CSE 599K
Empirical Research Methods

Winter 2025

Qualitative vs. Quantitative Research

Today

- More on artifact evaluation and replication
- Qualitative vs. quantitative research

Artifact evaluation and replication

Artifact evaluation and Replication

- Analysis grounded in a **conceptual model?**
- Clear **operationalization (implementation)?**
- **Implementation consistent with** the **model?**
- **Proper** use of **statistical methods?**
- Data interpreted in **context** of **prior knowledge?**
- Explored and validated **alternative hypotheses?**
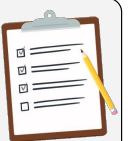
Transparency is key
- Transparent decision making (data collection and analysis)
- Shared instructions, data, and analyses (scripts)

Artifact evaluation and Replication

Transparency is key
- Transparent decision making (data collection and analysis)
- Shared instructions, data, and analyses (scripts)

What is the purpose of artifact evaluations?

https://www.acm.org/publications/policies/artifact-review-and-badging-current

Repeatability, reproducibility, and replicability
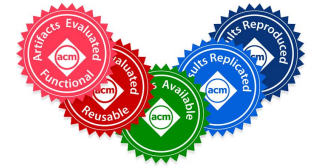
- **Repeatability**
  - Same research questions
  - Same experimental setup and artifacts
  - Same team
- **Reproducibility**
  - Same research questions
  - Same experimental setup and artifacts
  - Different team
- **Replicability**
  - Same research questions
  - Different experimental setup and artifacts
  - Different team

Note: the ACM defined replicability and reproducibility in the opposite way of most other scientific fields … now fixed!

Artifact badges

Pre-publication (You)

Post-publication (Others)

Artifact badges

	Repeated	Reproduced	Replicated
Team	same	different	different
Artifact	same	same	different

The purpose of artifact evaluations

ACM policies (https://tinyurl.com/352st3vm)

Highlight topics and inconsistencies for discussion/resolution
- Read *Overview*
- Skim *Badging*
- Read *Review Procedures*

The purpose of artifact evaluations

ACM policies (https://tinyurl.com/352st3vm)

Discussion
- 3 independent categories
  - *Evaluated*
  - *Available*
  - *Validated*
- Evaluation criteria
- Focus on Reproducibility
- …

Artifact evaluations: the good, the bad, and the ugly

The good
- Lots of **sharing and transparency** (data availability is now an expectation).
- **Rose festival** and **reproducibility** (RENE) **tracks.**
- Some venues **invite replication studies** (as technical papers).

The bad
- **Artifacts** remain largely an **afterthought**.
- Lots of **overhead** (artifact eval) and **questionable focus** (reproducibility).
- **Little progress** on replicability.

The ugly
- **Incentives**: Replicability isn't valued.
- **False sense of security** (artifact vs. conclusions).
- **Specification crisis**: emphasis is on the implementation, not the design.

The role of peer review, artifacts, and replication

- Analysis grounded in a **conceptual model?**
- Clear **operationalization (implementation)?**
- **Implementation consistent with** the **model?**
- **Proper** use of **statistical methods?**
- Data interpreted in **context** of **prior knowledge?**
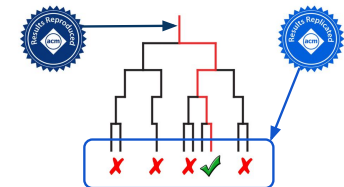- Explored and validated **alternative hypotheses**?
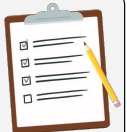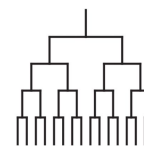
Design space   **Reported design**   **Reproduction/Replication**

Qualitative vs. Quantitative Research

Qualitative vs. quantitative research*

Define qualitative and quantitative research
- What are the characteristics of each methodology?
- What are examples for each methodology?
- What is the purpose/applicability of each methodology?
- What is the relationship of each methodology to deductive/inductive reasoning?

Define mixed-methods research
- What are examples for mixed-methods research?

* https://tinyurl.com/536az6tj