How can bibliometrics operationalise sociological concepts?

1. Perspective and problem
2. The concept and process of operationalisation
3. Operationalisations in the history of bibliometrics
4. Example 1: Reconstructing topics from networks of papers
5. Example 2: Measuring research quality
6. Conclusions

1. Perspective and problem

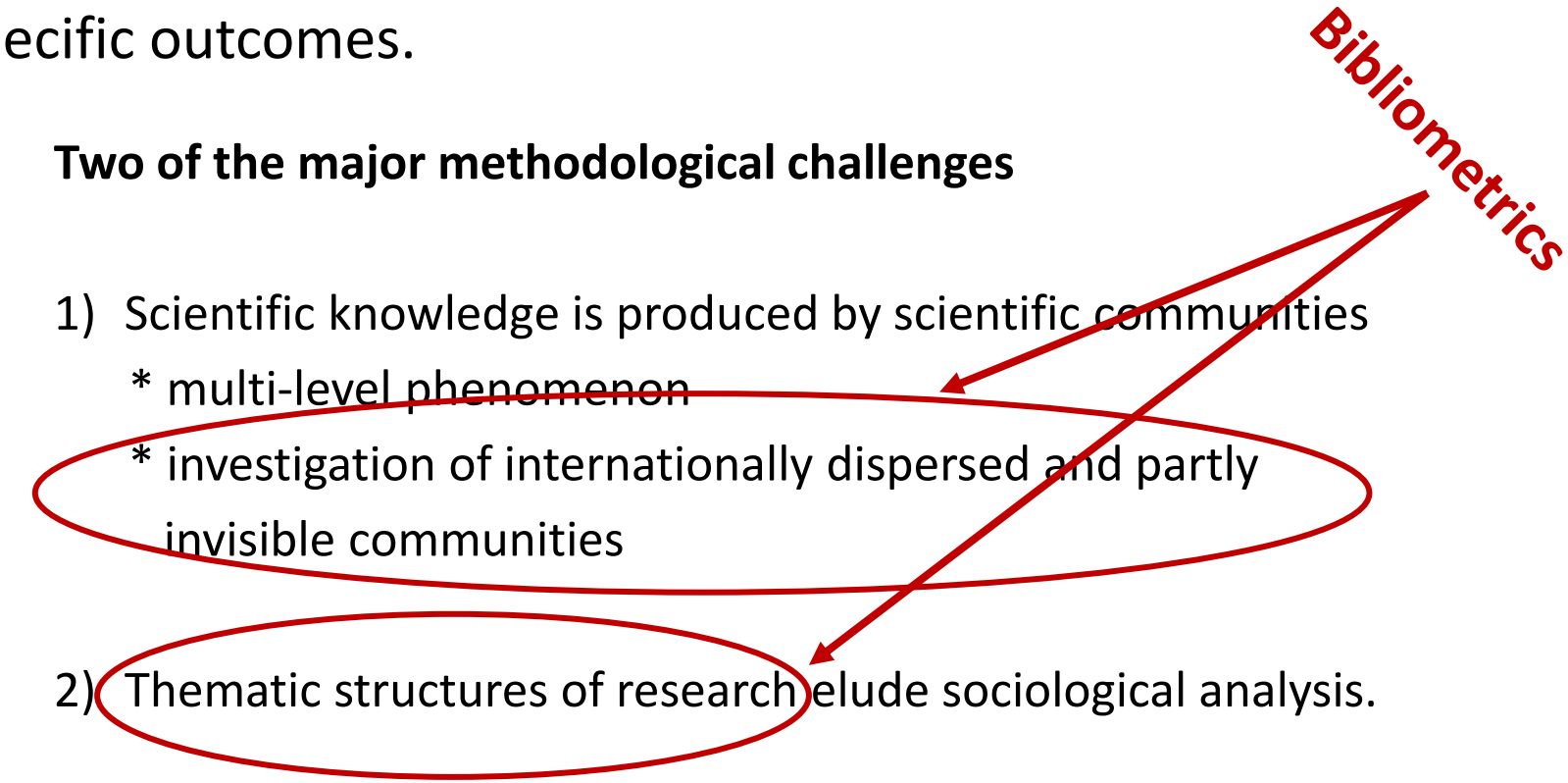
Sociology of Science:

Strives to explain the connection between conditions under which scientific* knowledge is produced and the content and properties of that knowledge by identifying the social mechanisms that operate under specific conditions and produce specific outcomes.

Two of the major methodological challenges

1) Scientific knowledge is produced by scientific communities
   * multi-level phenomenon
   * investigation of internationally dispersed and partly invisible communities

2) Thematic structures of research elude sociological analysis.



* Including knowledge production in the sciences, social sciences and humanities.

Bibliometric methods can be of invaluable help to the sociology of science …

… if they operationalise sociological concepts

If they don‘t …

… the **sociology of science and science studies in general** are deprived of valuable methods.

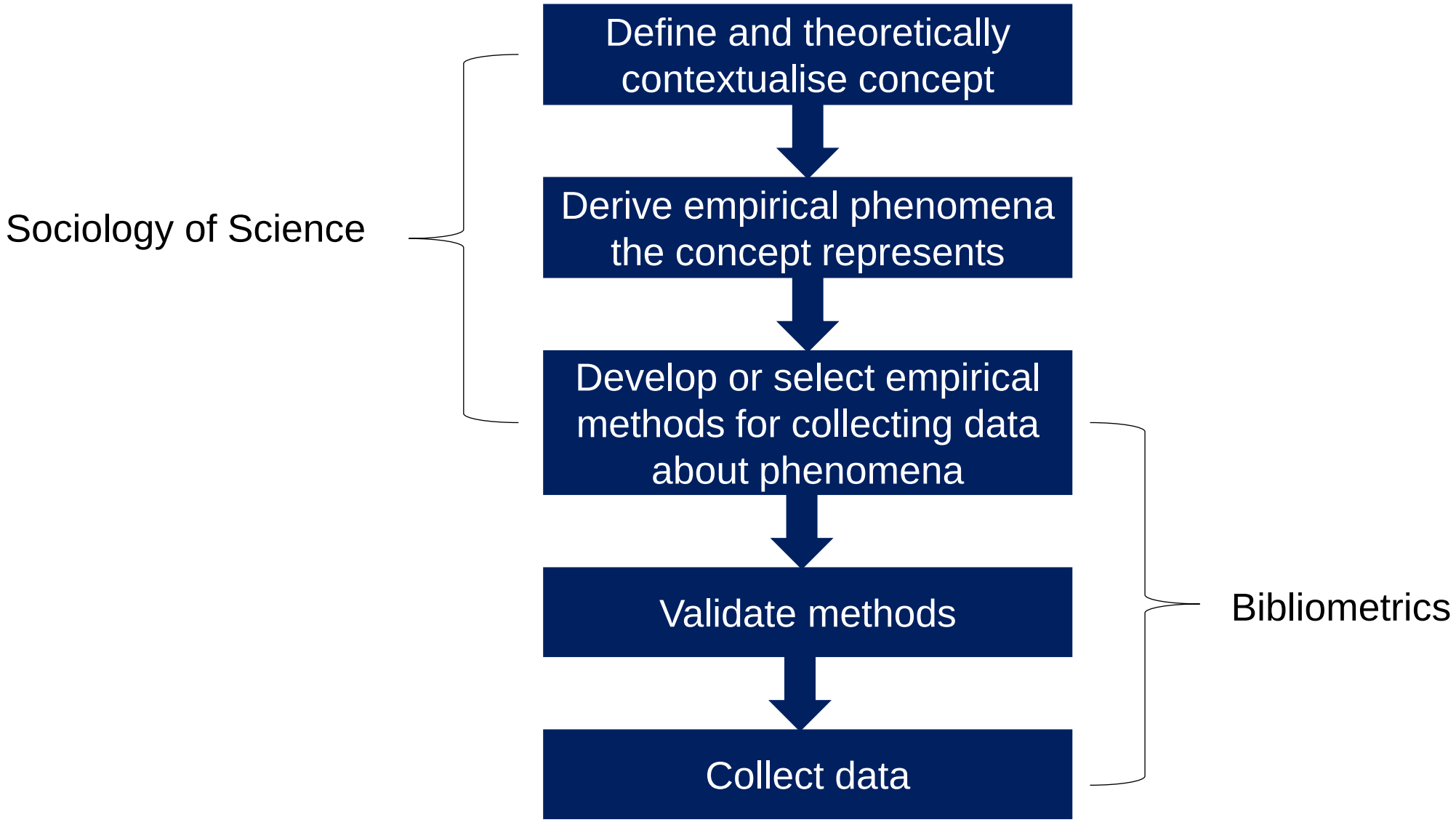
… bibliometrics is deprived of

- theoretical foundations,
- criteria for the assessment of methods,
- opportunities to be useful, and
- one source of legitimation.

2. The concept and process of operationalisation

Definition:

Operationalising a concept means deriving procedures for identifying the empirical phenomena the concept describes.


Define and theoretically contextualise concept
Derive empirical phenomena the concept represents
Develop or select empirical methods for collecting data about phenomena
Validate methods
Collect data
Sociology of Science
Bibliometrics

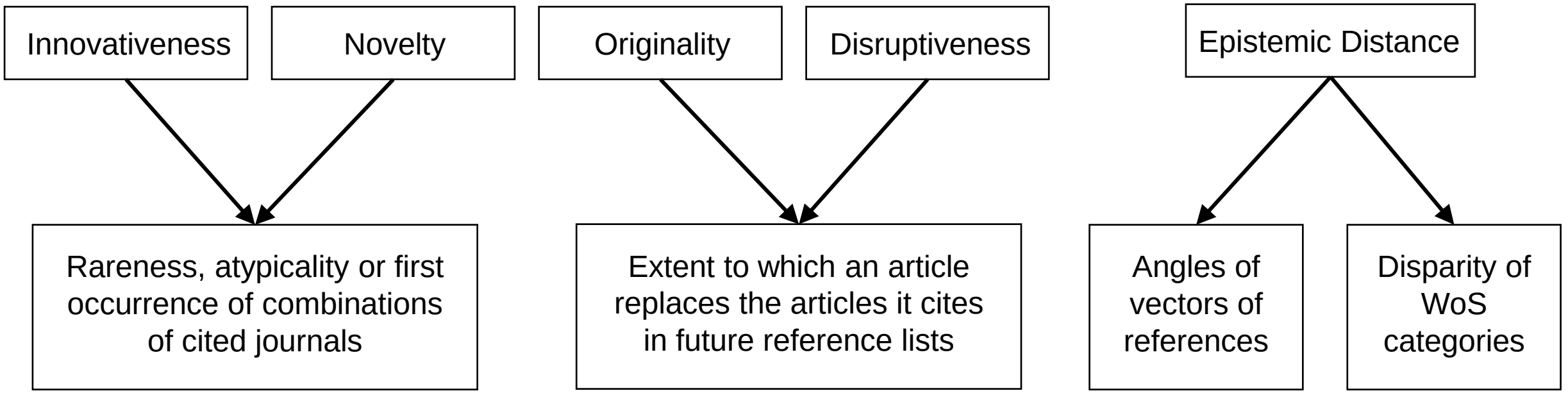
3. Operationalisations in the history of bibliometrics

- Citations as partial indicators of quality and recognition (Cole and Cole 1967)
- Attempts to empirically identify the Matthew Effect (Cole 1970)
- Attempts to (dis)prove the Newton/Ortega hypotheses (e.g. Cole and Cole 1987)
- Tracing the use of knowledge over time (Cozzens 1985)
- Correspondence between historical reonstruction of a theory selection process and changing citation patterns (Pickering and Nadel 1987)
- Co-authorships as partial indicators of research collaboration (Laudel 2001)
- Attempts to operationalise epistemic diversity (Rafols et al. since 2007)
- Attempt to operationalise the degree of codification of knowledge (Fanelli and Glänzel 2013)
- Bibliometric identification of knowledge transfer (Aman 2020)

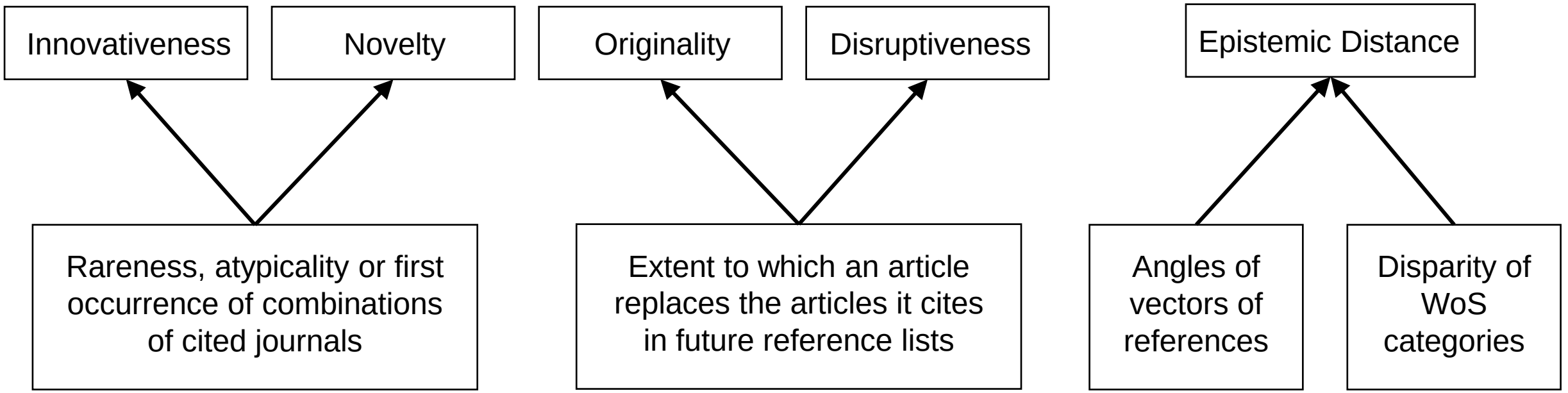
Two general problems:

- Bibliometric indicators are often partial operationalisations of sociological concepts, and their utility is difficult to determine without a full operationalisation.
- Bibliometric indicators may conflate dimensions of sociological concepts that need separate measurement.

Counterexamples



Counterexamples



4. Example 1: Reconstructing topics from networks of papers

Define concept

* Does not usually happen

* Common approach: apply a method and declare the outcome to be a topic („operational definition“)

* Choice of method not guided by the need to collect specific data

This sounds bad.

But it seems justified because there is no ready-made definition of „topic“ in the sociology of science.

This sounds bad.

However, there is a rich old literature on the link between thematic and social structures in science, from which definitions can be constructed.

Define concept

A topic is a focus on theoretical, methodological or empirical knowledge that is shared by a number of researchers and provides these researchers with a joint frame of reference for the formulation of problems, the selection of methods or objects, the organisation of empirical data, and the interpretation of data. (Havemann et al. 2017)

This needs to be developed! Don‘t operationalise one-sentence-definitions!

Derive empirical phenomena the concept represents

Shared focus on knowledge -> Knowledge that is in the focus.

= Part of scientific knowledge that a number of researchers perceive as jointly developing (extending and restructuring).

- Relatively full communication between researchers addressing the topic (Kuhn)
- High thematic similarity of publications addressing the topic
- Perceived differently by researchers working and not working on the topic
- Topics overlap in publications, researchers, journals, organisations …
- Topics are differently structured depending on the content of knowledge

(Held and Gläser in preparation)

Develop or select empirical methods for collecting data about phenomena

Reconstruction of topics in sets of publications

- Bibliometric data models as partial operationalisations:

* Direct citation (predominantly communication)
* Bibliographic coupling (predominantly thematic similarity)
* Text-based similarities (predominantly thematic similarity)

Conflation of dimensions

- Algorithms (here only for detection of communities in networks)

* Priority of local information
* Allow for pervasive overlaps
* Allow for structural variation.

Validate methods

No valid topic

No validation of topic reconstruction methods

No valid method of topic reconstruction

Validate methods

Currently popular workarounds (convergent validity):

Present the topics you found to researchers and ask their opinion: “Does this make sense to you?“
(e.g. Klavans, Boyack and Small 2012)

Highly plausible other methods as „gold standards“
(e.g. Klavans and Boyack 2017).

Validation for library and information science
(e.g. Šubelj, van Eck and Waltman 2016).

5. Example 2: Measuring research quality

Define concept

* Does not usually happen

* Common approach: start from bibliometric indicators as quantitative measures of „an important aspect of quality, namely international impact” (van Raan 1996: 404).

* Choice of method not guided by the need to collect specific data

This sounds bad.

However, there is no ready-made definition of „research quality“ in the sociology of science.

This sounds bad.

The sociology of science has so far not even tried to define the concept.

Which is worse.

Define concept

Common approach: avoid definition.

1) Ask your respondents

“What are criteria for good research of scholars in English literature, German literature and art history?” (Hug et al. 2013: 370)

2) Beat about the bush

“An overview of the, otherwise diverse, literature on quality yields three attributes of research considered important for the consensus of what is ‘good’ research. These are originality/novelty, plausibility/reliability, and the value or usefulness of the research.“ (Langfeldt et al. 2020: 120)

Define concept

Collective frame of a scientific community that organises the assessment of the utility of epistemic objects for knowledge production processes.

(Gläser submitted)

This needs to be developed! Don‘t operationalise one-sentence-definitions!

Derive empirical phenomena the concept represents

Collective interpretive scheme -> Properties of epistemic objects

* Epistemic object of assessment (researchers, publications, reviews, materials, data, software …)

* Functions of the object in the communal knowledge production process

* „Usability“ of object: How well does it fulfil these functions?

Develop or select empirical methods for collecting data about phenomena

Example: Measuring the quality of publications

Utilise enactments of collective frame by reviewers (e.g. open peer review)

Utilise enactments of the collective frame by users of a publication

Citation-based indicators: publication has been used in some of the instances of citation (and in non-cited instances)

Conflation of dimensions

Citation context analysis:
- used for what,
- assessment of utility
- assessment of reliability,
- other criteria (aesthetic)

Validate methods

Convergent validity again:

Citation-based indicators compared to each other, to peer review outcomes, to academic prizes …

6. Conclusions

Observations

Today, operationalisation rarely happens due to a combination of bibliometrics‘ indifference with the current theoretical weakness of the sociology of science.

This makes it difficult for bibliometrics to be theoretically useful.

4. Conclusions

Suggestions

If it wants to be an operationalisation (i.e. if it wants to be useful for science studies), bibliometric measurement must include or link to „upstream“ work.

Since the sociology of science doesn‘t deliver, it would be useful to provide an offer by finding out what bibliometric indicators could be an operationalisation of. (See upcoming workshop series „Bibliometric Indicators of Epistemic Change“)

This works best as an interdisciplinary research effort, not least because bibliometrics offers only partial operationalisations of sociological concepts.

New Years Resolution: I will validate my new methods before I use them to produce knowledge claims.