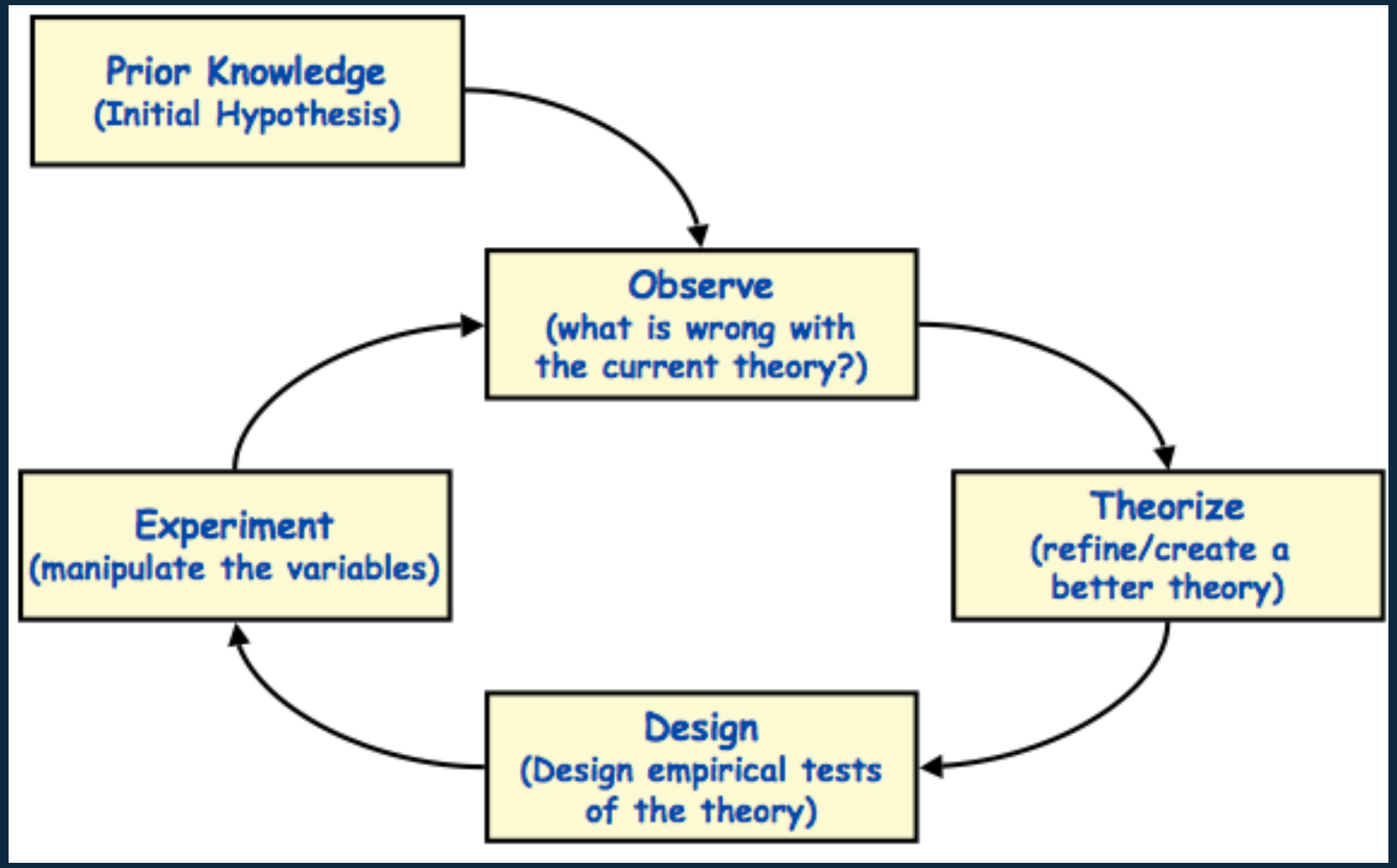
Empirical Studies to Evaluate Software Design

Sep 12, 2024


Prior Knowledge
(Initial Hypothesis)
Observe
(what is wrong with the current theory?)
Theorize
(refine/create a better theory)
Design
(Design empirical tests of the theory)
Experiment
(manipulate the variables)

Research Questions

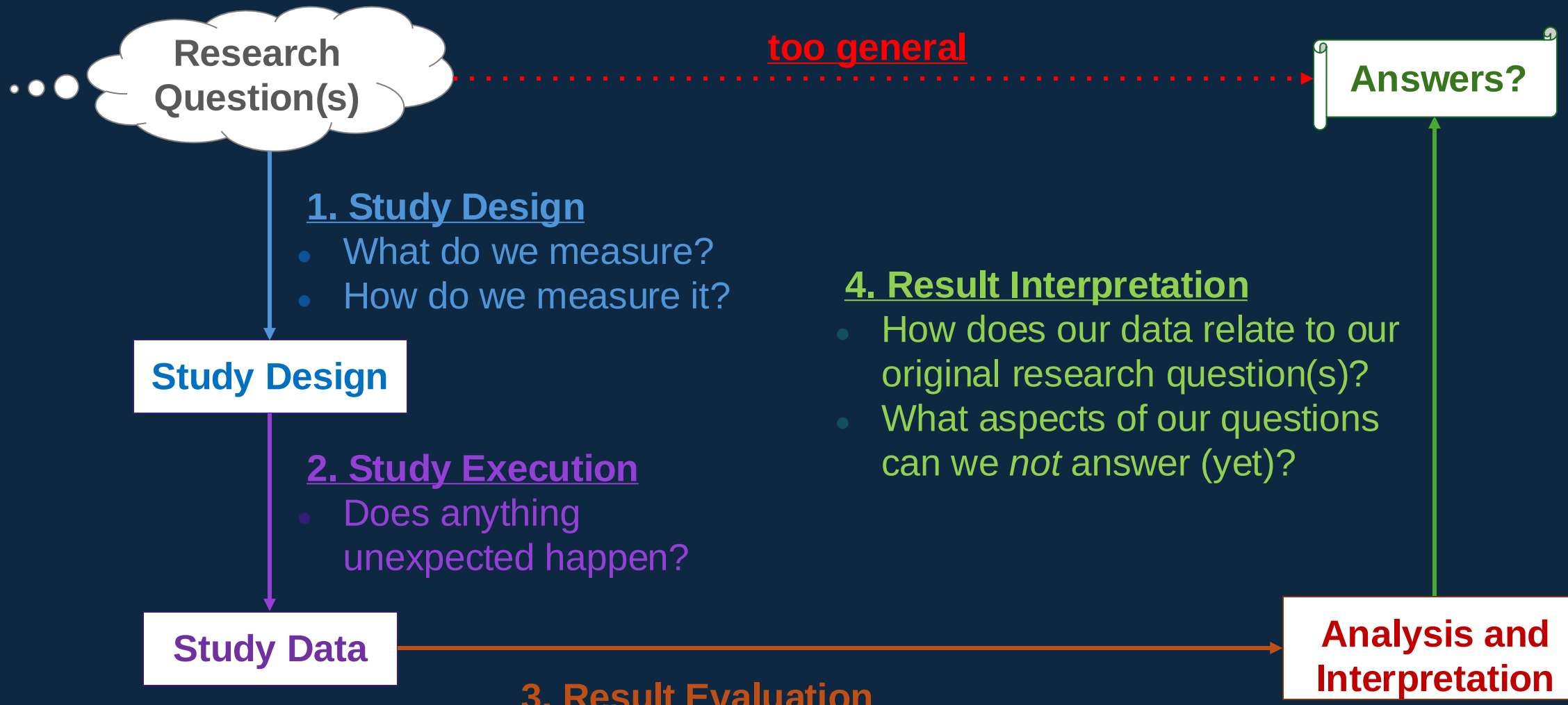
There are two general categories of research questions:

- Qualitative questions
  - What? Why? How?
- Quantitative questions
  - How much?
  - Yes/no questions can often be translated into more precise “how much” questions (e.g., “Is A better than B” → “How much better is A than B?”)

Research Questions

“Is my implementation of truth tables easy to use?”

- This question is impossible to answer definitively! Why?
  - Are we referring to instructors or students here?
  - Either way, how would we ask every potential user of their opinion?
  - We don’t even have a definition what “easy to use” means!
- We need to narrow down our question to:
  - *a specific type of user*
  - *a specific use case*
  - *a measurable definition of “easy to use”*


Research Question(s)
too general
Answers?
1. Study Design
What do we measure?
How do we measure it?
Study Design
2. Study Execution
Does anything unexpected happen?
Study Data
4. Result Interpretation
How does our data relate to our original research question(s)?
What aspects of our questions can we *not* answer (yet)?
Analysis and Interpretation
3. Result Evaluation
What data did we collect?
What causes or patterns can we infer?

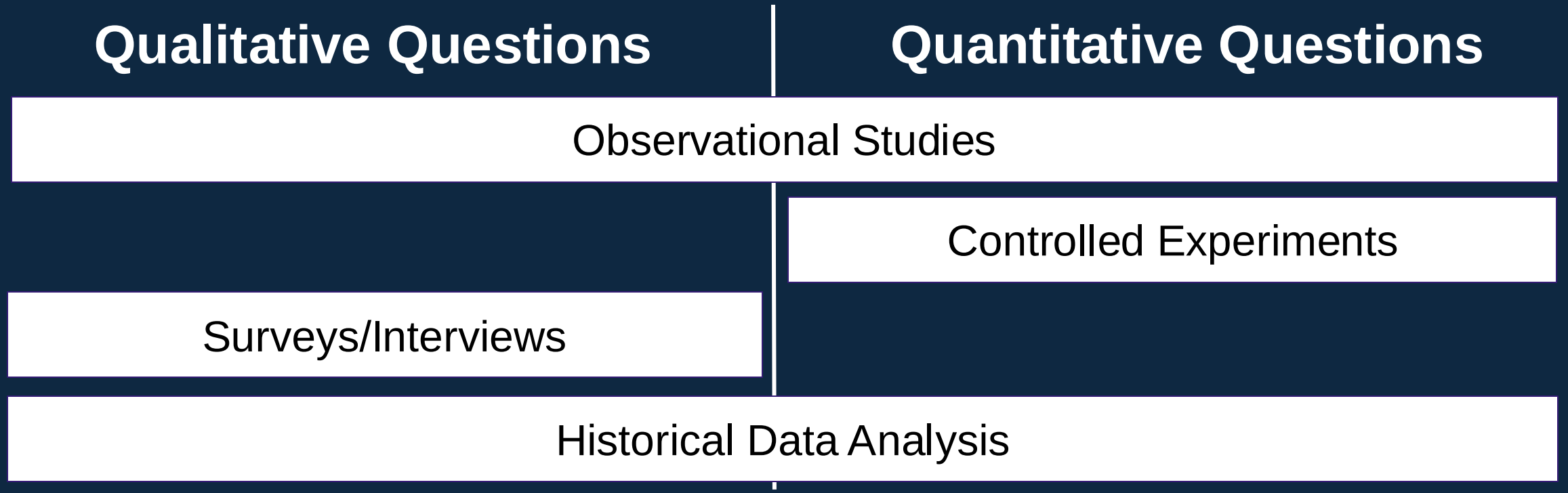
Types of Studies

Almost any study falls into one of the following categories:

1. Observational Study
   - Observe participants as they do a task
2. Controlled Experiment
   - Compare two (usually randomly assigned) groups of participants
3. Survey/Interview
   - Let participants directly answer question(s)
4. Historical Data Collection/Analysis
   - Find and analyze already existing data (e.g., from previous studies)

Types of Studies

Typical Fit for Research Questions



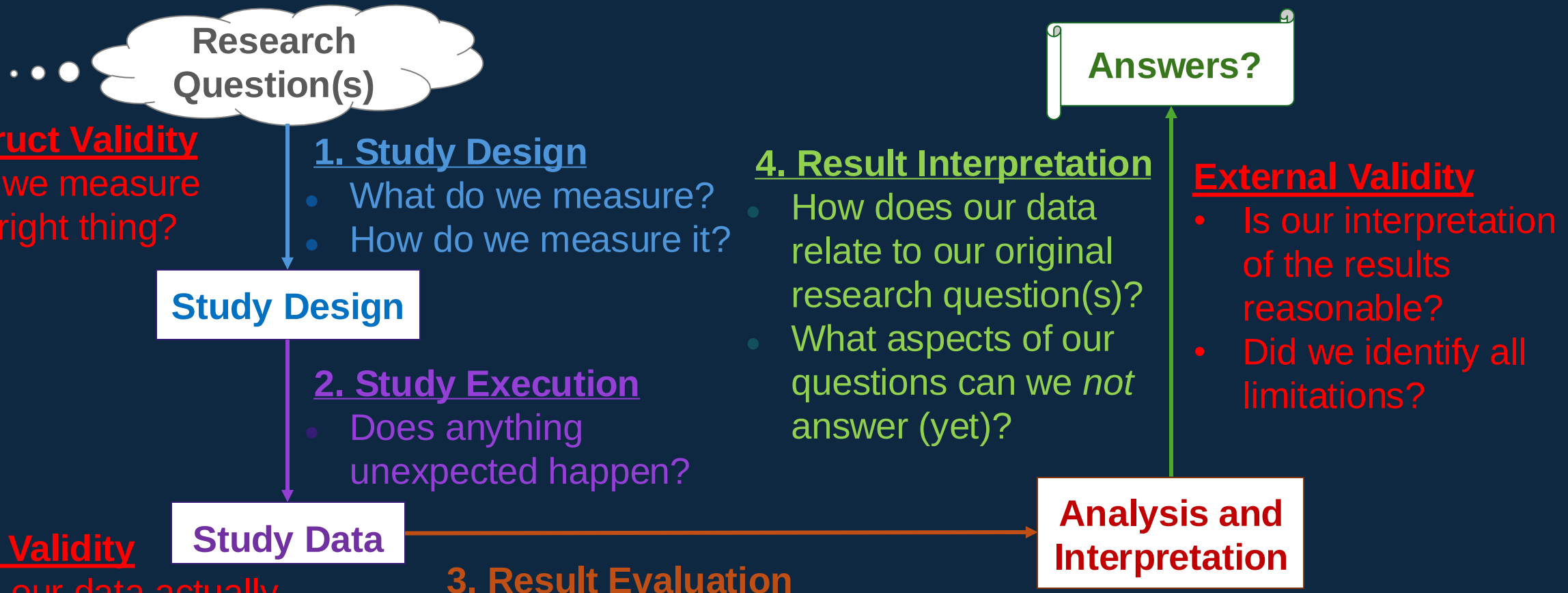
Types of Studies - Exercise

“Is my implementation of truth tables easy to use?”

How could you answer this question via a..?

1. Observational Study
   - Observe participants as they do a task
2. Controlled Experiment
   - Compare (usually randomly assigned) two groups of participants
3. Survey/Interview
   - Let participants directly answer our question(s)
4. Historical Data Collection/Analysis
   - Find and analyze already existing data (e.g., from previous studies)

Threats to Validity
(what could possibly go wrong?)



Threats to Validity - Exercise

Suppose our research question is

“How do instructors want truth tables in PrairieLearn to work?”

… and we conducted the interview with Geoffrey Herman as a study to answer this question

What threats to validity can you think of?

- Construct validity threats
- Internal validity threats
- External validity threats

Risks and Informed Consent

By now, all of you should have completed the IRB training!

Reminders for real studies:

- You always need to evaluate the potential risks for participants *before* you run any study
- Participants need to be informed of these risks and and the purpose/benefits of your study
  - There are few exceptions, but don’t assume that one applies to your study just because you think the risks are minimal!
- Most research institutions (including UIUC) have an IRB that needs to review and approve studies

Homework for Next Week

1. Read a paper about the System Usability Scale
   - Think about it especially from a perspective of "what threats to validity does it try to address?"
2. Get started on your programming project