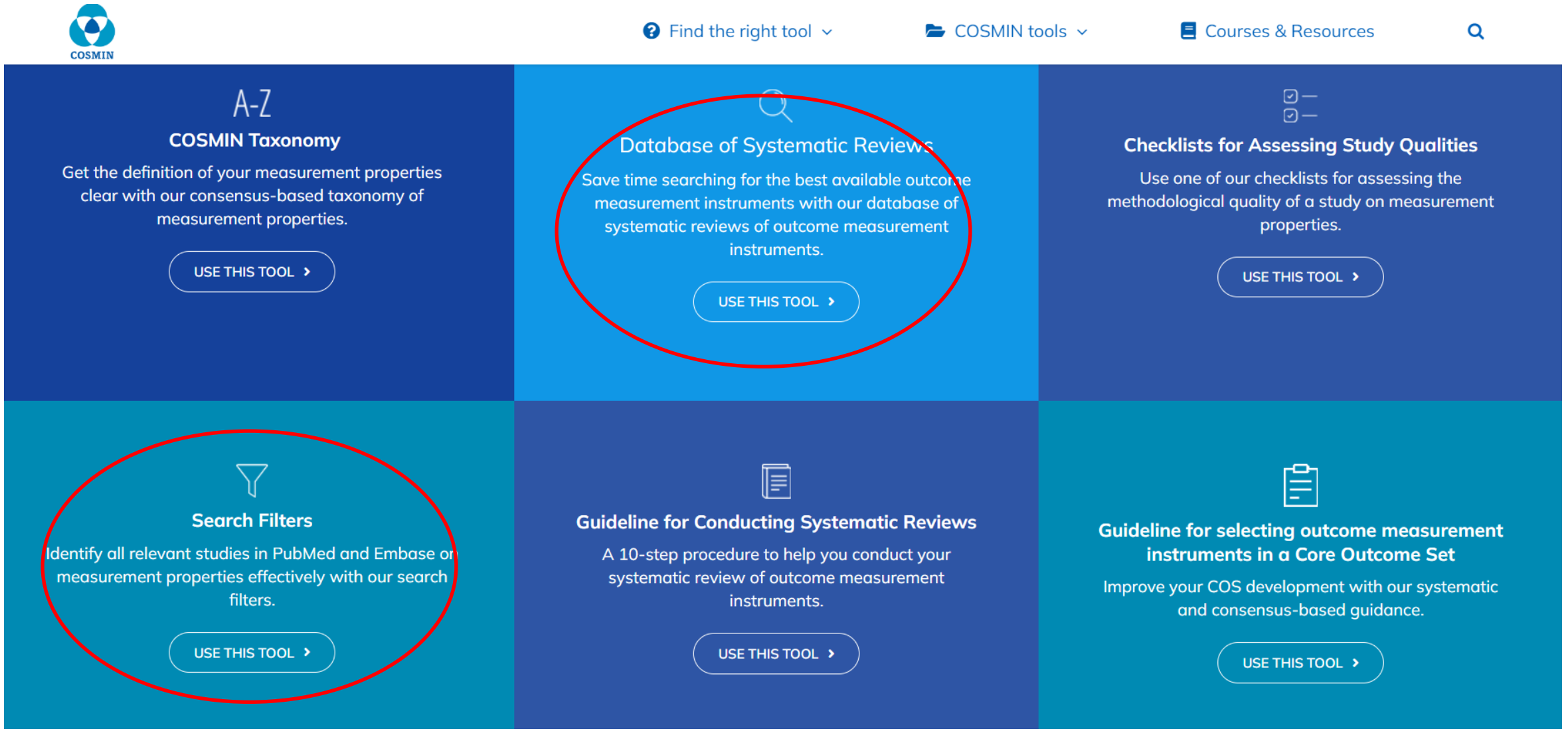

COSMIN
Find the right tool
COSMIN tools
Courses & Resources
A-Z
COSMIN Taxonomy
Get the definition of your measurement properties clear with our consensus-based taxonomy of measurement properties.
USE THIS TOOL
Database of Systematic Reviews
Save time searching for the best available outcome measurement instruments with our database of systematic reviews of outcome measurement instruments.
USE THIS TOOL
Checklists for Assessing Study Qualities
Use one of our checklists for assessing the methodological quality of a study on measurement properties.
USE THIS TOOL
Search Filters
Identify all relevant studies in PubMed and Embase on measurement properties effectively with our search filters.
USE THIS TOOL
Guideline for Conducting Systematic Reviews
A 10-step procedure to help you conduct your systematic review of outcome measurement instruments.
USE THIS TOOL
Guideline for selecting outcome measurement instruments in a Core Outcome Set
Improve your COS development with our systematic and consensus-based guidance.
USE THIS TOOL

The COSMIN taxonomy

Overview of definitions in COSMIN

COSMIN definitions of domains, measurement properties, and aspects of measurement properties

Term			Definition
Domain	Measurement property	Aspect of a measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true'† differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses *(for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups)* based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability*			Interpretability is the degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in scores.


COSMIN
Find the right tool
COSMIN tools
Courses & Resources
A-Z
COSMIN Taxonomy
Get the definition of your measurement properties clear with our consensus-based taxonomy of measurement properties.
USE THIS TOOL
Database of Systematic Reviews
Save time searching for the best available outcome measurement instruments with our database of systematic reviews of outcome measurement instruments.
USE THIS TOOL
Checklists for Assessing Study Qualities
Use one of our checklists for assessing the methodological quality of a study on measurement properties.
USE THIS TOOL
Search Filters
Identify all relevant studies in PubMed and Embase on measurement properties effectively with our search filters.
USE THIS TOOL
Guideline for Conducting Systematic Reviews
A 10-step procedure to help you conduct your systematic review of outcome measurement instruments.
USE THIS TOOL
Guideline for selecting outcome measurement instruments in a Core Outcome Set
Improve your COS development with our systematic and consensus-based guidance.
USE THIS TOOL

The COSMIN checklists

Checklists developed to

Improve the design of a study
Assess the methodological quality of a study
Improve the reporting of a study

3 checklists

1. COSMIN Study Design checklist
   Optimal study design and statistical method

2. COSMIN Risk of Bias checklist
   A list of standards to assess the methodological quality (design requirements and preferred statistical methods) of single studies included in systematic reviews of PROMs)

3. COSMIN Reporting checklist
   A guide on how to report studies on measurement properties.

An example

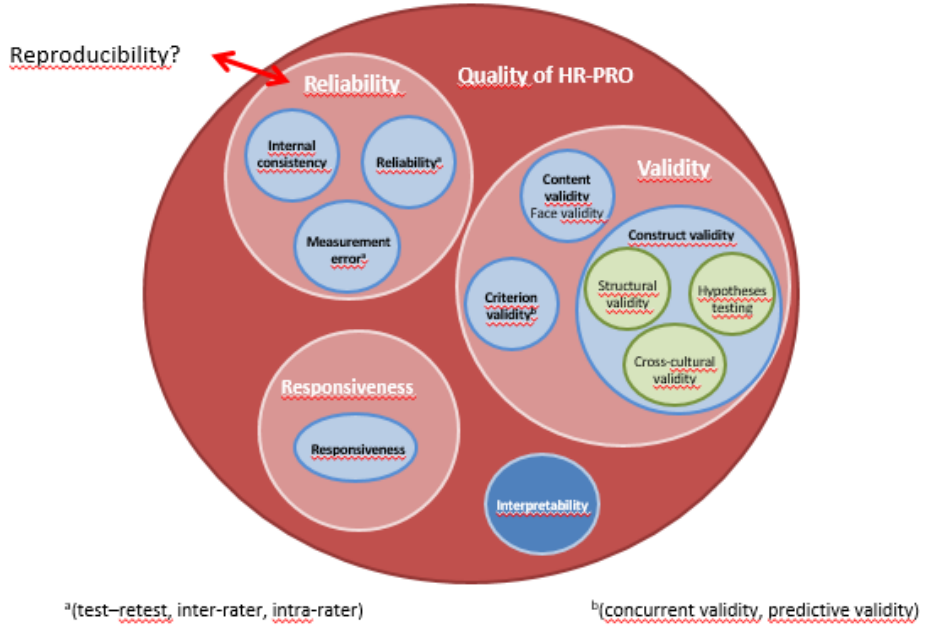
You want to measure ‘Internal consistency’ of a PROM in a study you are about to design

COSMIN

Domain: Reproducibility (reliability)

Measurement property: Internal consistency

Definition: ”The degree of interrelatedness among the items”



COSMIN Study Design Checklist

- Recommended for designing studies to evaluate measurement properties of **existing** PROMs
- Can also be used for designing protocols for developing **new** PROMs
- Linked to the Risk of Bias checklist
- **What are the design requirements for determining 'Internal consistency'?**

Internal consistency

Like structural validity, internal consistency is only relevant for PROMs based on a reflective model. Furthermore, internal consistency should be assessed for unidimensional (sub)scales. Therefore, unidimensionality or structural validity using e.g. factor analysis should be assessed for each scale or subscale in the study or evidence for structural validity obtained in a previous study in a sample from a comparable target population should be available.

Important introductory information on each measurement property

Internal consistency	very good	adequate	doubtful	inadequate	NA	Justification
Design requirements						
1 Check whether a scale or a subscale is unidimensional	Evidence provided that each scale or subscale is unidimensional		Unclear whether each scale or subscale is unidimensional	the scale or subscale is NOT unidimensional		RoB Box 4 (1)
2 Perform the analysis in a sample with an appropriate number of patients (taking into account expected number of missing values)	≥100 patients	50-99 patients	30-49 patients	<30 patients		Sample size
3 Provide a clear description of how missing items will be handled	The way missing items will be handled is clearly described		The way missing items will be handled is not clearly described			Original CC
Statistical methods						
4 For continuous scores: calculate Cronbach's alpha or Omega for each unidimensional scale or subscale	Cronbach's alpha, or Omega will be calculated		Only item-total correlations will be calculated	No Cronbach's alpha and no item-total correlations will be calculated	Not applicable	RoB Box 4 (2)
5 For dichotomous scores: calculate Cronbach's alpha or KR-20 for each unidimensional scale or subscale	Cronbach's alpha or KR-20 will be calculated		Only item-total correlations will be calculated	No Cronbach's alpha or KR-20 and no item-total correlations will be calculated	Not applicable	RoB Box 4 (3)

NB: Item 6 has been removed due to limited space

Justification for choices:

RoB – Risk of Bias checklist
CC – original COSMIN Checklist[1]
Sample size considerations
Newly added standard

[1] MokkinkLB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4): 539-49

The COSMIN checklists

3 checklists

1. COSMIN Study Design checklist
   Optimal study design and statistical method

2. COSMIN Risk of Bias checklist
   A list of standards to assess the methodological quality (design requirements and preferred statistical methods) of single studies included in systematic reviews of PROMs)

3. COSMIN Reporting checklist
   A guide on how to report studies on measurement properties.

COSMIN Risk of Bias Checklist

Most often used in systematic reviews of the quality of a given PROM

Standards

- Refers to design requirements and preferred statistical methods for evaluating the ***methodological quality of studies*** on measurement properties

Criteria

- Refers to what constitutes good measurement properties – ***quality of the PROM***

Quality of the study (the standards)

- Each measurement property has a COSMIN box containing all standards
- 10 boxes for good methodological quality of development and measurement properties:

1. *PROM development**
2. *Content validity*
3. *Structural validity*
4. *Internal consistency*
5. *Cross-cultural validity/Measurement invariance*
6. *Reproducibility (reliability)*
7. *Measurement error*
8. *Criterion validity*
9. *Hypotheses testing for construct validity*
10. *Responsiveness*

1 on development of a PROM

9 measurement properties

* Taken into account when evaluating content validity

Prinsen et al, 2018; Mokkink et al., 2017; Terwee et al., 2018

The COSMIN Risk of Bias checklist

	COSMIN Risk of Bias Checklist
	Box 1. PROM development
	Box 2. Content validity
	Box 3. Structural validity
X	Box 4. Internal consistency
	Box 5. Cross-cultural validity/Measurement invariance
	Box 6. Reproducibility (reliability)
	Box 7. Measurement error
	Box 8. Criterion validity
	Box 9. Hypotheses testing for construct validity
	Box 10. Responsiveness

- *Domain:* Reproducibility (reliability)
- *Measurement property:* Internal consistency
- *Definition:*

”The degree of interrelatedness among the items”

Box 4. Internal consistency

Box 4. Internal consistency

Does the scale consist of effect indicators, i.e. is it based on a reflective model? [1] yes / no

	very good	adequate	doubtful	inadequate	NA
Design requirements					
1 Was an internal consistency statistic calculated for each unidimensional scale or subscale separately?	Internal consistency statistic calculated for each unidimensional scale or subscale		Unclear whether scale or sub scale is unidimensional	Internal consistency statistic NOT calculated for each unidimensional scale or sub scale	
Statistical methods					
2 For continuous scores: Was Cronbach's alpha or omega calculated?	Cronbach's alpha, or Omega calculated		Only item-total correlations calculated	No Cronbach's alpha and no item-total correlations calculated	Not applicable
3 For dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	Cronbach's alpha or KR-20 calculated		Only item-total correlations calculated	No Cronbach's alpha or KR-20 and no item-total correlations calculated	Not applicable
4 For IRT-based scores: Was standard error of the theta (SE (θ)) or reliability coefficient of estimated latent trait value (index of (subject or item) separation) calculated?	SE(θ) or reliability coefficient calculated			SE(θ) or reliability coefficient NOT calculated	Not applicable
Other					
5 Were there any other important flaws in the design or statistical methods of the study?	No other important methodological flaws		Other minor methodological flaws	Other important methodological flaws	

Quality of the PROM (the criteria)

Internal consistency	+	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or subscale[6]
	?	Criteria for "At least low evidence[4] for sufficient structural validity[5]" not met
	–	At least low evidence[4] for sufficient structural validity[5] AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale[6]

+	= sufficient
?	= indeterminate
-	= insufficient

[4] As defined by grading the evidence according to the GRADE approach
[5] This evidence may come from different studies

COSMIN systematic review manual for PROMs (Table 4)

GRADE approach

Quality level	Definition
High	We are very confident that the true measurement property lies close to that of the estimate* of the measurement property
Moderate	We are moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different
Low	Our confidence in the measurement property estimate is limited: the true measurement property may be substantially different from the estimate of the measurement property
Very low	We have very little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property

Grading the quality of the evidence

Summary of study quality (standards) AND PROM quality (criteria)

Standards

Refers to design requirements and preferred statistical methods for evaluating the ***methodological quality of studies*** on measurement properties

Criteria

Refers to what constitutes good measurement properties – ***quality of the PROM***

COSMIN and your project

Have you used it?

Will you use it?

Join the COSMIN network

Lunch