The Art and Science of Transportation Research in the AI Era

Key Concepts in Research Design and Methodology

Meng Cai

Learning goals

#1 Understand how to formulate a research question

#2 Understand key concepts in research design and methodology

#3 Understand research limitations

What we did in previous weeks

https://transport.ec.europa.eu/news-events/news_en

✊ **DO THIS:** Here is an example project. We would like to find out what the European Union has done recently (let's say since 2023) to advance sustainable mobility and transport. One possible data source is the news we just scraped, but we need more information other than the title of the news.

#1 What is a research question

- A clear, focused, and answerable question that guides the research process.

- Our example:
  - **General question**: What has the European Union done recently to advance sustainable mobility and transport?
  - **Research question**: How do the European Union's mobility and transport initiatives align with the United Nations Sustainable Development Goals?

- Broad and exploratory
- Initiate an inquiry

- Precise
- Guide the research

#2 What is an operational definition

- Operational definitions clarify how abstract concepts are measured and observed within the context of a study.

- Our example:
  - Mobility and transport: The projects, policies, programs, or actions related to mobility and transport as reported in the news articles.
  - Sustainability: the United Nations Sustainable Development Goals (no poverty, zero hunger, good health and well-being, etc).

image source: https://ifapa.net/new-un-enable-webpage/sdgwheel1/

#3 What is a unit of study

- It is the primary element that is observed, analyzed, or measured in a study.
- Our example:
  - The unit of study is each individual news article (specifically, its title and short description) published on the European Union's mobility and transport news website.

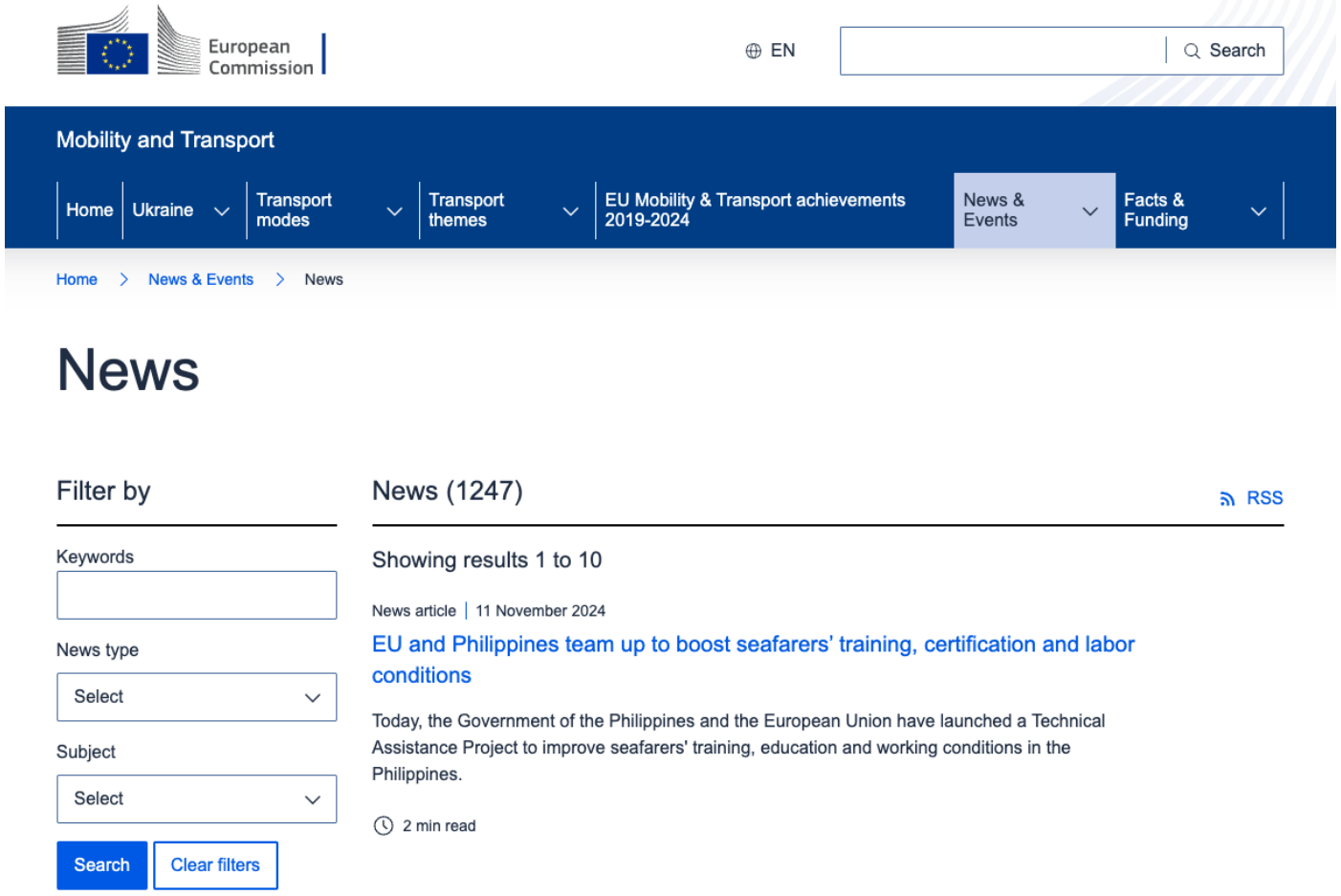
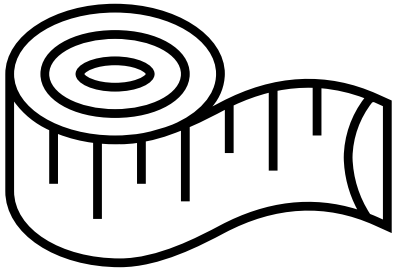


image source: https://transport.ec.europa.eu/news-events/news_en

#4 What is measurement



- Measurement is the systematic process of assigning numbers or labels to objects, events, or characteristics according to specific rules. In other words, it's about deciding *what* we are going to count, observe, or record in a study and then doing so in a systematic way. This process allows us to quantify abstract concepts based on their operational definitions.

- Our example:

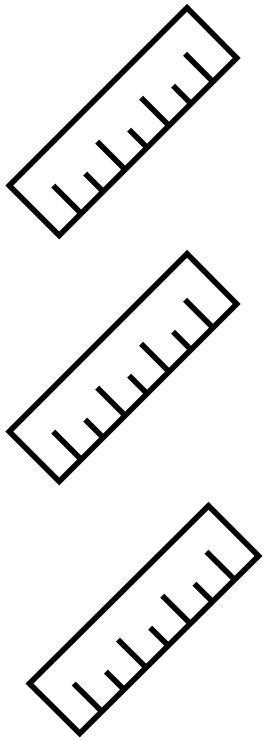
  - The measurement is the frequency of each SDG appearing in the news titles and descriptions on the EU's transport website.

#5 What is reliability

- Reliability is the consistency of a measure; if a study is repeated under the same conditions, it should yield the same results.

- Our example:

  – Reliability would be achieved if seesus, the computational tool we used, consistently labels the same news items with the same SDGs every time it is applied.



#6 What is validity

- Validity is the degree to which a study accurately captures what it intends to measure. It's about how "true" or accurate the findings are concerning the real-world concepts being studied.

- Our example:

  – Validity means that the methodology we used to identify alignment with the SDGs accurately reflects the EU's actual priorities in mobility and transport initiatives.

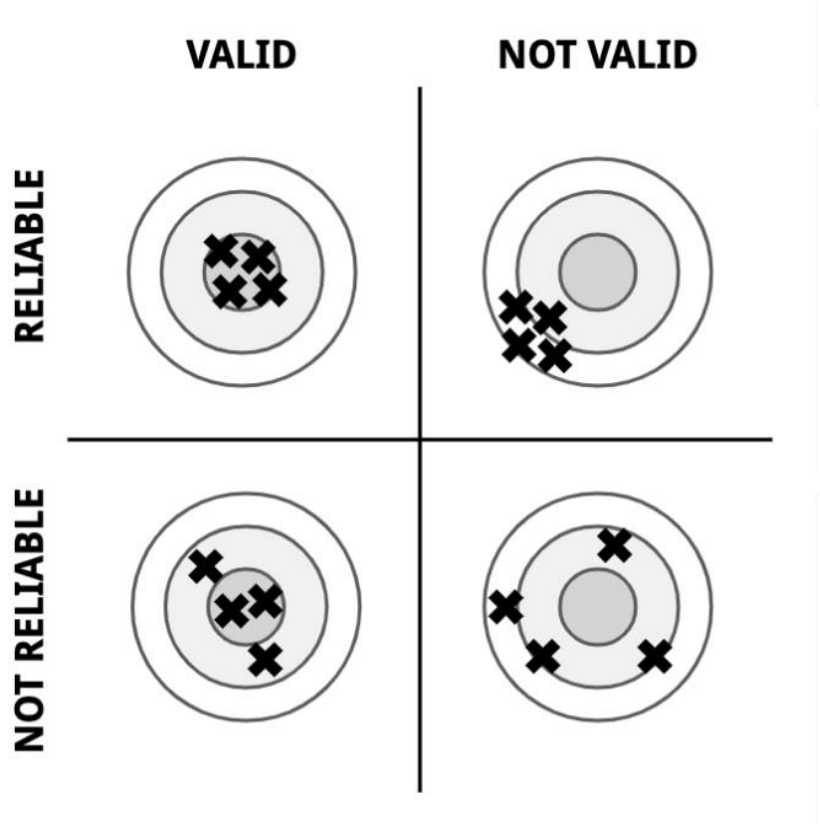


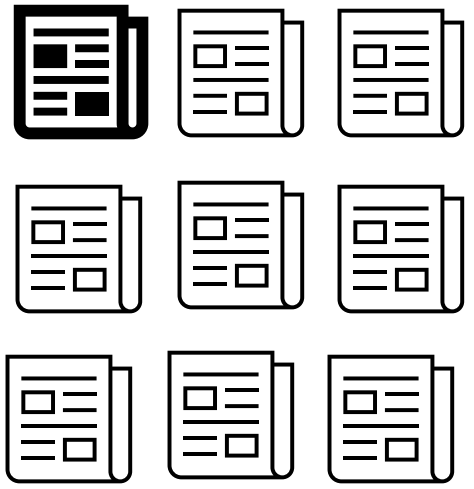
image source: https://nesslabs.com/validity-reliability-mental-models

#7 What is generalizability

- Generalizability is the extent to which the findings of a study apply to contexts beyond the specific data or sample analyzed. It is about how well the results might represent a larger population or different settings.

- Our example:

  – Generalizability depends on whether the news data from 2023 and 2024 represents the EU's overall transport and sustainability goals beyond this specific time frame (i.e. if these years are reflective of longer-term patterns in EU policy).



#8 What is bias

- Bias refers to **systematic** errors in a study that can lead to incorrect or skewed results. Bias can come from the data, analysis methods, or researcher expectations.
- Our example:
  - Bias occurs if certain SDGs are more likely to be identified than others because of the matching algorithms of seesus. In addition, if the news focuses more on the positive aspects, the results are biased.

#9 What is a limitation

- A limitation refers to any factors or constraints that impact the study's design, implementation, or interpretation. Limitations are factors that researchers cannot fully control, but that may affect the validity, reliability, or generalizability of the findings.

- Our example:

  – The data is limited to the news articles available on the transport.ec.europa.eu website, which may not cover all EU activities related to sustainable mobility and transport. In addition, the reliance on published news articles may introduce a bias toward topics deemed newsworthy by the EU's communication department. Furthermore, the analysis is constrained by the depth of the information in the news, which may not capture the full scope of all initiatives.

Notice

The deadline for submitting the final assignment (i.e., a one-page research proposal) is

February 21, 2025

Source: https://x.com/John_Attridge/status/1815970347965165649