METHODOLOGY MATTERS

Is There a Method Choice Bias in Software Engineering?

Courtney Williams, Alexey Zagalsky, Margaret-Anne Storey

NIER ICSE 2018

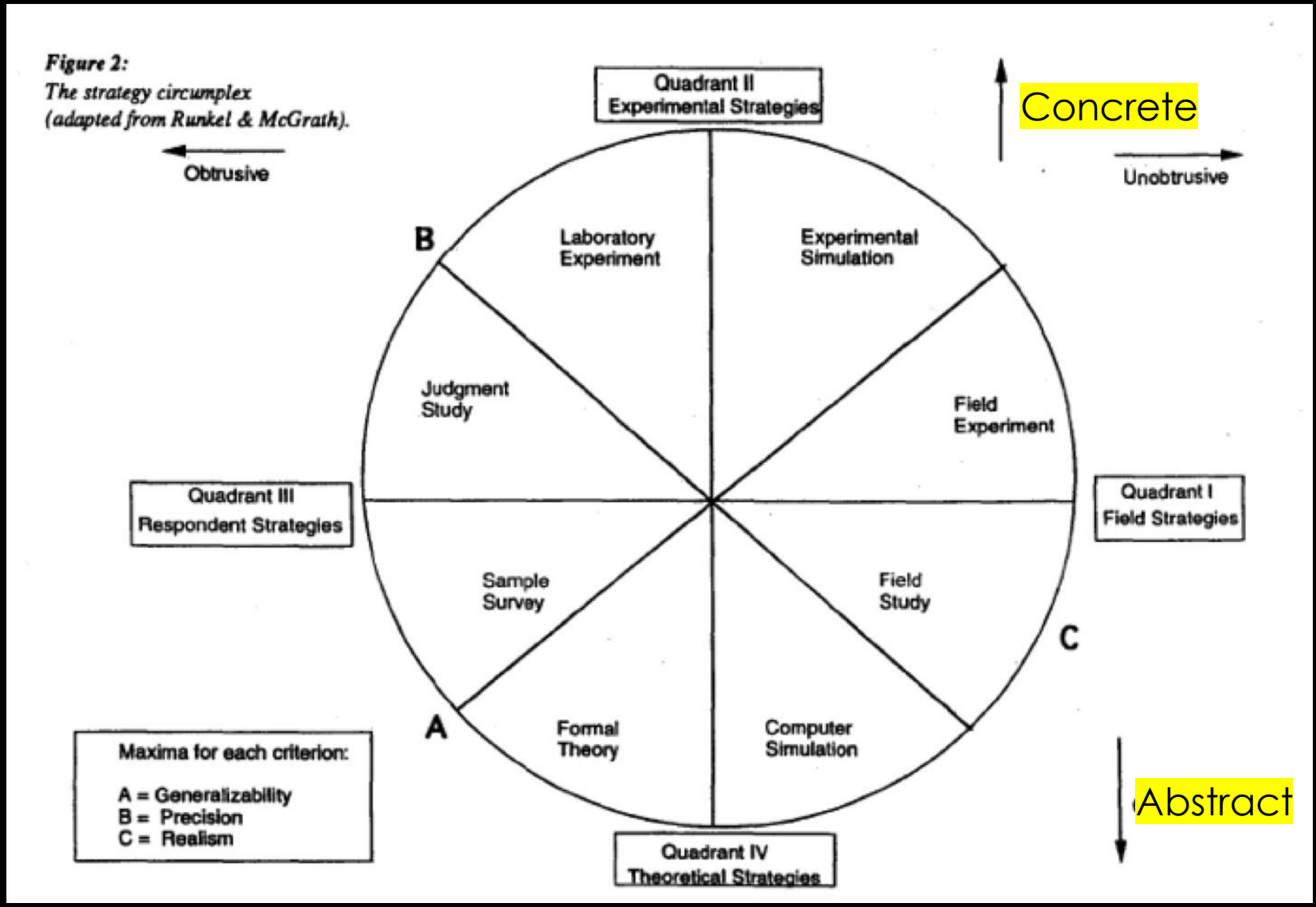
- Reflections on the past
  - Startling results that call current research directions into question;
  - Bold arguments on current research directions that may be somehow misguided;
  - Results that disregard established results or believe of evidence that call for fundamentally new directions.
- Visions of the future
  - Bold visions of new directions that may not yet be supported by solid results, but rather by a strong and well-motivated scientific intuition.

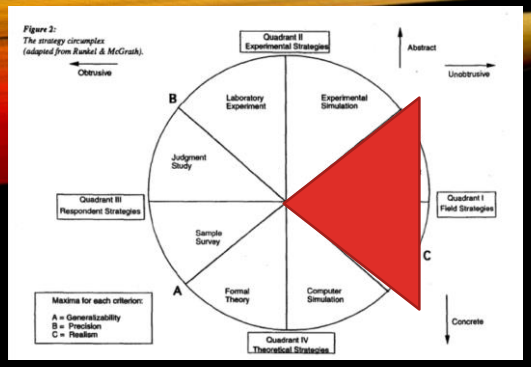
RESEARCH QUALITY

- What makes good research in software engineering?
- Mary Shaw
- Focuses on research questions, methods, and evaluation criteria

- Grounded theory in software engineering research: a critical review and guidelines
- Stol, Ralph, and Fitzgerald
- Focuses in on Grounded Theory studies and aspects of quality in GT work

METHOD CHOICE

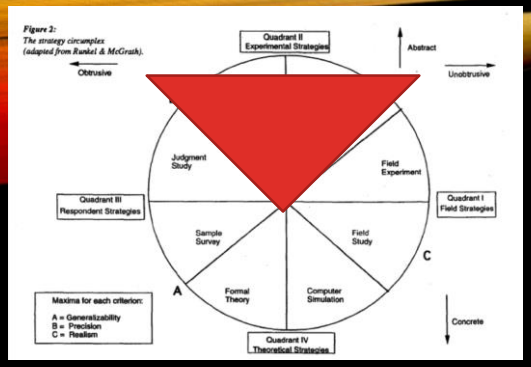




FIELD STRATEGIES

- FIELD STUDIES
- No manipulation
- Observing participants in their "natural environment"

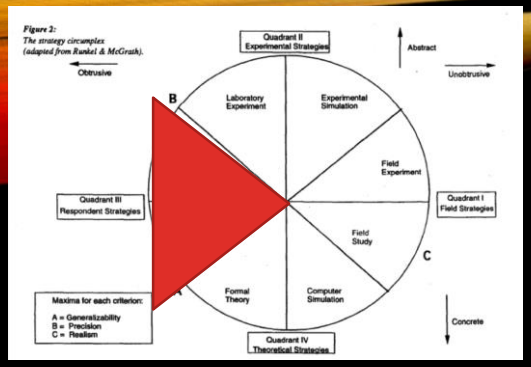
- FIELD EXPERIMENTS
- Introduce a controlled variable to the natural environment



EXPERIMENTAL STRATEGIES

- LABORATORY EXPERIMENTS
- Controlled situations
- Outside of the participant's natural environment

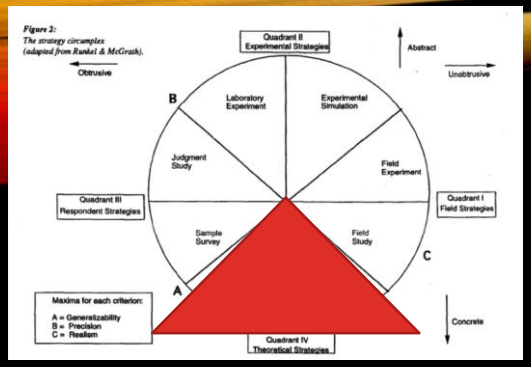
- EXPERIMENTAL SIMULATIONS
- Controlled situations
- Simulating the participant's natural environment in the lab setting



RESPONDENT STRATEGIES

- SAMPLE SURVEYS
- Investigate the effects of a phenomenon on a population
- Relies on self-reports of participants
- Questionnaires, surveys, interviews

- JUDGMENT STUDIES
- Investigate aspects of a phenomenon using a population
- Relies on self-reports
- Typically used to evaluate a tool or technique's efficacy



FORMAL METHODS

No active human participation

- COMPUTER SIMULATIONS
- Complete and closed system
- Data mining studies
- Computerized analysis of software
- Automatic tool evaluations using repository data
- Prediction and classification models

- FORMAL THEORY
- No gathering of new empirical evidence
- The creation of models and theories
- Systematic literature reviews, meta-analysis, etc.

HUMAN PARTICIPANTS

- ACTIVE
- Self-reports
- Visible observer
- Hidden observer

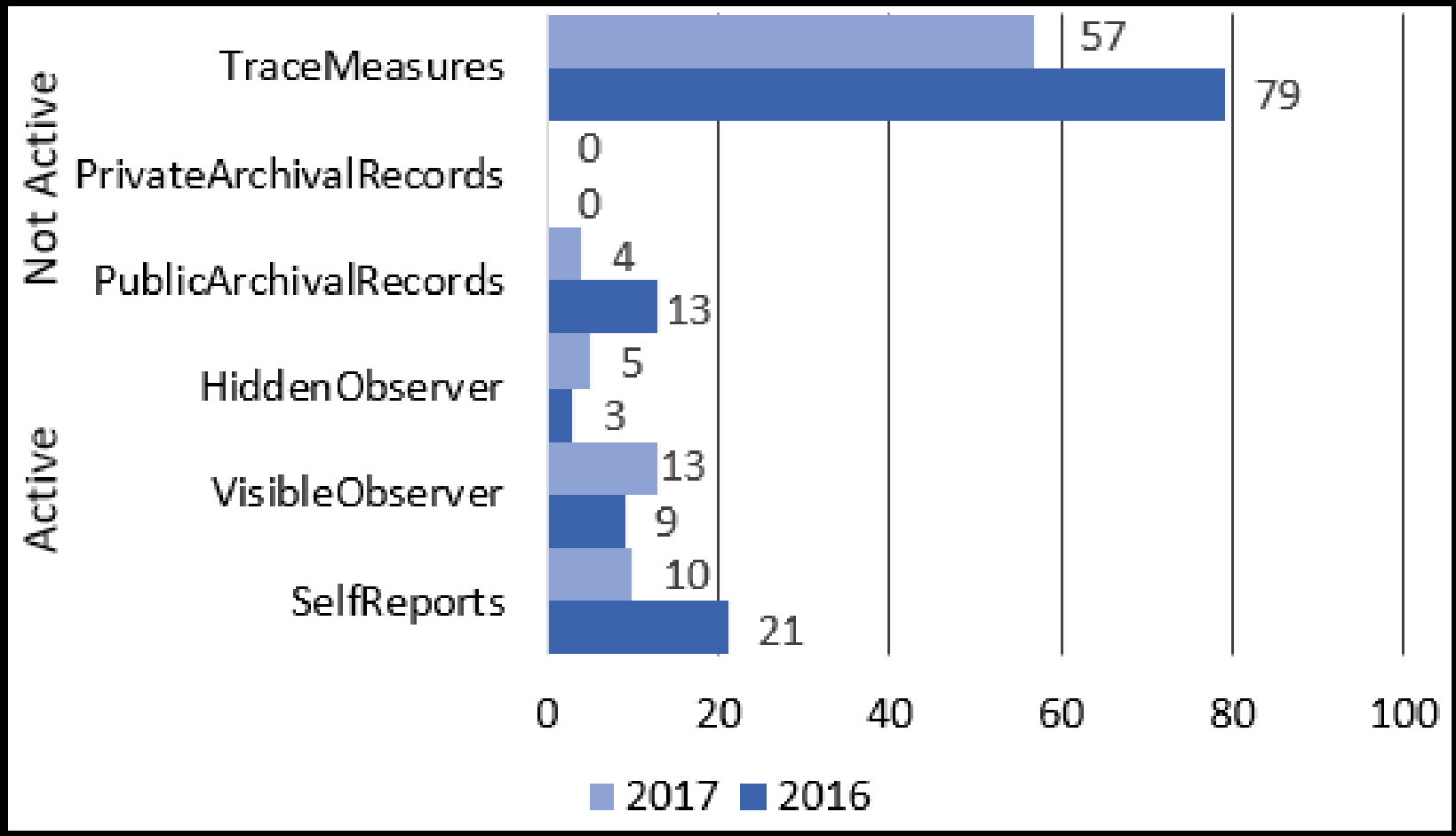
- INACTIVE
- Public archival records
- Private archival records
- Trace measures

THE STUDY

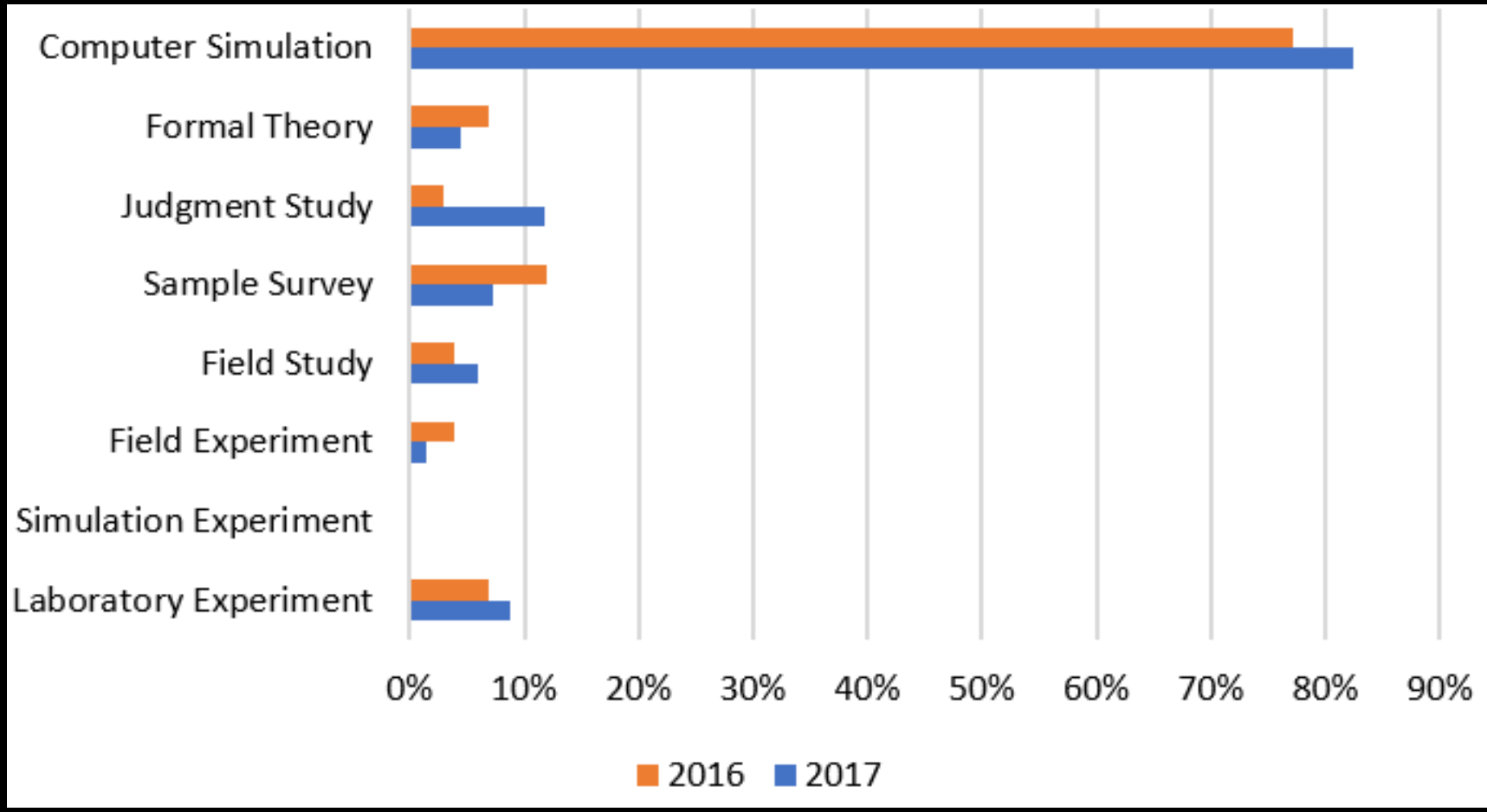
- Applied McGrath's models to SE
- Descriptions of these methods in the SE domain

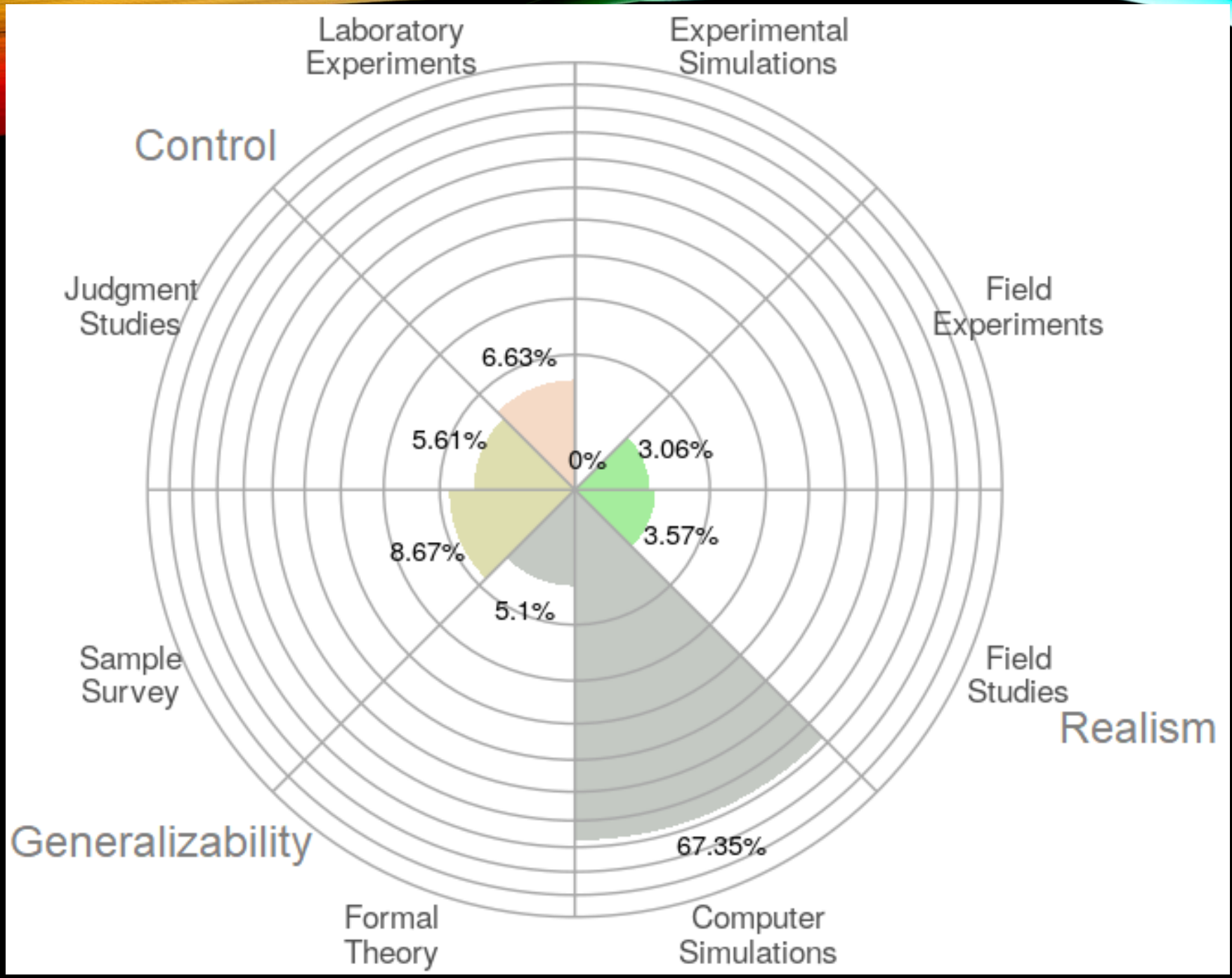
- ICSE 2017 and 2016
- Technical track papers
- 68 (2017) + 101 (2016) = 169 papers
- Classified in excel spreadsheet
- Research method, human involvement
- Inter-rater reliability: 72%

FINDINGS

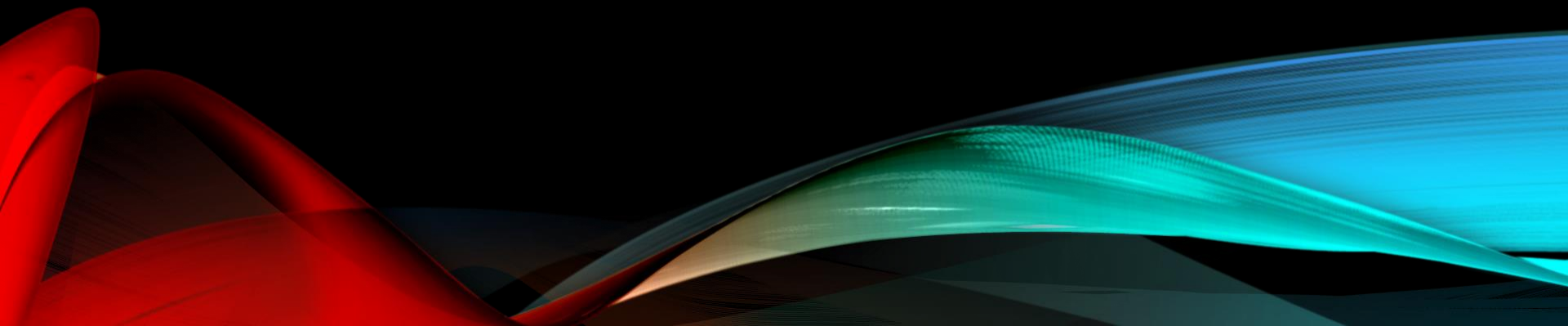

Not Active
Active
TraceMeasures
PrivateArchivalRecords
PublicArchivalRecords
HiddenObserver
VisibleObserver
SelfReports
57
79
0
0
4
13
5
3
13
9
10
21
0
20
40
60
80
100
2017
2016

FINDINGS


Computer Simulation
Formal Theory
Judgment Study
Sample Survey
Field Study
Field Experiment
Simulation Experiment
Laboratory Experiment
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
2016
2017


Laboratory Experiments
Experimental Simulations
Control
Judgment Studies
Field Experiments
6.63%
5.61%
0%
3.06%
3.57%
8.67%
5.1%
Sample Survey
Field Studies
Realism
Generalizability
67.35%
Formal Theory
Computer Simulations

DISCUSSION



BIG DATA

When does it become inappropriate to conduct software engineering research using only big data resources and repositories?

NEW TECHNOLOGIES

In what circumstances is it inappropriate to conduct human research remotely?

NEW TECHNOLOGIES

Will future technologies make remote research as rigorous as in-person interaction?

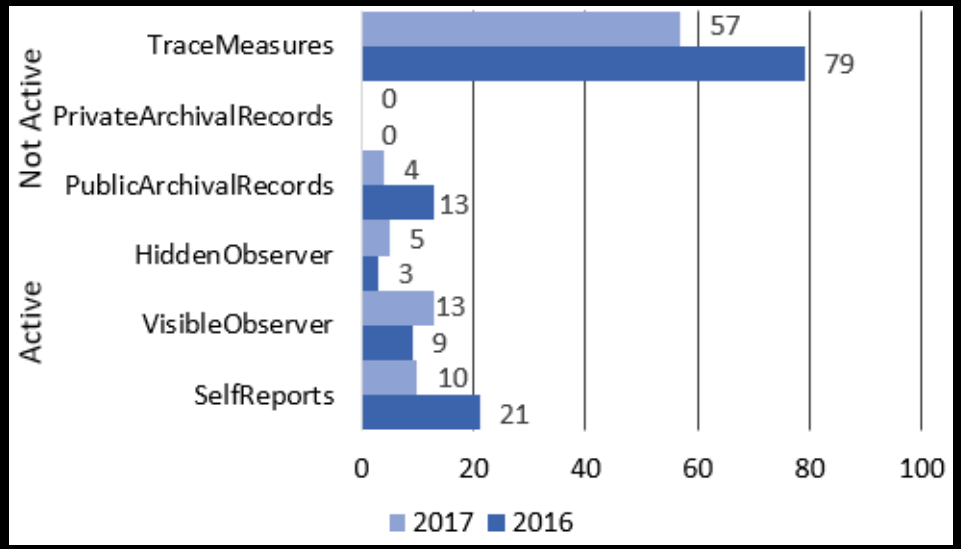
NEW TECHNOLOGIES

How should we approach the study of virtual development environments?

HUMAN INVOLVEMENT

What are the implications of using inactive forms of human participation in the majority of our research?

Is this how we want to move forward as a community?



METHOD BALANCE

What are the implications of this method "imbalance"?

Is a balance even desired?