2IMP40 EMPIRICAL METHODS IN SOFTWARE ENGINEERING

WHERE DO WE START?

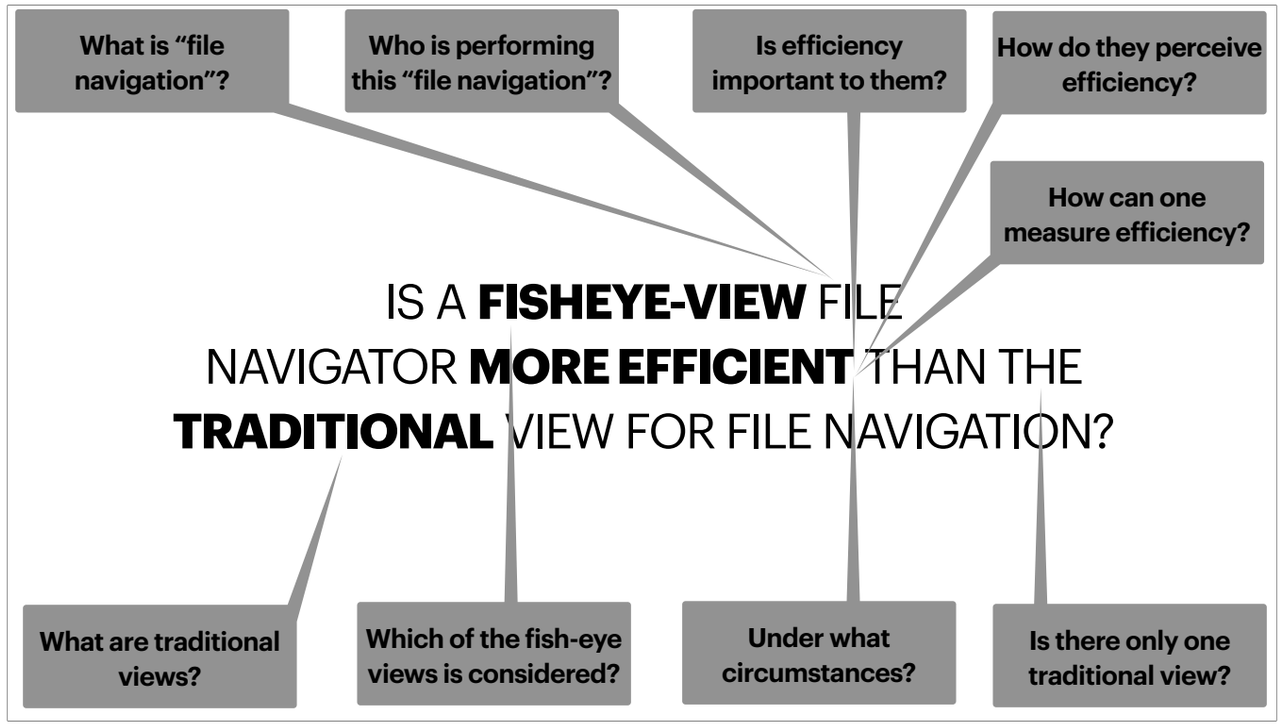
ALEXANDER SEREBRENIK

BOTH images are fisheye visualisations of the source code the second one hides “irrelevant” lines, see the triangles

This is Jane. Jane is a new PhD student interested in the effectiveness of a novel fisheye-view file navigator. Her research is motivated by the fact that navigation is a primary activity of software developers requiring a lot of scrolling and many clicks to find files. ‘Fisheyeviews’ use a distortion technique that, if applied correctly, display information in a compact format that could potentially reduce the amount of scrolling required. Jane’s intuition is that the fisheye-view file navigator is more efficient for file navigation, but critics argue that the more compact information is difficult to read and that developers will not adopt it over the traditional file navigator. Her research goal, therefore, is to find evidence that supports or refutes her intuition that fisheye-view file navigators are more efficient than traditional file navigators for navigation.

IS A **FISHEYE-VIEW** FILE
NAVIGATOR **MORE EFFICIENT** THAN THE
TRADITIONAL VIEW FOR FILE NAVIGATION?

Jane wants to know…

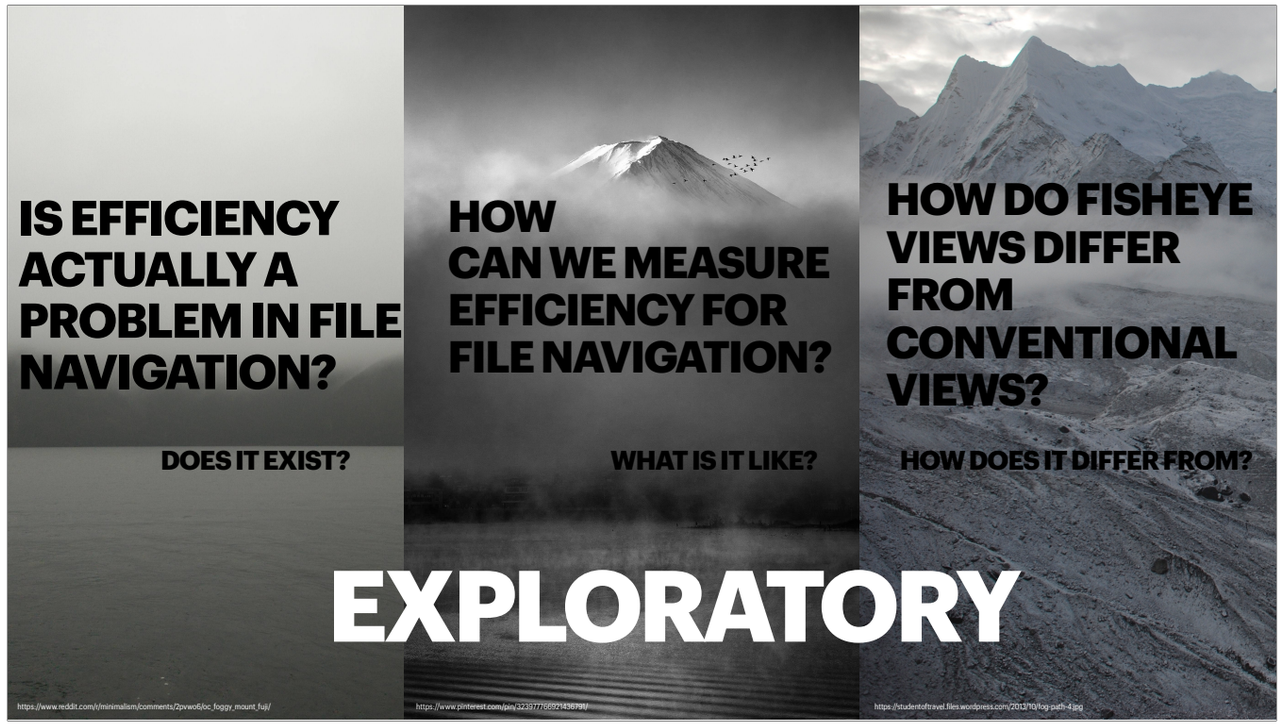


However, this is not a good research question. It is vague and it is based on hidden assumptions

HOW TO ASK A RESEARCH QUESTION?

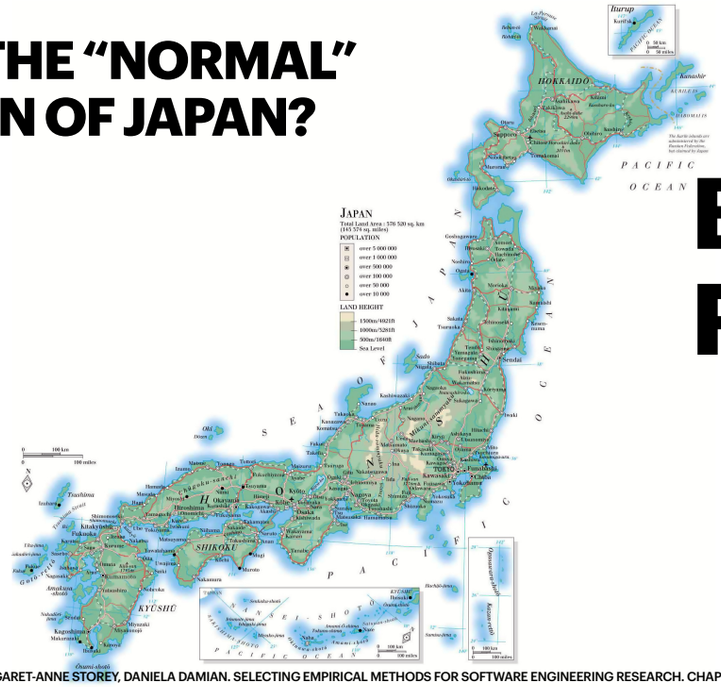
Let us assume that we have observed the view on the right. We are not even sure what we see, we suspect it to be a mountain but we are not sure whether it exists. If we observe the view in the middle, we know that the mountain exists but we do not know much about it. We want to describe it, understand how big it is, compare with another mountain we have seen before.

The first and the second are Mount Fuji in Japan. The third one is Mount Everest.



The first and the second are Mount Fuji in Japan. The third one is Mount Everest.

WHAT IS THE “NORMAL” ELEVATION OF JAPAN?



BASE RATE

STEVE EASTERBROOK, JANICE SINGER, MARGARET-ANNE STOREY, DANIELA DAMIAN. SELECTING EMPIRICAL METHODS FOR SOFTWARE ENGINEERING RESEARCH. CHAPTER 11 IN GUIDE TO ADVANCED SOFTWARE ENGINEERING.

But then we want to understand how special this mountain is for what is “normal” for this region. Let us assume that we know that our observation was made in Japan. We see that most of Japan is green, meaning it is at the sea level or lower than 1000 meters.

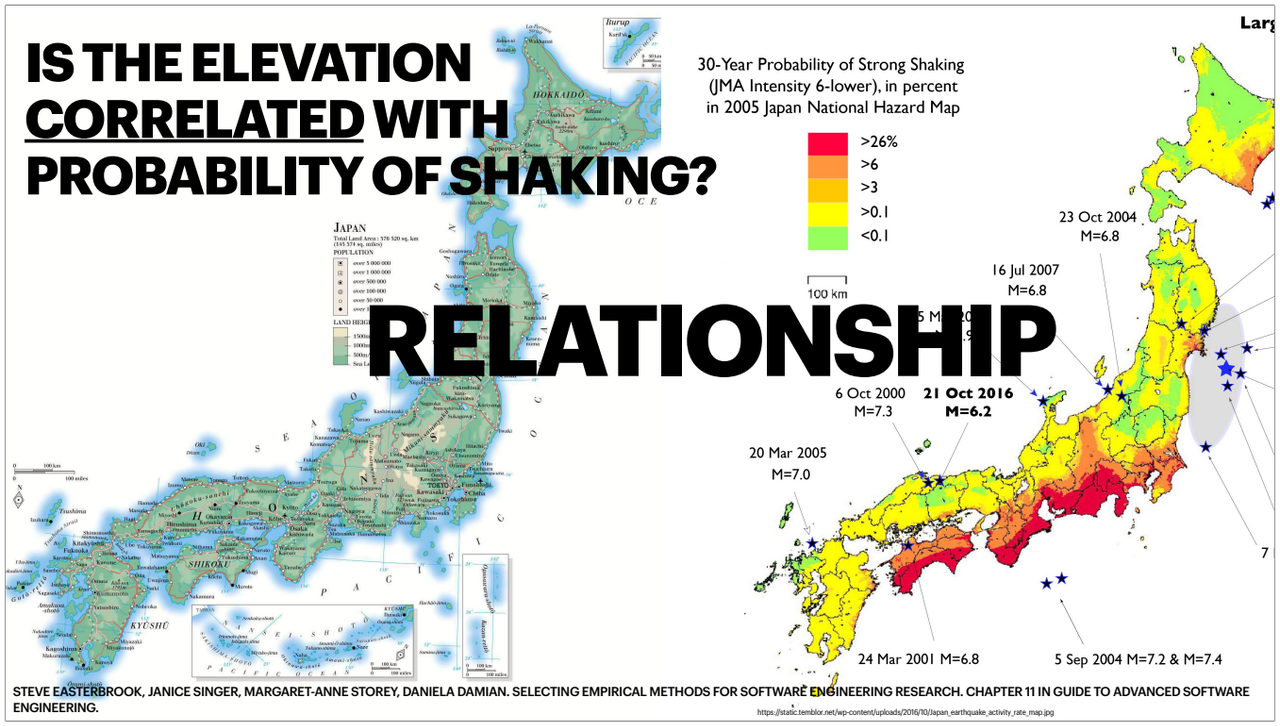
This means that we need to understand the base rate. This is the next group of research questions.

HOW DO PROGRAMMERS NAVIGATE FILES USING EXISTING TOOLS?

STEVE EASTERBROOK, JANICE SINGER, MARGARET-ANNE STOREY, DANIELA DAMIAN. SELECTING EMPIRICAL METHODS FOR SOFTWARE ENGINEERING RESEARCH. CHAPTER 11 IN GUIDE TO ADVANCED SOFTWARE ENGINEERING.

BASE RATE

Back to Software Engineering an example of a base rate question is…

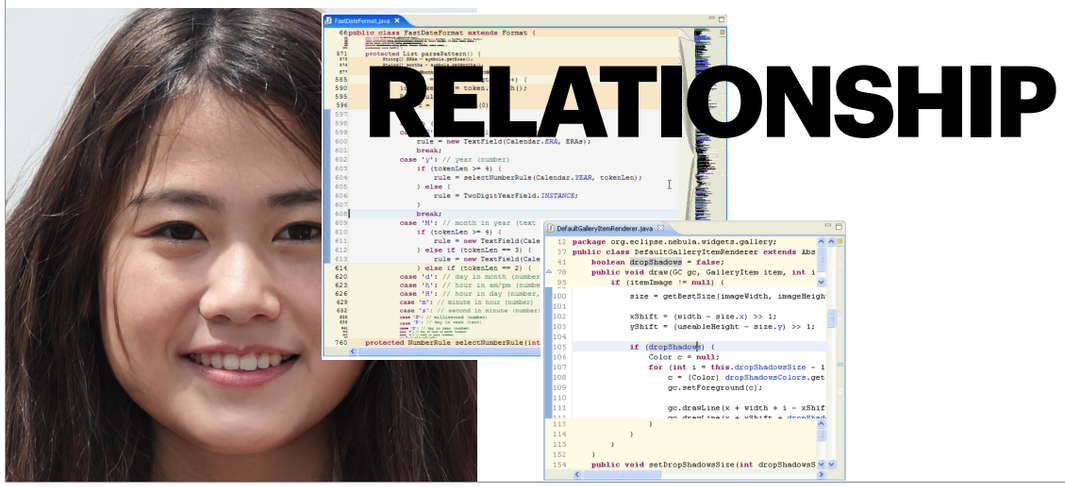


STEVE EASTERBROOK, JANICE SINGER, MARGARET-ANNE STOREY, DANIELA DAMIAN. SELECTING EMPIRICAL METHODS FOR SOFTWARE ENGINEERING RESEARCH. CHAPTER 11 IN GUIDE TO ADVANCED SOFTWARE ENGINEERING.

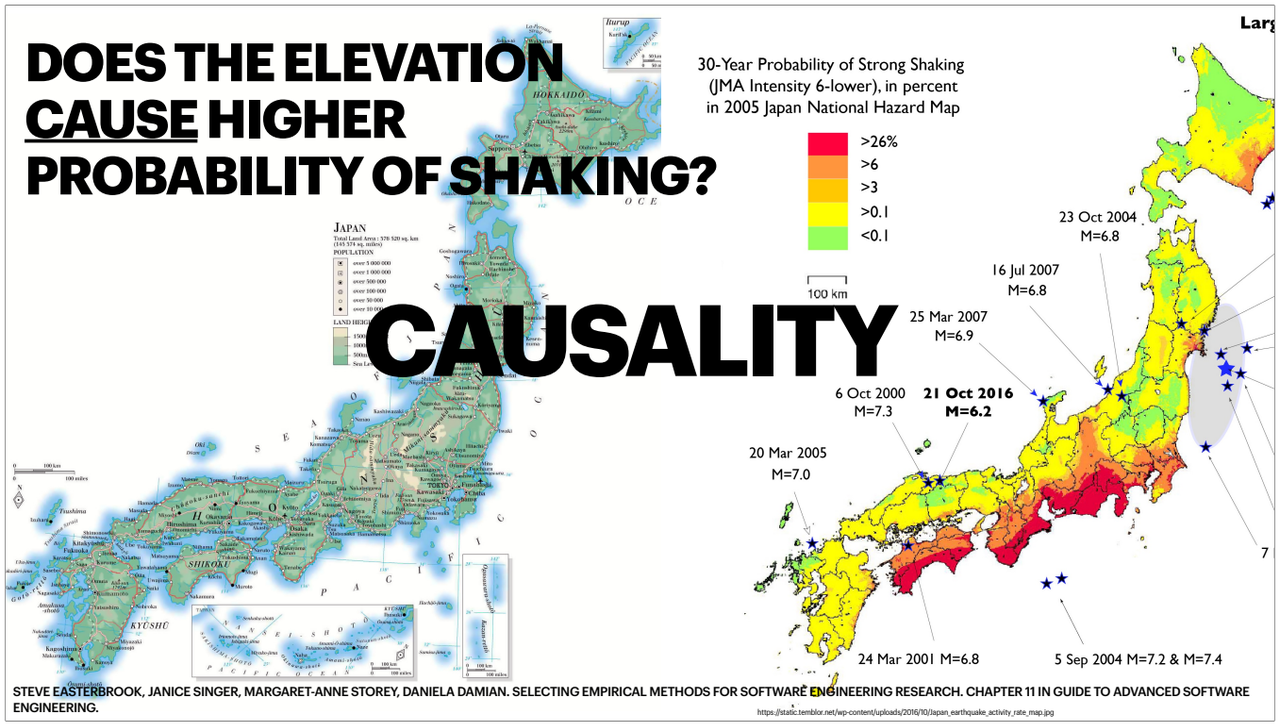
https://static.temblor.net/wp-content/uploads/2016/10/Japan_earthquake_activity_rate_map.jpg

another group of questions is related to relationship between variables
On the left is the map of elevation of Japan, on the right the map of the seismic hazard. We want to understand whether these two variables are related.

DOES EFFICIENCY IN FILE NAVIGATION CORRELATE WITH THE PROGRAMMER'S FAMILIARITY WITH THE PROGRAMMING ENVIRONMENT?



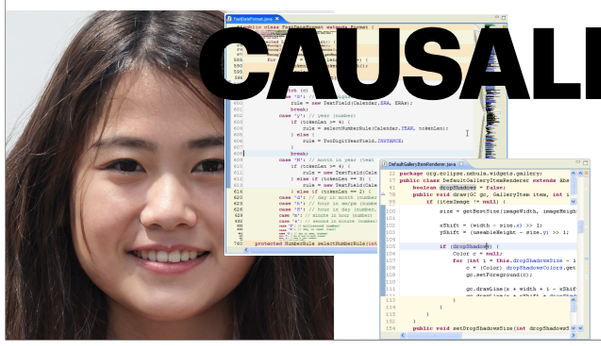
Back to Software Engineering an example of a question about a relationship is…



Of course, in this case the answer is “no”. Correlation is not the same as causation. In this case, the collision of tectonic plates that forms the tallest and steepest mountains on Earth produces large and destructive earthquakes.

DO FISHEYE-VIEWS CAUSE AN IMPROVEMENT IN EFFICIENCY FOR FILE NAVIGATION?

DO FISHEYE-VIEWS CAUSE PROGRAMMERS TO BE MORE EFFICIENT AT FILE NAVIGATION THAN CONVENTIONAL VIEWS?



CAUSALITY

DO FISHEYE-VIEWS CAUSE PROGRAMMERS TO BE MORE EFFICIENT AT FILE NAVIGATION THAN CONVENTIONAL VIEWS WHEN PROGRAMMERS ARE DISTRACTED, BUT NOT OTHERWISE?

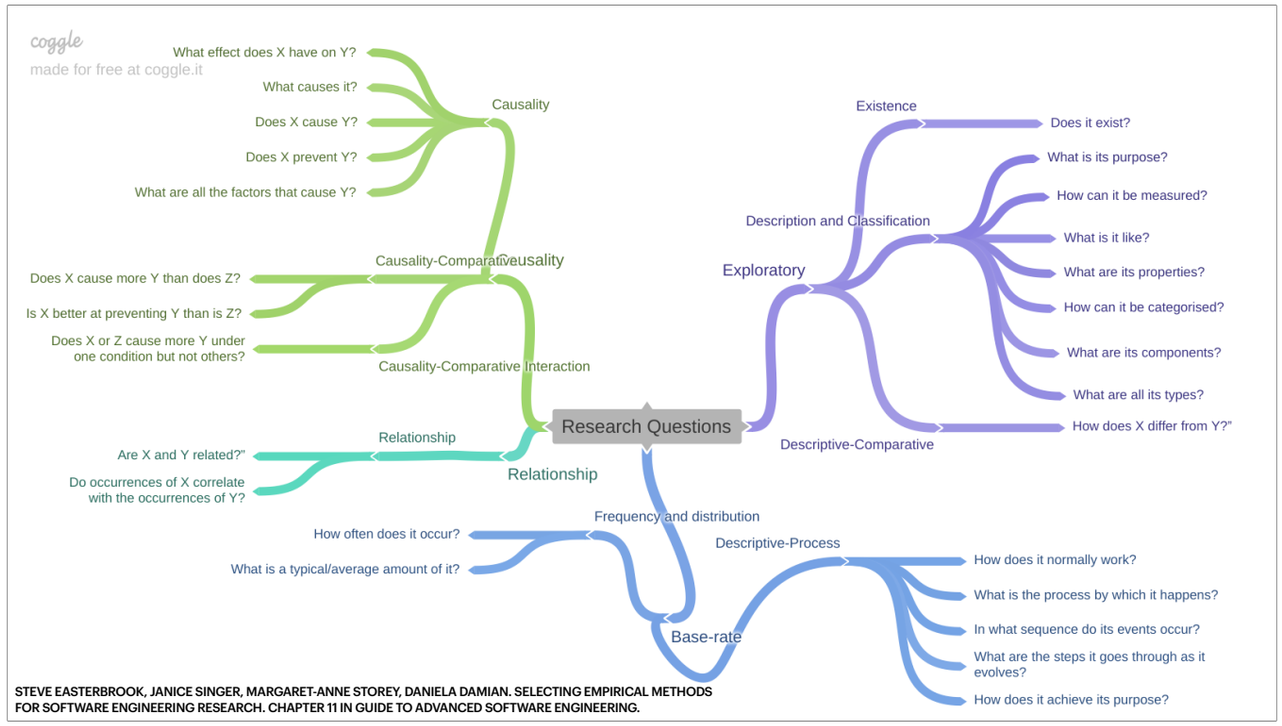
DO FISHEYE-VIEWS CAUSE AN IMPROVEMENT IN EFFICIENCY FOR FILE NAVIGATION?

Requires base-rate to evaluate improvement

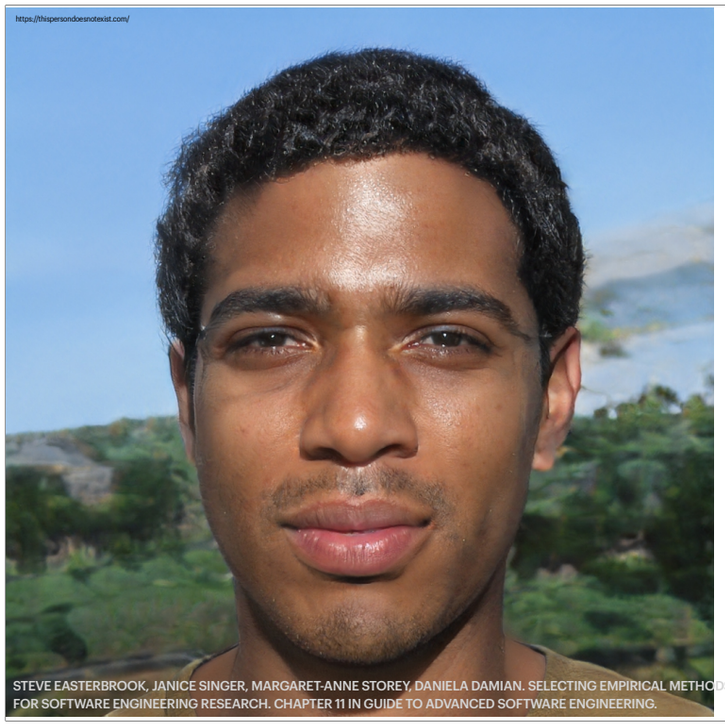
DO FISHEYE-VIEWS CAUSE PROGRAMMERS TO BE MORE EFFICIENT AT FILE NAVIGATION THAN CONVENTIONAL VIEWS?

Comparison

DO FISHEYE-VIEWS CAUSE PROGRAMMERS TO BE MORE EFFICIENT AT FILE NAVIGATION THAN CONVENTIONAL VIEWS WHEN PROGRAMMERS ARE DISTRACTED, BUT NOT OTHERWISE?



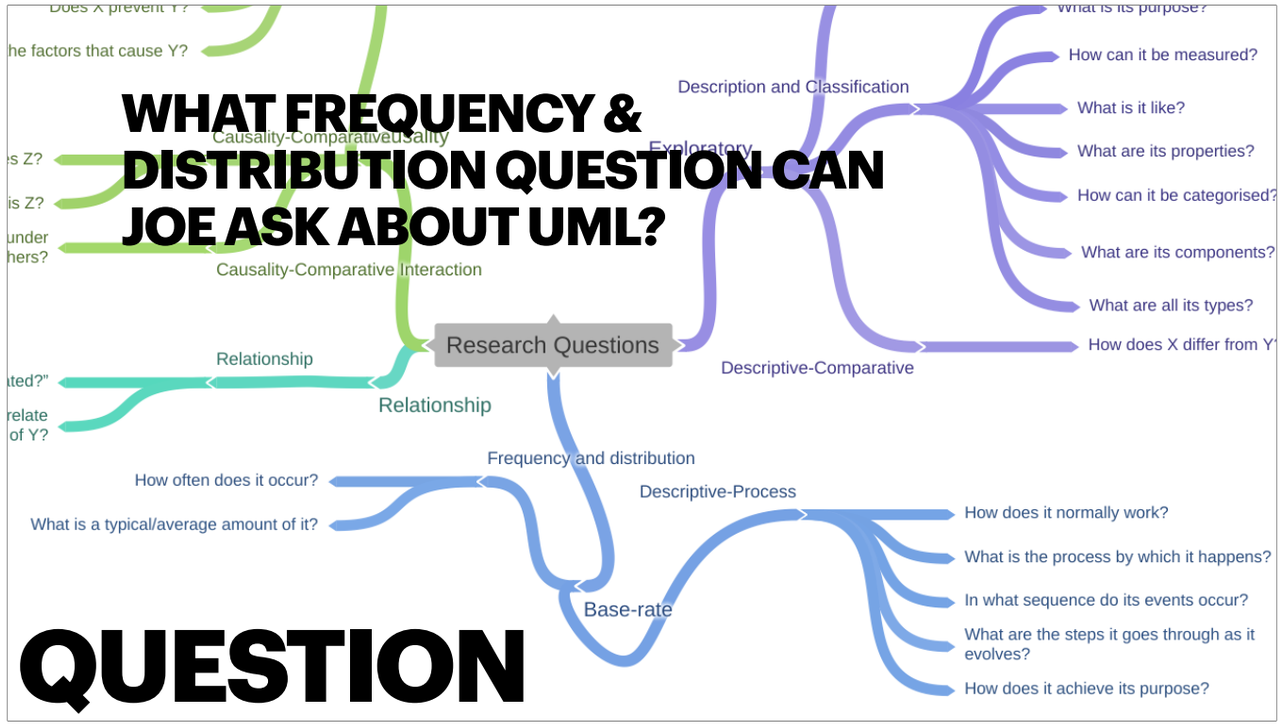
Exploratory, Base-rate, Relationship and Causality



STEVE EASTERBROOK, JANICE SINGER, MARGARET-ANNE STOREY, DANIELA DAMIAN. SELECTING EMPIRICAL METHODS FOR SOFTWARE ENGINEERING RESEARCH. CHAPTER 11 IN GUIDE TO ADVANCED SOFTWARE ENGINEERING.

UML IS TAUGHT AT THE UNI BUT IS IT USED IN THE INDUSTRY?

This is Joe



“How many distinct UML diagrams are created in software development projects in large software companies?”

WE KNOW WHAT WE WANT TO KNOW.
BUT HOW DO WE KNOW IT?

DIFFERENT QUESTIONS REQUIRE DIFFERENT RESEARCH STRATEGIES

QUESTION

(DOES IT EXIST, WHAT EFFECT DOES X HAVE ON Y, ...)

STRATEGY

(WE WILL DISCUSS THEM NEXT)

METHODS

(INTERVIEW, SURVEY, REPO MINING,...)

KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

4 WHEELS
2 WING MIRRORS

SPEED?
COMFORT?

KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

This is a generic car, it is not a recognisable brand. Observations that we can derive from such a generic model can be expected to hold very **broadly**; but they are also **shallow, not** necessarily **useful** and might **miss corner cases**.

The car on the right is a prototype vehicle designed in China. It only has two wheels! We can answer more specific questions and derive more specific conclusions but they might only hold for this particular car!

KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

https://www.ecagroup.com/media-picture/5573-3162-1653-eca-group-driving-simulation-ef-car-14.jpg

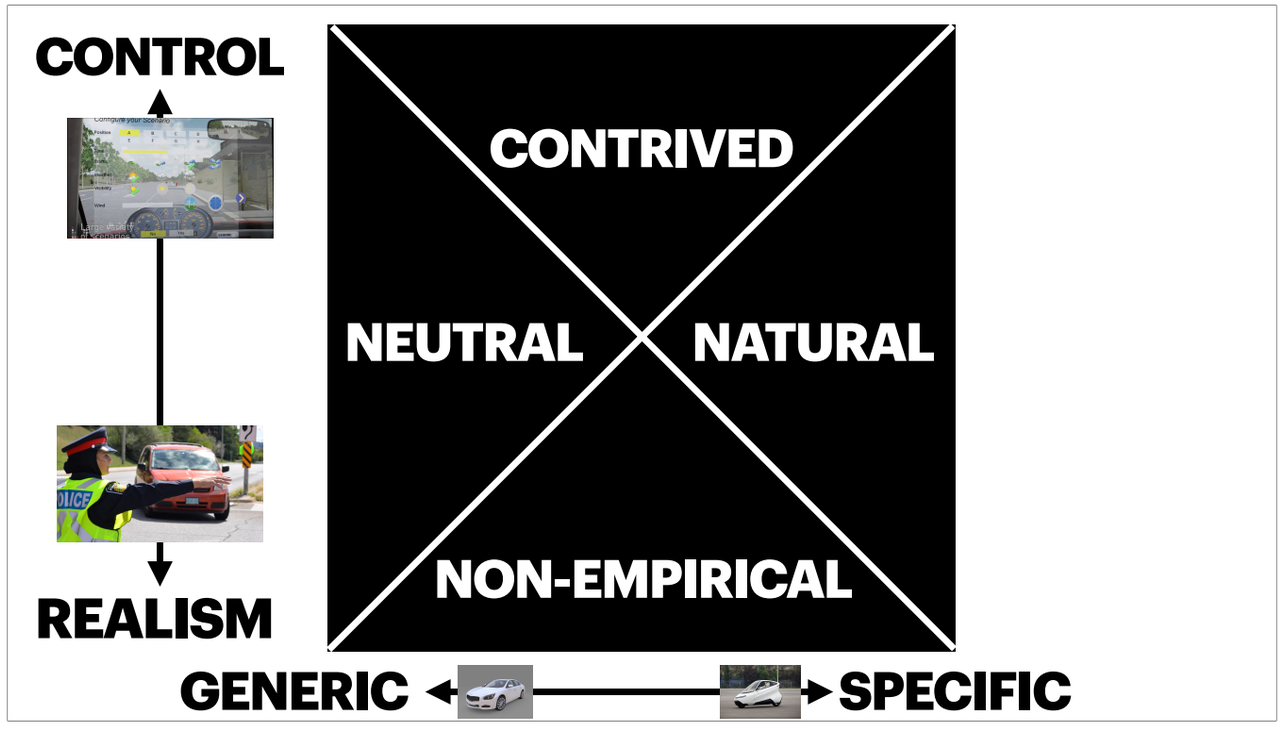
This is not even a car but a car simulator.

CONTROLLABLE BUT REALISTIC?

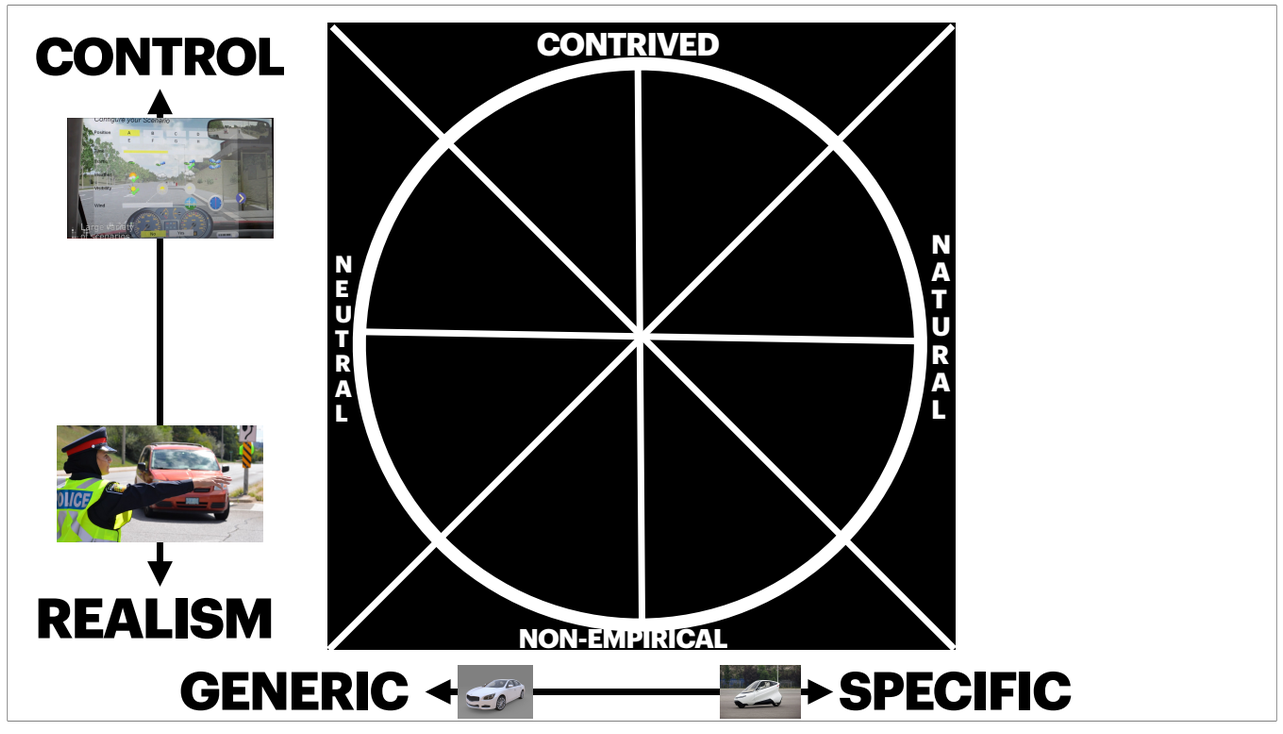
https://www.ecagroup.com/media-picture/5573-3162-1653-eca-group-driving-simulation-ef-car-14.jpg

An important advantage of the simulator is that we can control parameters that we would never be able to control in reality such as traffic, weather, visibility or wind, and understand how these parameters affect drivers' driving style. However, this is an unrealistic situation since drivers do not sit in a chair, or reality might be complex than a simulation.

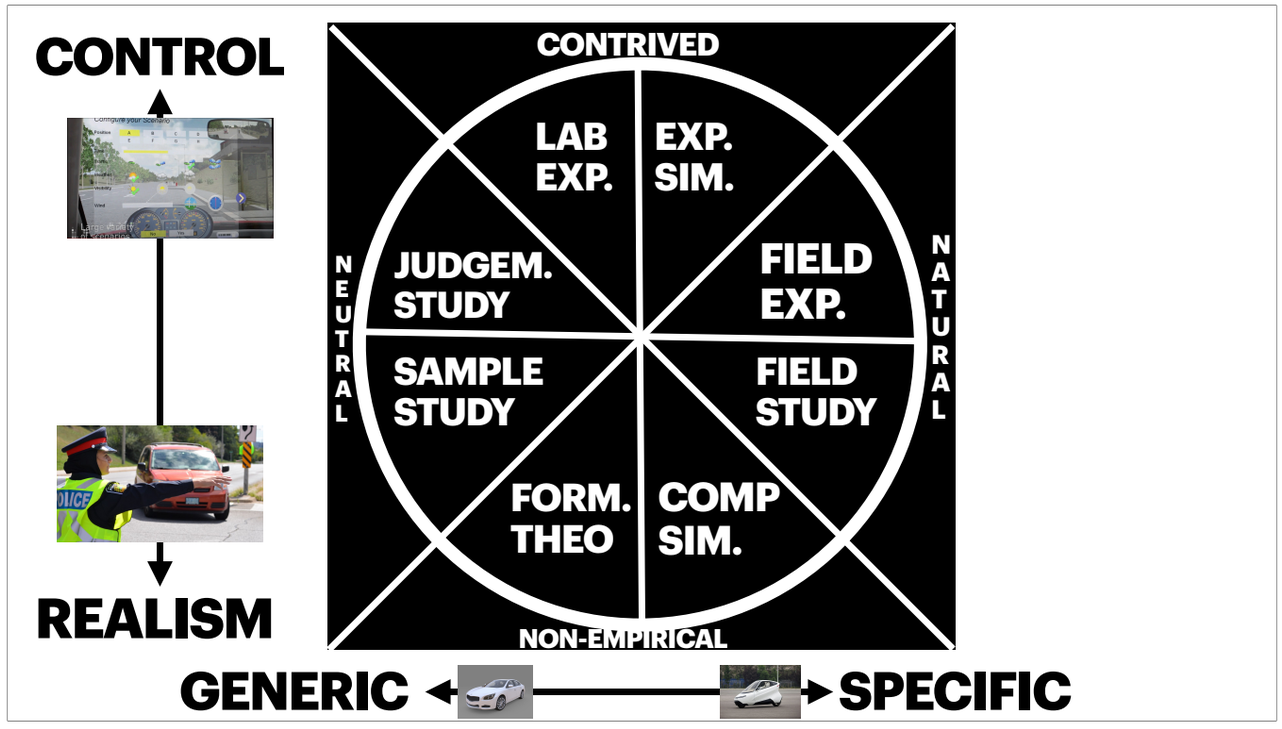
A very different scenario is on this photo. Here a police officer checks the behaviour the car driver. This is a real driver in the real context but one cannot control the factors that can affect the driver’s behaviour such as weather, wind, visibility or traffic…



Here we have four quadrants of research settings: natural, contrived, neutral and non-empirical.
Natural: rich in context and no manipulations
Contrived: all under control but unrealistic
Neutral: everything is neutralised, individual differences are blurred
Non-empirical: not immediately based on observations



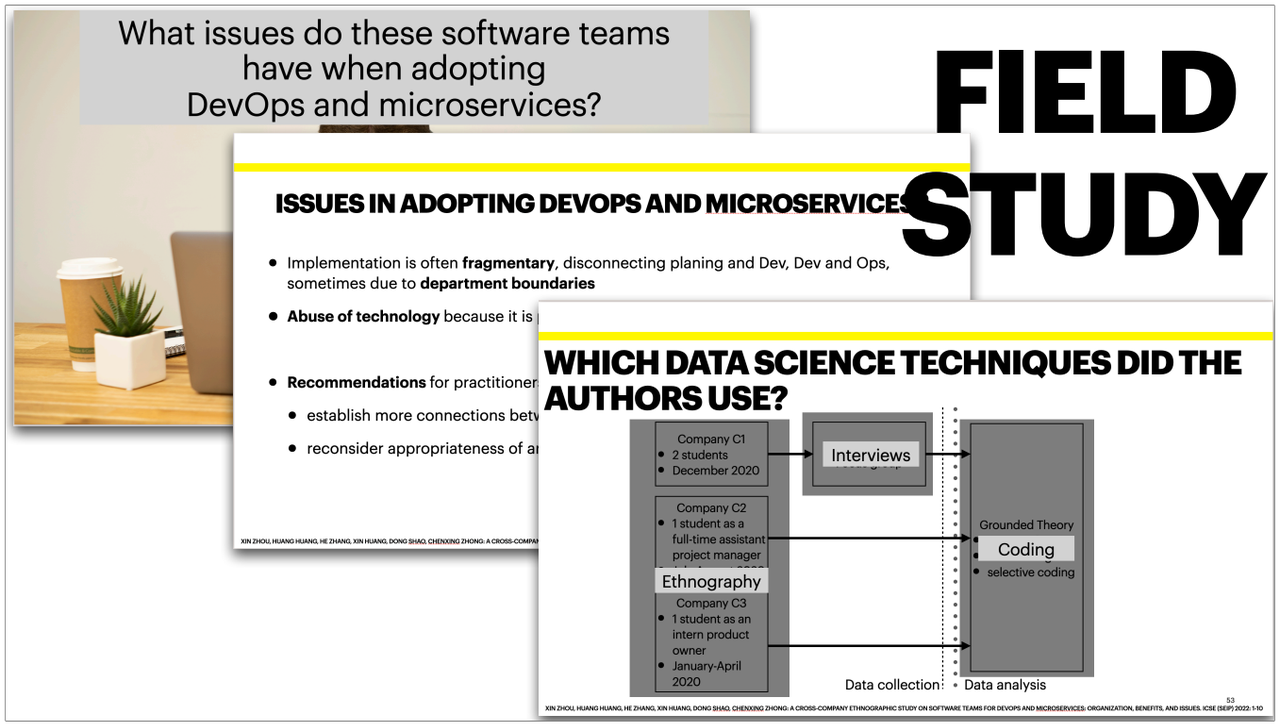
For each quadrant we have two research strategies



For each quadrant we have two research strategies: laboratory experiment, experimental simulation, field experiments, field studies, computer simulation, formal theory, sample studies, judgement studies. We are going to discuss them one by one.

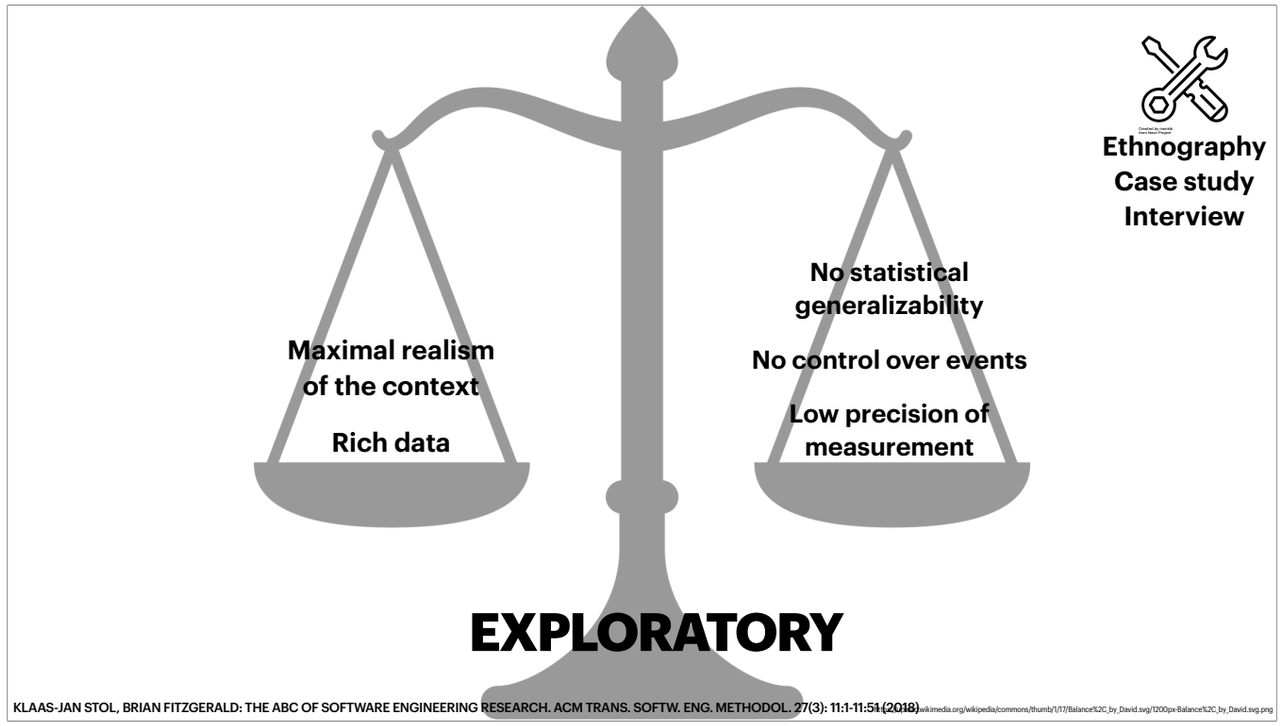
Field study is a **jungle**. Natural setting that exists before the researcher enters it. Minimal intrusion of the setting so as not to disturb realism, only to facilitate data collection. Facilitates study of phenomena and actors and their behaviour in natural contexts. Exploratory, to understand what's going on, how things work, or to generate hypotheses.
Case study, ethnography, observational study; qualitative data incl. interviews, field notes, archival documents, may include quantitative data.

- No statistical generalizability
- No control over events
- Low precision of measurement



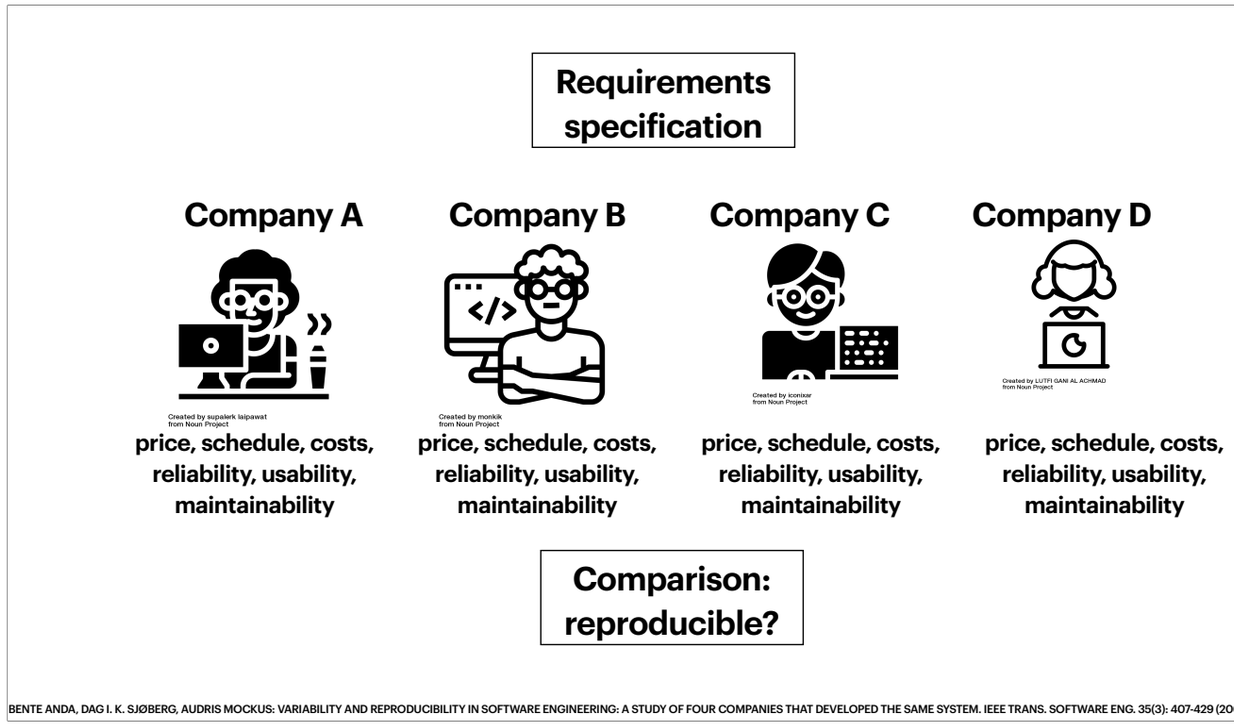
FIELD STUDY

This is an example of a field study: recall, “Minimal intrusion of the setting so as not to disturb realism, only to facilitate data collection.” In their study the research method used is ethnography. Other examples can be a case study, or an observational study; **qualitative** data incl. interviews, field notes, archival documents, may include quantitative data.” Similarly, in the master thesis of Samar Jameel she has done her very best as to not disturb the developers.



Field study is usually an exploratory strategy, to understand what’s going on, how things work, or to generate hypotheses. These hypotheses should be then verified in a separate study

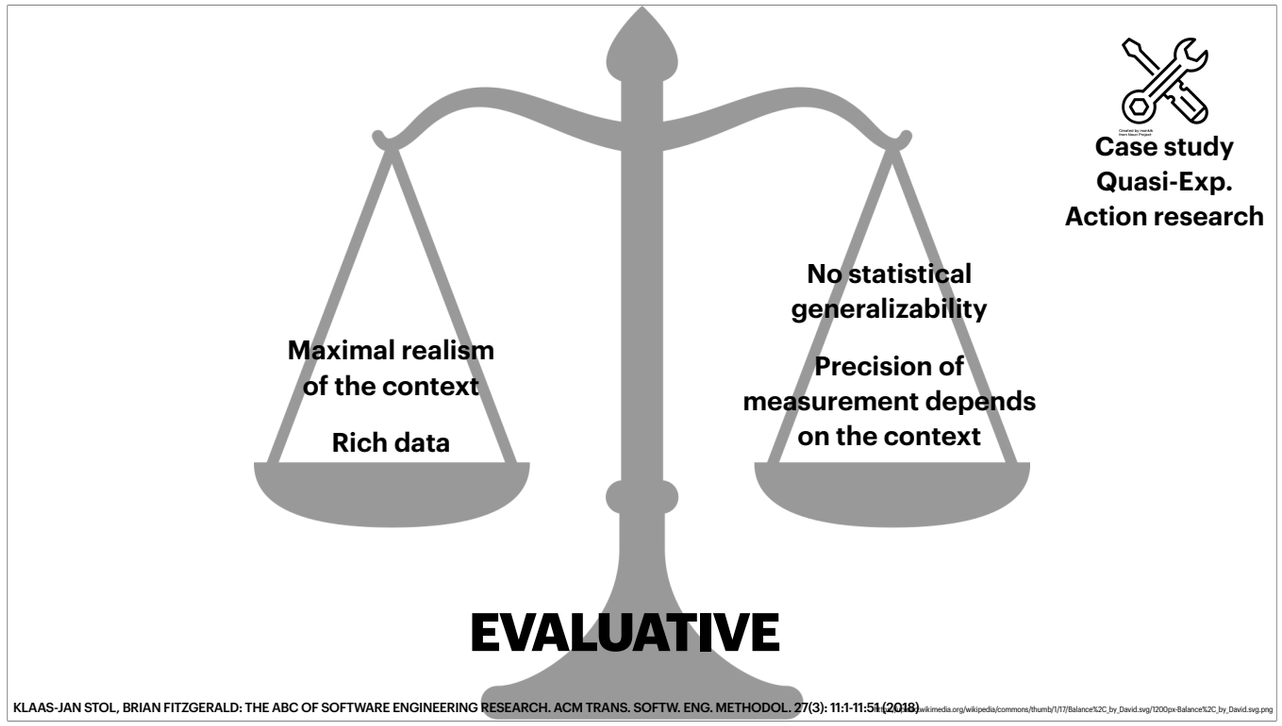
Nature reserve: Natural, pre-existing setting (in vivo), but some level of intrusion due to the deliberate manipulation of aspects of the setting; study affected by confounding factors. In a nature reserve, flora and fauna can still thrive as normal, but the reserve facilitates the conduct of research, for example, by placing fences so as to separate the wildlife into different treatment groups and evaluate the
effects of those treatments.
To investigate, evaluate, or compare techniques, practices, processes, or approaches within a real-world and pre-existing setting.
Evaluative case study, quasi-experiment, Action Research; studies may use either quantitative data or qualitative data.
• No statistical generalizability
• Precision of measurement affected by confounding contextual factors



Example of a field experiment.

A longitudinal multiple-case study of variations and reproducibility in software development, from bidding to deployment, on the basis of the same requirement specification. In a call for tender to 81 companies, 35 responded. **Four** of them developed the system independently.

Common methods in a field experiment: evaluative case study (like here), quasi-experiment, Action Research



Field experiment is usually an evaluative strategy, to investigate, evaluate, or compare techniques, practices, processes, or approaches within a real-world and pre-existing setting.

QUESTION

Caneill et al. present a case study of Debian, which investigates its evolution, rate of change, use and popularity of programming languages, and use and evolution of licenses within Debian. No attempt at generalisation is made.

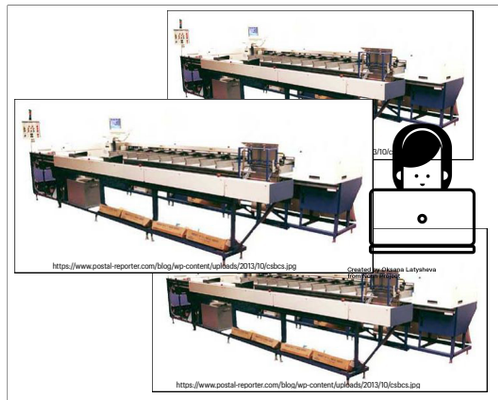
(A) FIELD STUDY

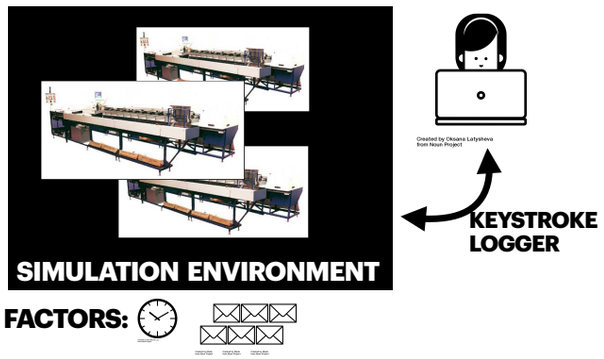
(B) FIELD EXPERIMENT

There is no discussion of experimentation at all, so this is a field study

Greenhouse, flight simulator: Contrived setting (in virtuo) created specifically for a study to represent a concrete type of setting. Environment is created by the researcher to study behavior of actors.
To study behavior of participants or systems in a controlled setting that resembles a real-world, concrete class of settings as closely as possible.
Simulation games, management games, instrumented multiplayer games; quantitative or qualitative data, depending on the simulation instrument.

- Generalizability reduced as setting is designed to mirror a specific type of setting
- Realism reduced due to artificial setting





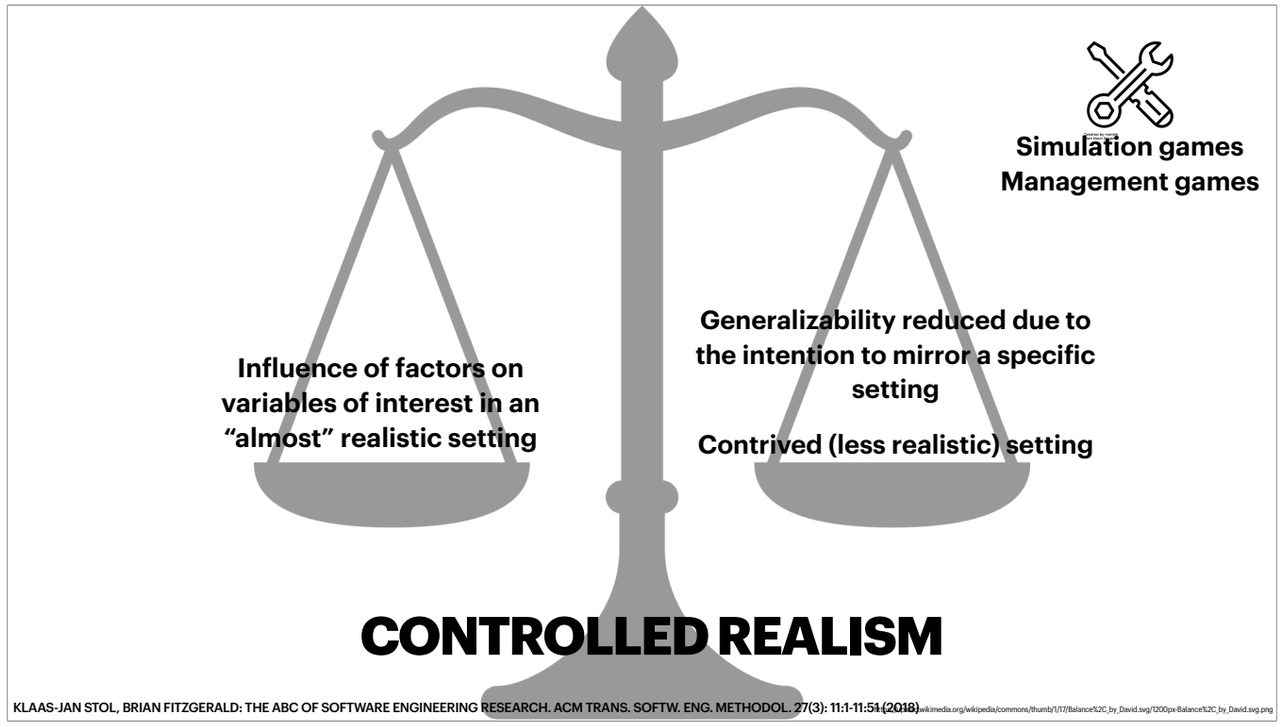
(1) INTERVIEWS + OBSERVATION STUDY

(2) EXPERIMENTAL SIMULATION STUDY

This is an old sorting machine of the US Postal Service. BCS (bar code sorter) supervisors make a series of sort program assignments (between 80 and 100) throughout their shift. That is, they decide when mail should be assigned to available BCS machines, and which sort programs to use. These decisions are highly inter-dependent; for example, decisions about which mail to process at any given point depend on the timing and the nature of prior decisions. Furthermore, the BCS environment changes as a consequence of mail arriving at the GMF (General Mail Facilities).

Factors that can be influenced by researchers: time, heavy duty/light duty day

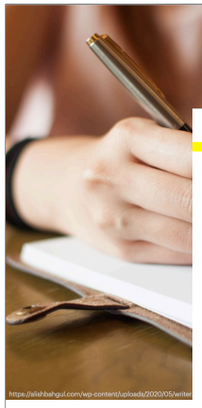
Crucial



KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018). https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Balance%2C_by_David.svg/1200px-Balance%2C_by_David.svg.png

Cleanroom, test tube: Contrived setting (in vitro) created specifically for a study, with high degree of control of all measured variables.
To study with a high degree of precision relationships between variables, or comparisons between techniques; may allow establishment of causality between variables.
Randomized controlled experiments and quasi experiments, comparative evaluations with benchmark studies; usually quantitative
data exclusively.

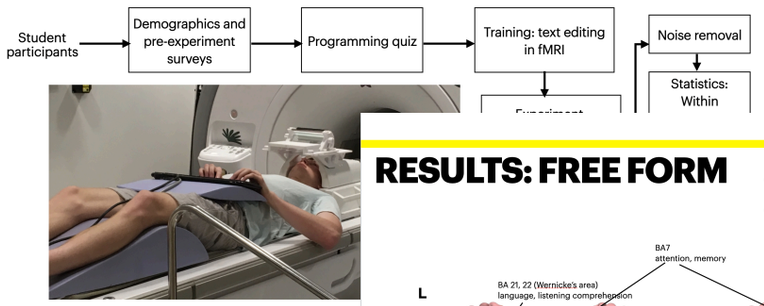
- Abstract or unrealistic context due to highly artificial setting
- Typically scope of problem reduced to study the “essence,” optimizing internal validity at cost of external validity



Is prose writing the same as code writing?

LABORATORY EXPERIMENT

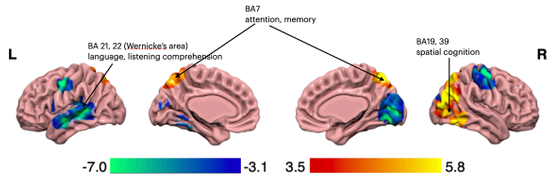
RESEARCH METHOD



RYAN KRUEGER, YU HUANG, XINYU LIU, TYLER SANTANDER, WESTLEY WEIMER, KEVIN LEACH. NEUROLOGICAL DIVIDE

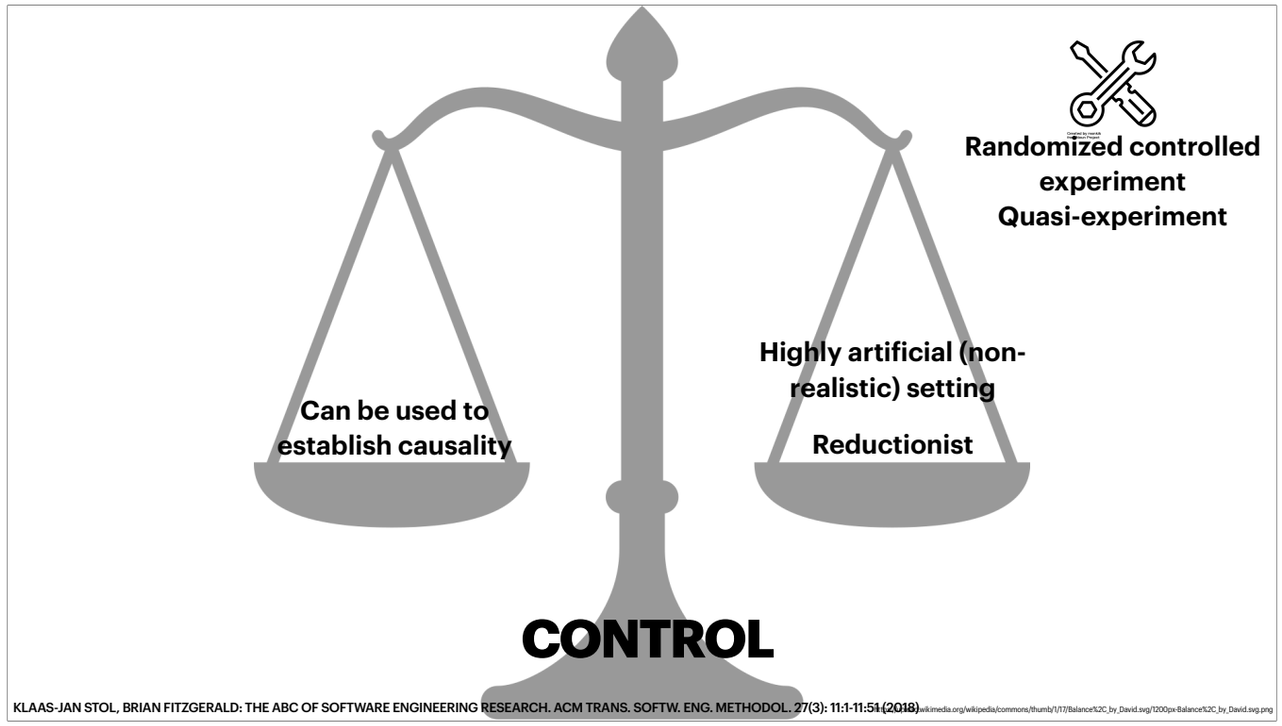
RESULTS: FREE FORM

- Cold: more prose than code
- Hot: more code than prose



KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

47



KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018). https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Balance%2C_by_David.svg/1200px-Balance%2C_by_David.svg.png

QUESTION

Jongeling et al. investigate to what extent results from software engineering studies using sentiment analysis depend on the choice of sentiment analysis tool. Through experimental comparison, finds that tools do not compare with manual labelling, nor do different tools agree with each other.

(A) FIELD STUDY

(B) FIELD EXPERIMENT

(C) EXPERIMENTAL SIMULATION

(D) LAB EXPERIMENT

This might appear odd but this is a lab experiment: participants are systems (sentiment analysis tools) rather than people!

https://images.squarespace-cdn.com/content/v1/56b4be6ec2ea51ac0c659ba4/1555916735706-6O1PC7P8QLT2IP2AEAOK/kef72wd3BTodd8pOm48t6NRtQcg4WhanWrojHLplx7gQa3H78H3Y0txjaiv_0ID0oOvwcdMmMkkDoyUqMSsMWxHk725yiiHCCLfrh8O1z5QPOohDIaIeljMHgDF5CVlOqpeNLcJ8ONk65_fV7StUS6lfA3z_hZ3gTROHKg40oiauNtSL88ZRQhKg2xy4MQPt_AAlcqPvsV/GVaSGkImcw/image-asset.jpeg?format=2500w

Courtroom: Neutral setting; may be actively designed to nullify the context, so that “responses” are in relation to some stimulus (question or instructions), independent of setting.

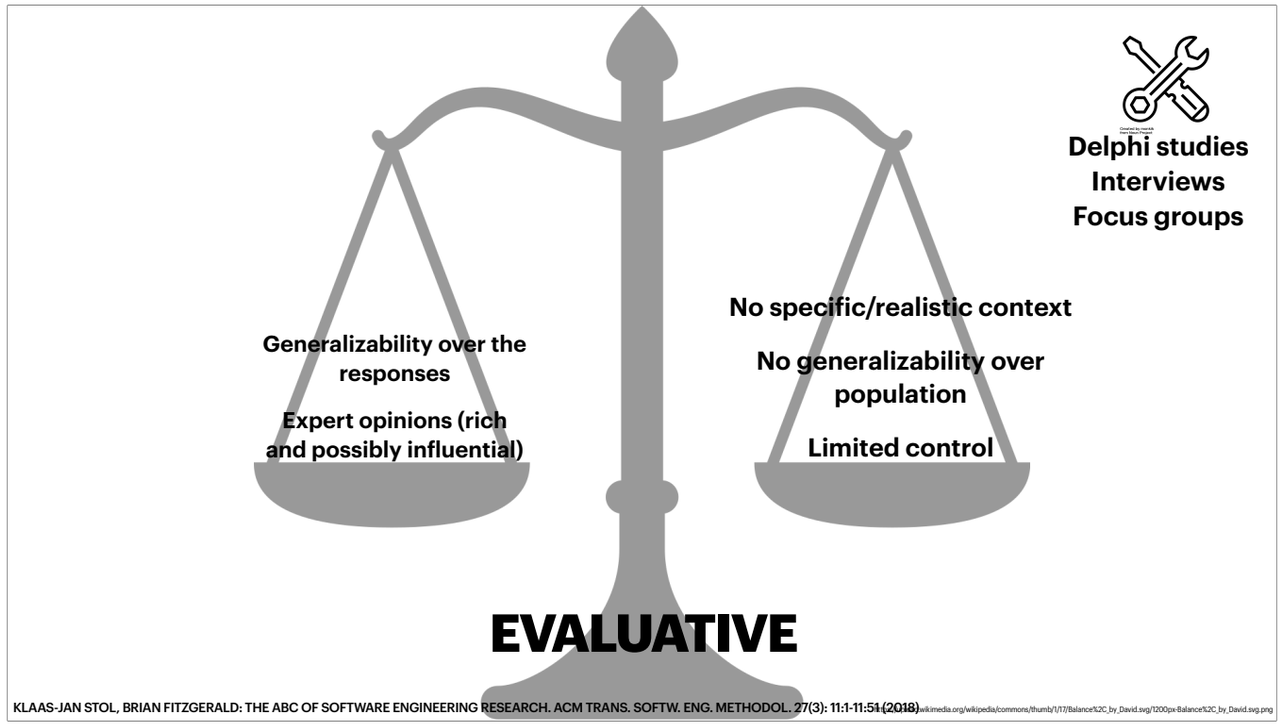
To elicit information from subjects for purposes of evaluation or study of an object. A judgment study involves gathering empirical data from a group of participants who are asked to judge or rate behaviors, to respond to a request or “stimulus” offered by a researcher, or to discuss a given topic of interest. Judgment studies rely on systematic sampling rather than representative sampling and should involve experts, appropriately informed to respond to a certain question or stimulus (**jury**). The goal of a judgment study is to seek generalizability over the responses, rather than generalizability to a population of actors.

Delphi studies, interview studies, focus group, evaluation studies; use of qualitative and/or quantitative data.

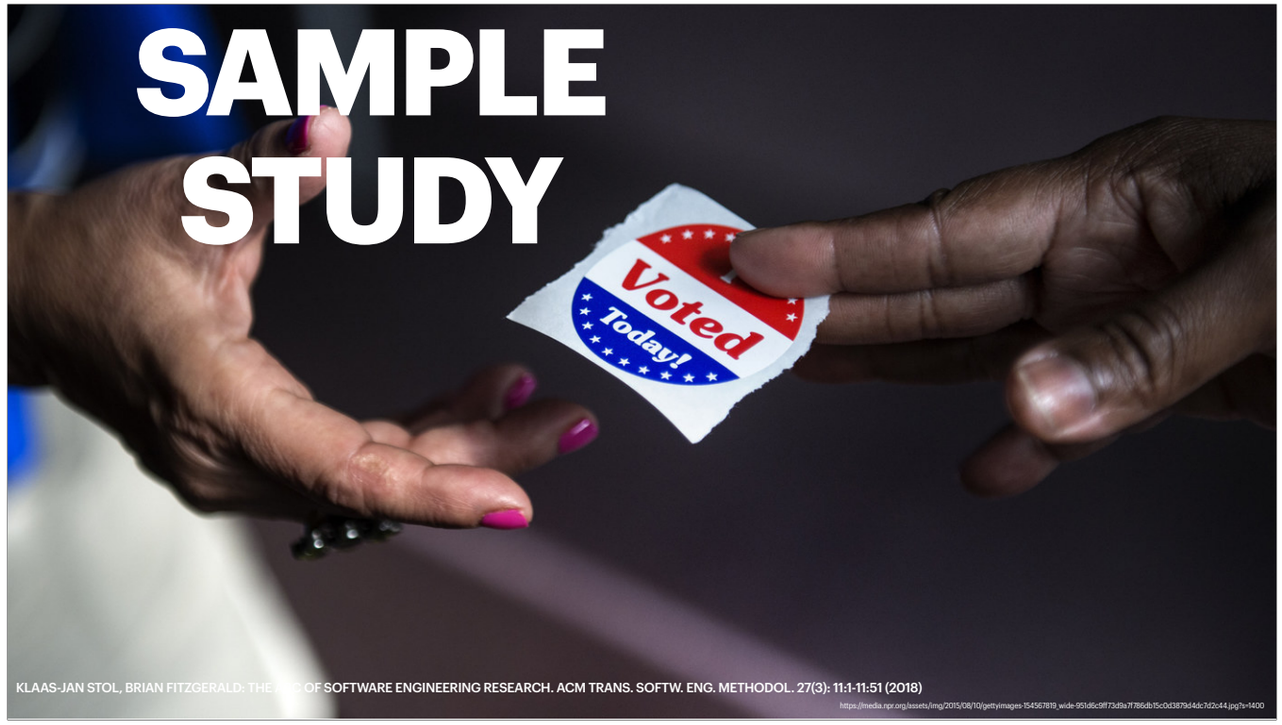
• Responses not related to any specific or realistic context
• Less generalizability than sample studies due to lack of representative sampling
• Less control and precision of measurement than a lab. exp.

CHARALAMBOS L. IACOVOU, ROBBIE T. NAKATSU: A RISK PROFILE OF OFFSHORE-OUTSOURCED DEVELOPMENT PROJECTS. COMMUN. ACM 51(6): 89-94 (2008)

Three rounds of expert interviews (so called Delphi method).



KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018). https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Balance%2C_by_David.svg/1200px-Balance%2C_by_David.svg.png



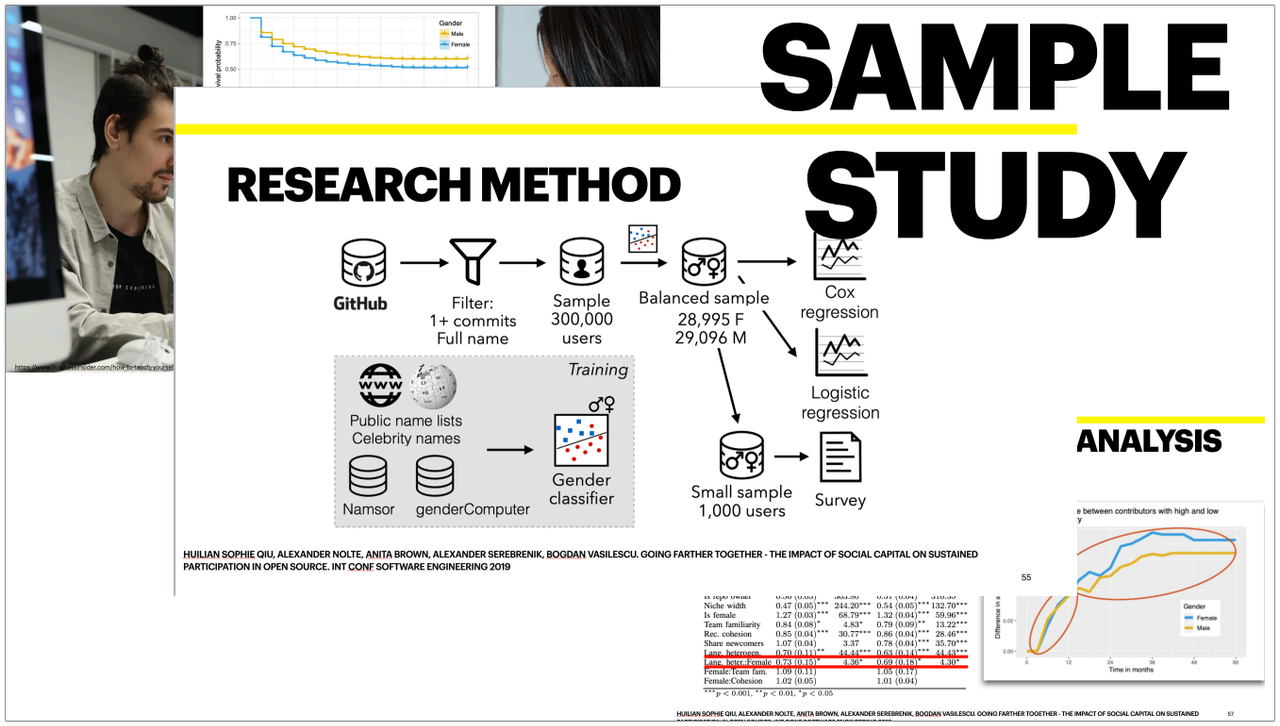
Referendum: Neutral setting. Limited level of precision of measurement; no variables are manipulated. The researcher must deal with whatever data is collected.

To study the distribution of a particular characteristic in a population (of people or systems), or the correlation between two or more characteristics in a population. Information is sought of the subjects.

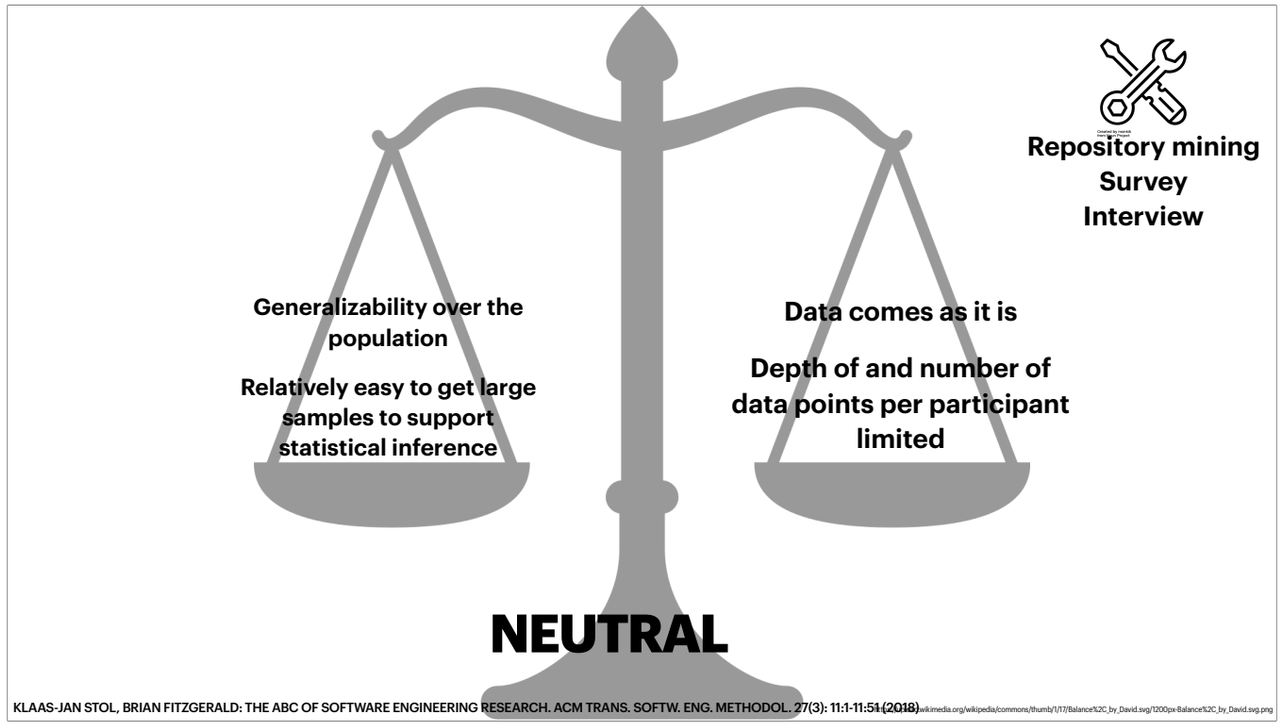
Software repository mining, questionnaires, interviews; analysis includes correlational methods, e.g., regression. Typically, quantitative data (e.g., Likert scales) but can include qualitative data.

• Reductionist—depth of and number of data points per participant limited

• Data collection not “interactive”: no option to clarify questions; repository data comes as is, no opportunity to manipulate variables, only to correlate them



Both GitHub analysis and survey are sample studies: they address a certain (sub)population of people or projects and aim to derive conclusions that can be generalised to the entire population. The context is neutral since differences between projects or participants are blurred or completely muted.



KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018). https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Balance%2C_by_David.svg/1200px-Balance%2C_by_David.svg.png

QUESTION

Joblin et al. investigates the evolution of developer coordination in open-source projects, through an analysis of 18 large open-source software projects, with findings that aim to generalize.

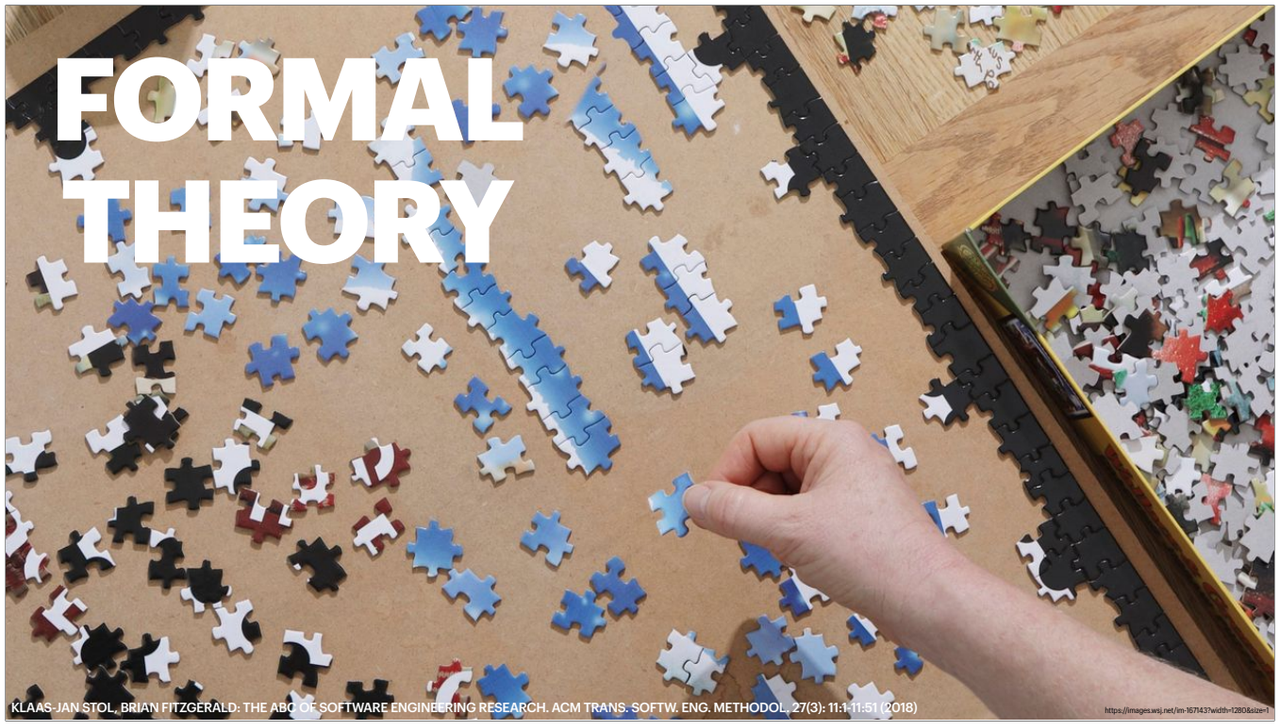
(A) FIELD STUDY

(B) SAMPLE STUDY

(C) JUDGEMENT STUDY

(D) LAB EXPERIMENT

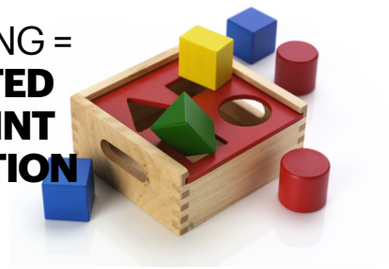
This is a sample study, no discussion of specially selected experts



Jigsaw puzzle: Nonempirical setting; typically a research office or library.
To develop a conceptualization, framework, or theory on a topic. Focus is on formulating relations among concepts, or explanations that hold for a wide range of contexts.
Conceptual reasoning, concept development, development of propositions and/or hypotheses; framework development.
• Low on realism: does not consider a specific context but rather abstract concepts
• No manipulation of variables or measurement (no empirical information is gathered)

SOFTWARE ENGINEERING =

DISTRIBUTED CONSTRAINT SATISFACTION PROBLEM



https://onwego.co.uk/blog/wp-content/uploads/2012/10/square-peg.jpg

JAMES D. HERBSLEB: BUILDING A SOCIO-TECHNICAL THEORY OF COORDINATION: WHY AND HOW (OUTSTANDING RESEARCH AWARD). SIGSOFT FSE 2016: 2-10

Some pieces fit in some of the holes, some do not

SOFTWARE ENGINEERING = **DISTRIBUTED CONSTRAINT SATISFACTION PROBLEM**

https://onwego.co.uk/blog/wp-content/uploads/2012/10/square-peg.jpg

https://ae01.alicdn.com/kf/H5cdb699c00fc44cb85f388f4d38e01b4E.jpg

PEOPLE MUST ORGANIZE TO SOLVE THIS PROBLEM BY USING **CAPABILITIES** AND **COORDINATION MECHANISMS** AT THEIR DISPOSAL.

JAMES D. HERBSLEB: BUILDING A SOCIO-TECHNICAL THEORY OF COORDINATION: WHY AND HOW (OUTSTANDING RESEARCH AWARD). SIGSOFT FSE 2016: 2-10

SOFTWARE ENGINEERING = **DISTRIBUTED CONSTRAINT SATISFACTION PROBLEM**

https://onwego.co.uk/blog/wp-content/uploads/2012/10/square-peg.jpg

https://ae01.alicdn.com/kf/H5cdb699c00fc44cb85f388f4d38e01b4E.jpg

PEOPLE MUST ORGANIZE TO SOLVE THIS PROBLEM BY USING **CAPABILITIES** AND **COORDINATION MECHANISMS** AT THEIR DISPOSAL.

CONGRUENCE = DEGREE OF MATCH BETWEEN THE COORDINATION PROBLEM AND THE ORGANIZATION'S COORDINATING ACTIVITIES.

JAMES D. HERBSLEB: BUILDING A SOCIO-TECHNICAL THEORY OF COORDINATION: WHY AND HOW (OUTSTANDING RESEARCH AWARD). SIGSOFT FSE 2016: 2-10

SOFTWARE ENGINEERING = **DISTRIBUTED CONSTRAINT SATISFACTION PROBLEM**

https://onwego.co.uk/blog/wp-content/uploads/2012/10/square-peg.jpg

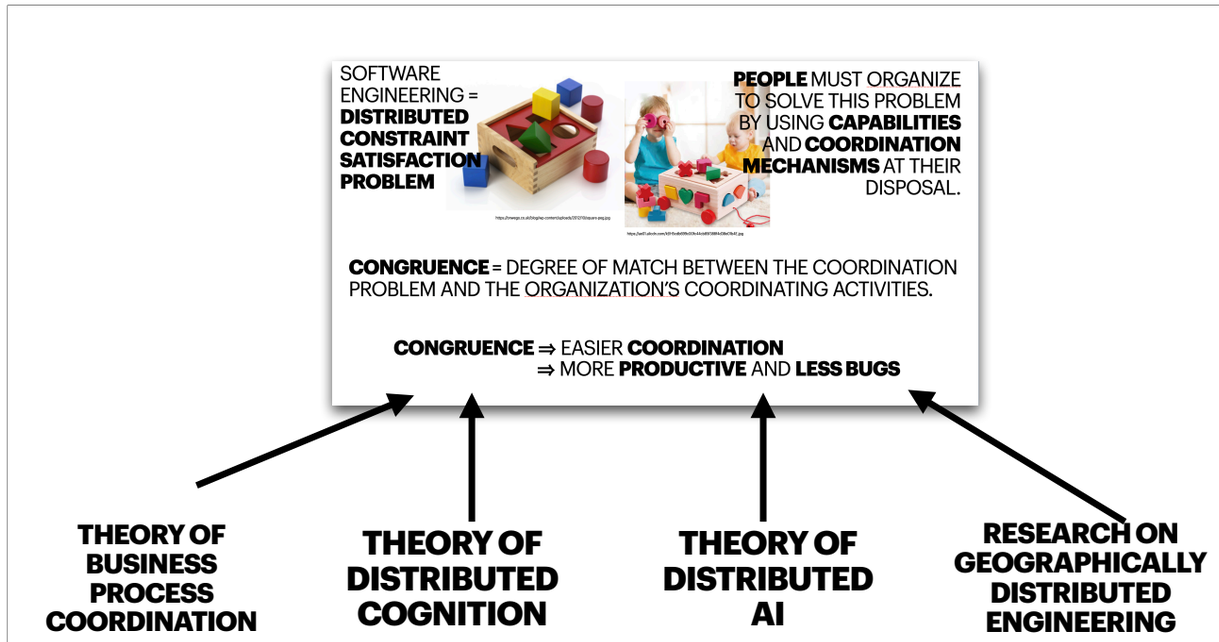
https://ae01.alicdn.com/kf/H5cdb699c00fc44cb85f388f4d38e01b4E.jpg

PEOPLE MUST ORGANIZE TO SOLVE THIS PROBLEM BY USING **CAPABILITIES** AND **COORDINATION MECHANISMS** AT THEIR DISPOSAL.

CONGRUENCE = DEGREE OF MATCH BETWEEN THE COORDINATION PROBLEM AND THE ORGANIZATION'S COORDINATING ACTIVITIES.

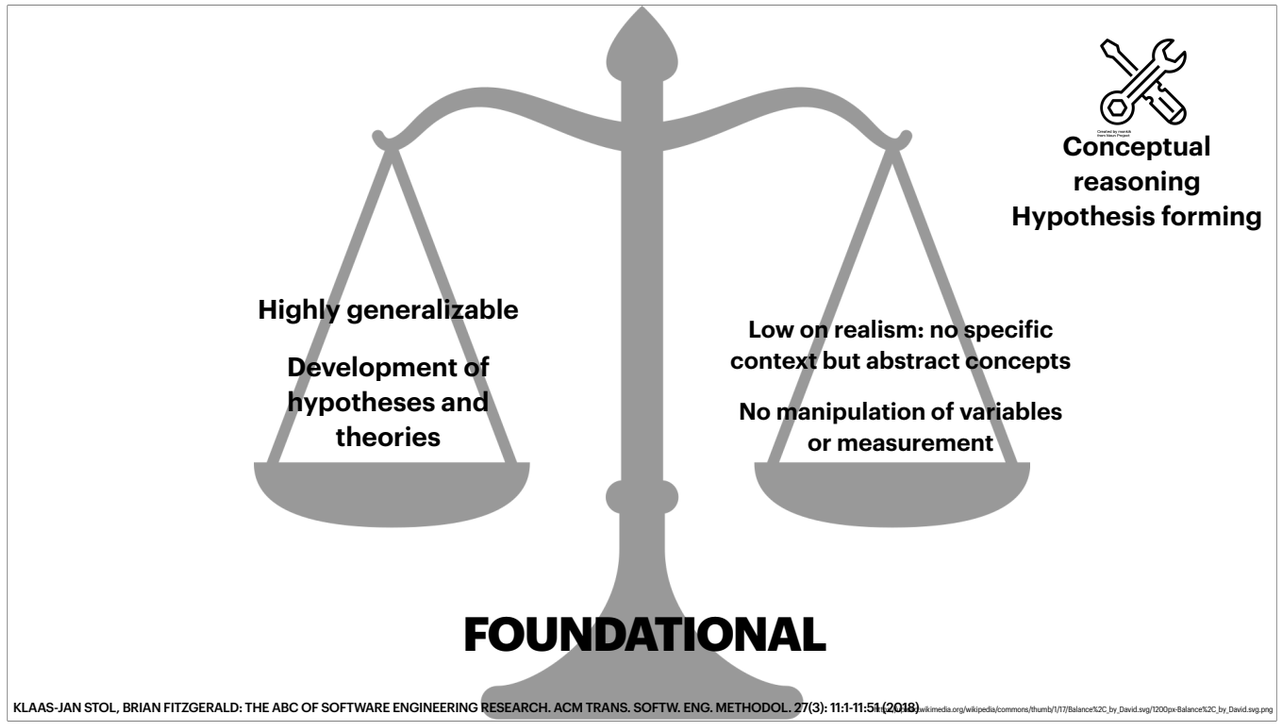
CONGRUENCE ⇒ EASIER **COORDINATION**
⇒ MORE **PRODUCTIVE** AND **LESS BUGS**

JAMES D. HERBSLEB: BUILDING A SOCIO-TECHNICAL THEORY OF COORDINATION: WHY AND HOW (OUTSTANDING RESEARCH AWARD). SIGSOFT FSE 2016: 2-10

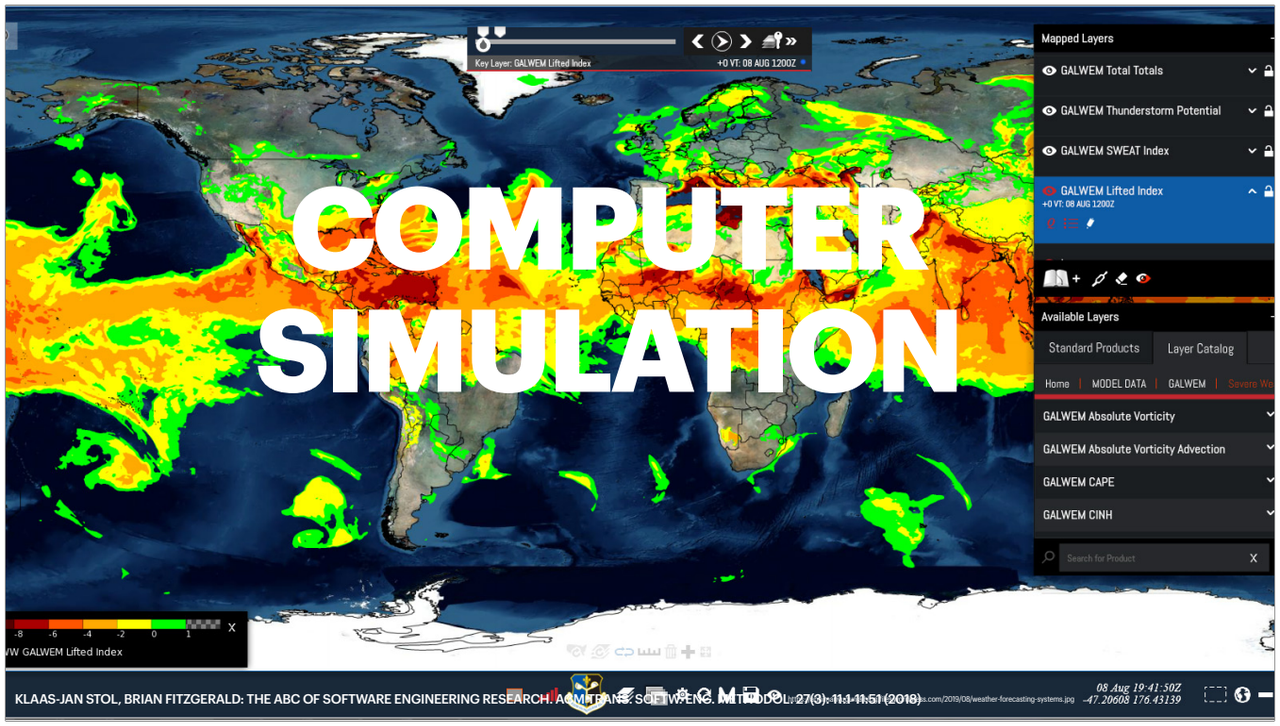


JAMES D. HERBSLEB: BUILDING A SOCIO-TECHNICAL THEORY OF COORDINATION: WHY AND HOW (OUTSTANDING RESEARCH AWARD). SIGSOFT FSE 2016: 2-10

This theory has been based on the scientific literature from adjacent fields, i.e., not really empirically, (and partially inspired by the author's experience) but it has consequences for us as empirical researchers since we can design empirical studies based on theories

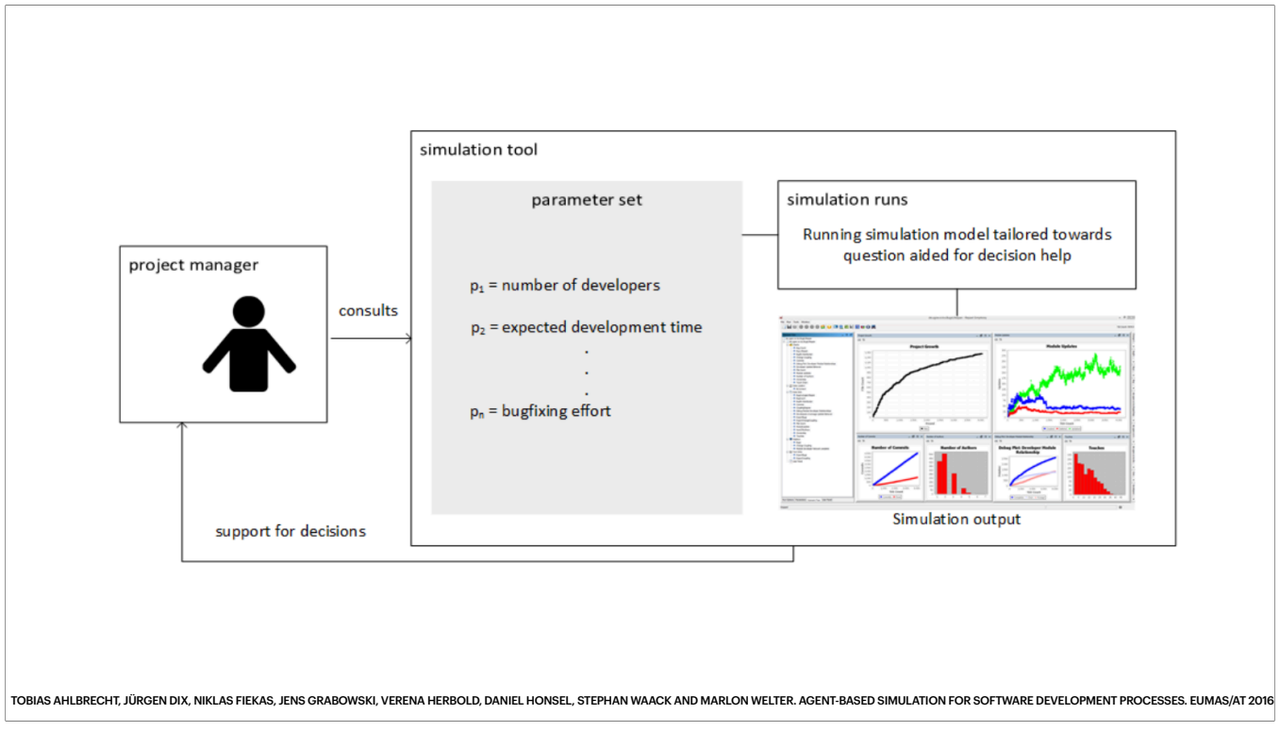


KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018). https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Balance%2C_by_David.svg/1200px-Balance%2C_by_David.svg.png



Forecasting system: Nonempirical setting (in silico); no recording of observations in the real world. There are no actors (people, real-world systems) or real-world behavior: everything is specified in the simulation.
To model a particular system or phenomenon that facilitates evaluation of a large number of complex scenarios that are captured
in the preprogrammed model.
Development of software programs that contain symbolic representations of all variables a researcher considers important; usually these variables are derived and calibrated based on prior empirical studies.

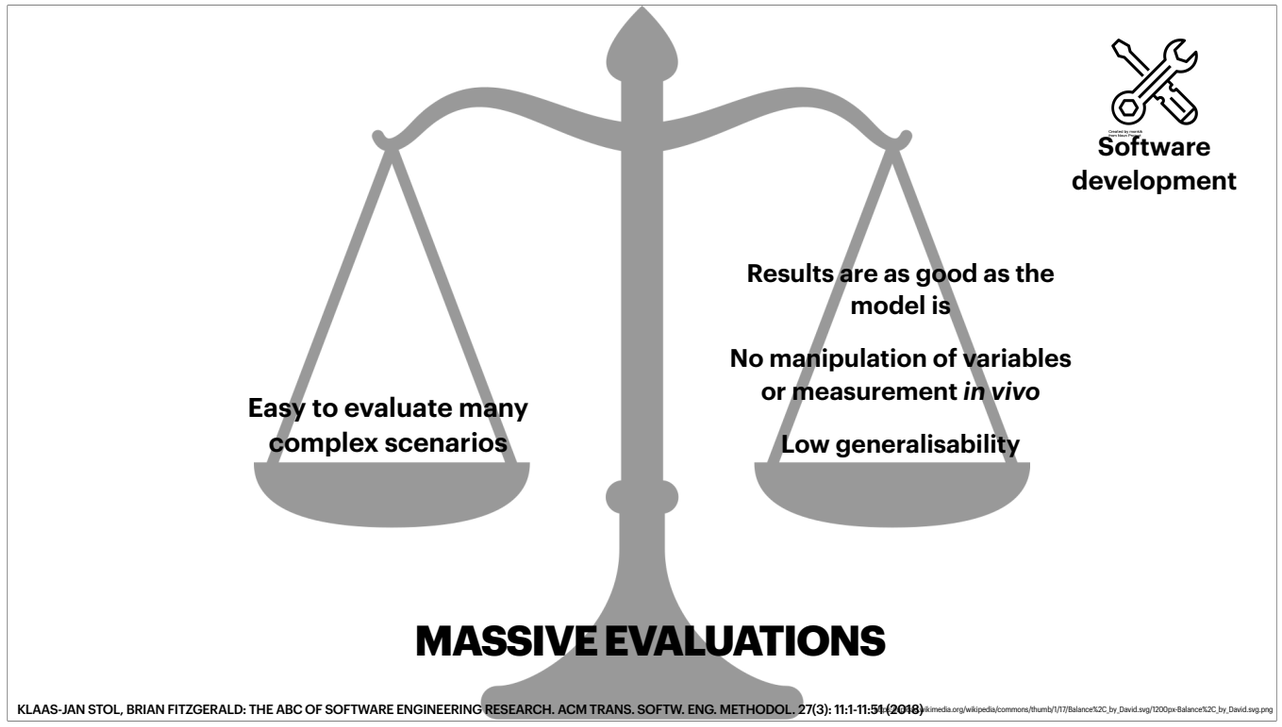
- No manipulation of variables or precision of measurement (no empirical data is gathered)
- Results will be as good as the accuracy of the model representing the simulated system
- Low generalizability as it attempts to model a specific class of real-world systems



TOBIAS AHLBRECHT, JÜRGEN DIX, NIKLAS FIEKAS, JENS GRABOWSKI, VERENA HERBOLD, DANIEL HONSEL, STEPHAN WAACK AND MARLON WELTER. AGENT-BASED SIMULATION FOR SOFTWARE DEVELOPMENT PROCESSES. EUMAS/AT 2016

The main idea is to view the software process as agent-based simulation in a multiagent system (MAS). This approach requires to combine three different areas: (1) mining data and patterns from projects done in the past, (2) modeling the software development process in a multiagent environment (3) running the simulation on a dedicated and scalable multiagent platform.

Similarly to the experimental simulation study we discussed before (remember the postal services?) the data comes from previous observations. However, in the experimental simulation actual people performed the actual albeit manipulated tasks, while here the process is simulated by agents.



No manipulation of variables or precision of measurement (no empirical data is gathered)

- Results will be as good as the accuracy of the model representing the simulated system
- Low generalizability as it attempts to model a specific class of real-world systems

QUESTION

Using Gregor's taxonomy, Nguyen and Shanks create a framework to analyze requirements engineering creativity research. The authors identified two main implications for further research and practice. First, they called for empirical studies on howto integrate the framework elements into requirements engineering methods to support creativity. Second, they recommended that organizations establish an environment to encourage creative people.

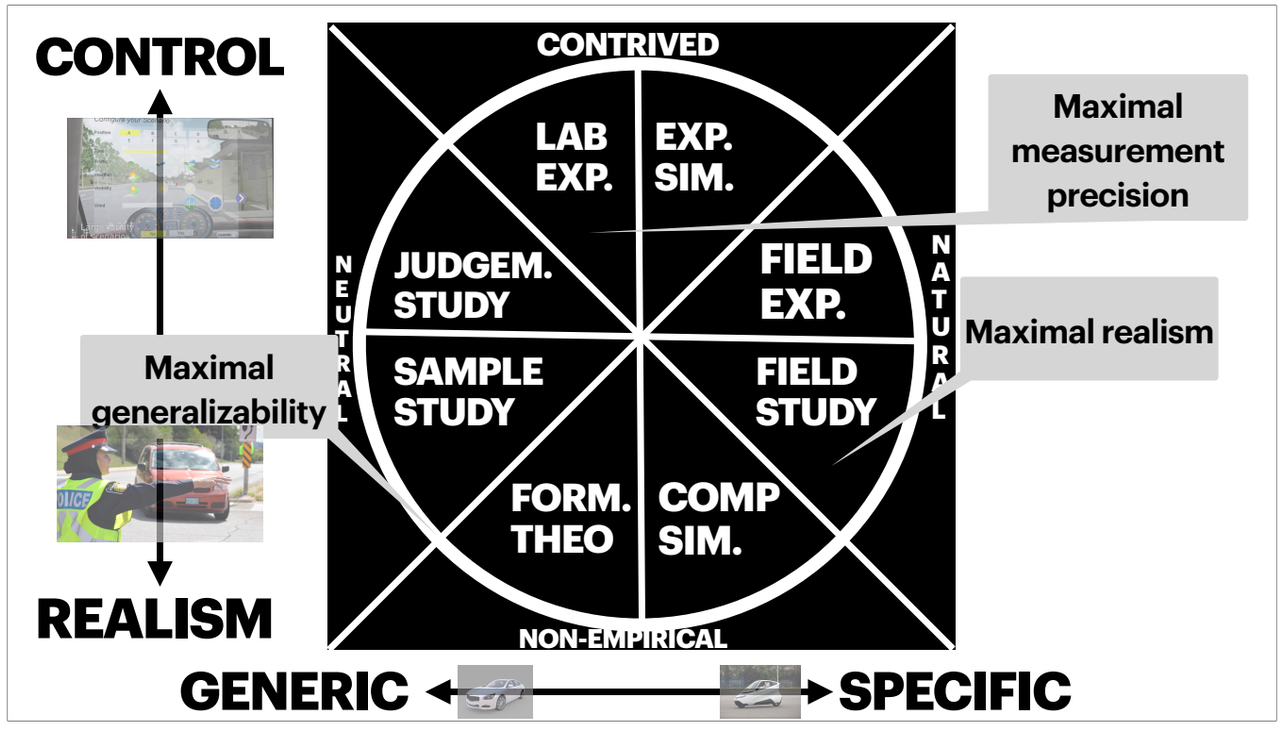
(A) EXPERIMENTAL SIMULATION

(B) COMPUTER SIMULATION

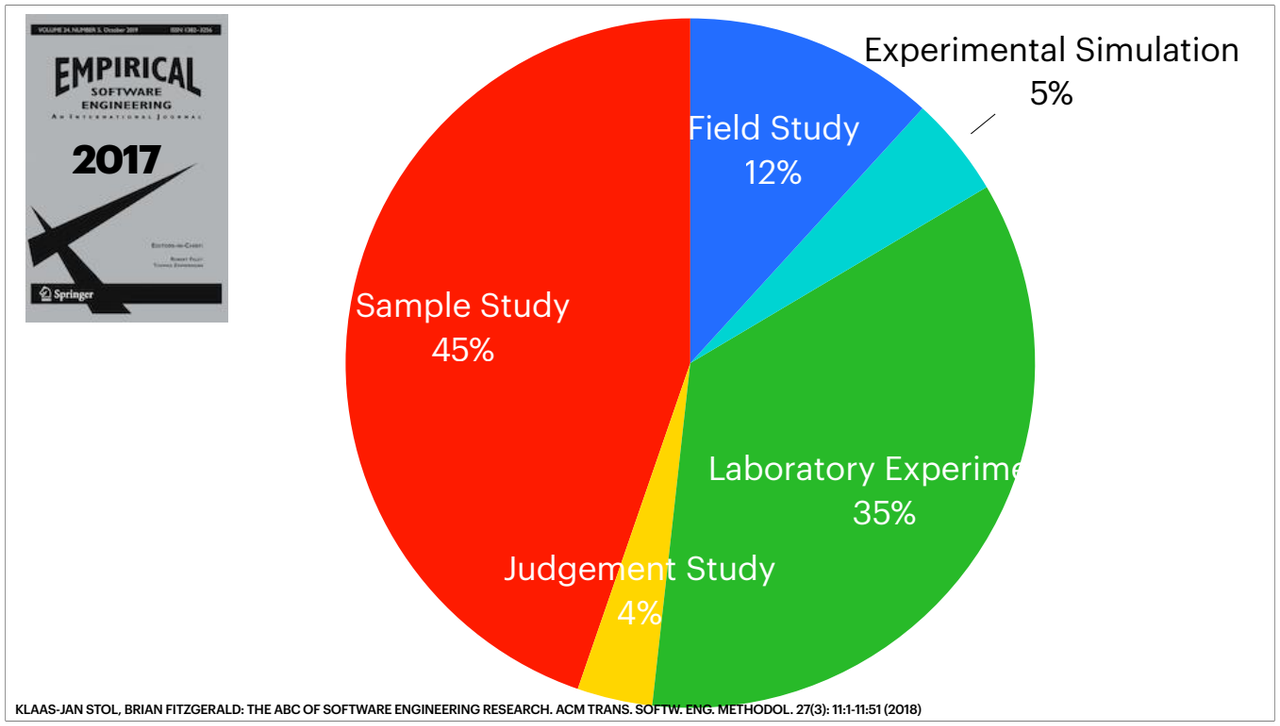
(C) SAMPLE STUDY

(D) FORMAL THEORY

Formal theory



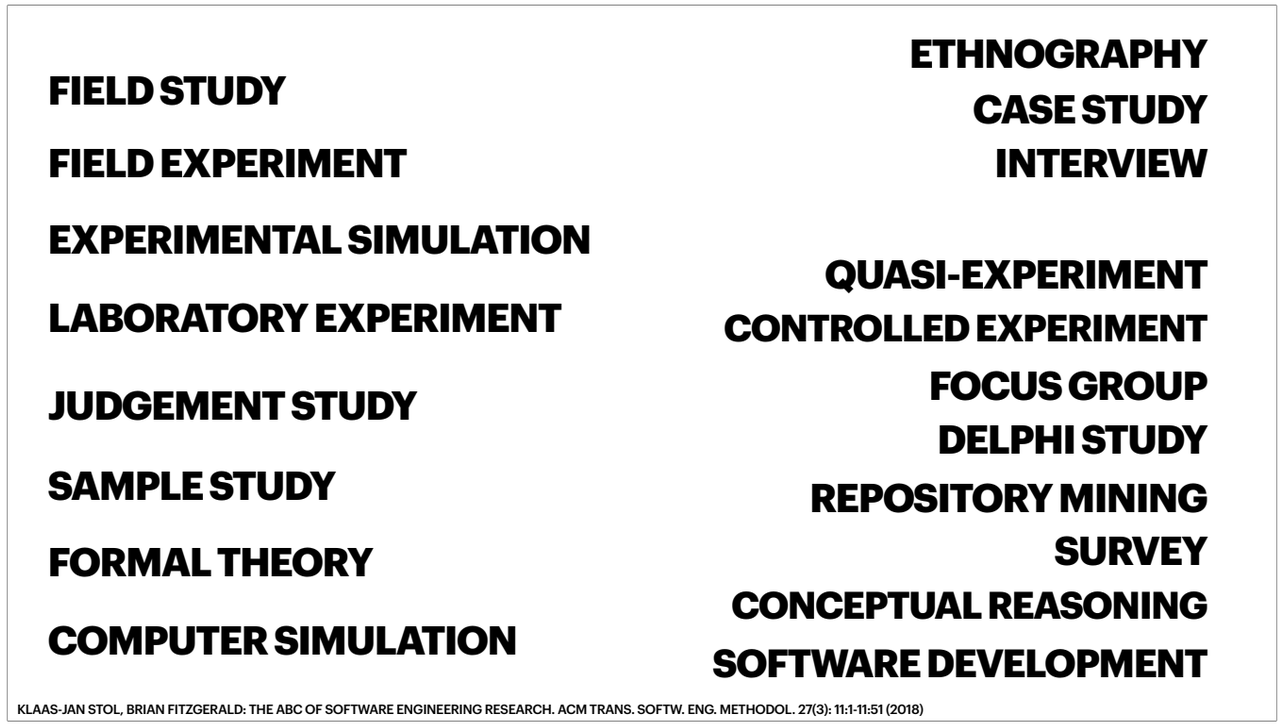
For each quadrant we have two research strategies: laboratory experiment, experimental simulation, field experiments, field studies, computer simulation, formal theory, sample studies, judgement studies. We have discussed them one by one.



So how popular are different research strategies? No Field Experiments, Formal Theories or Computer Simulations in EMSE 2017

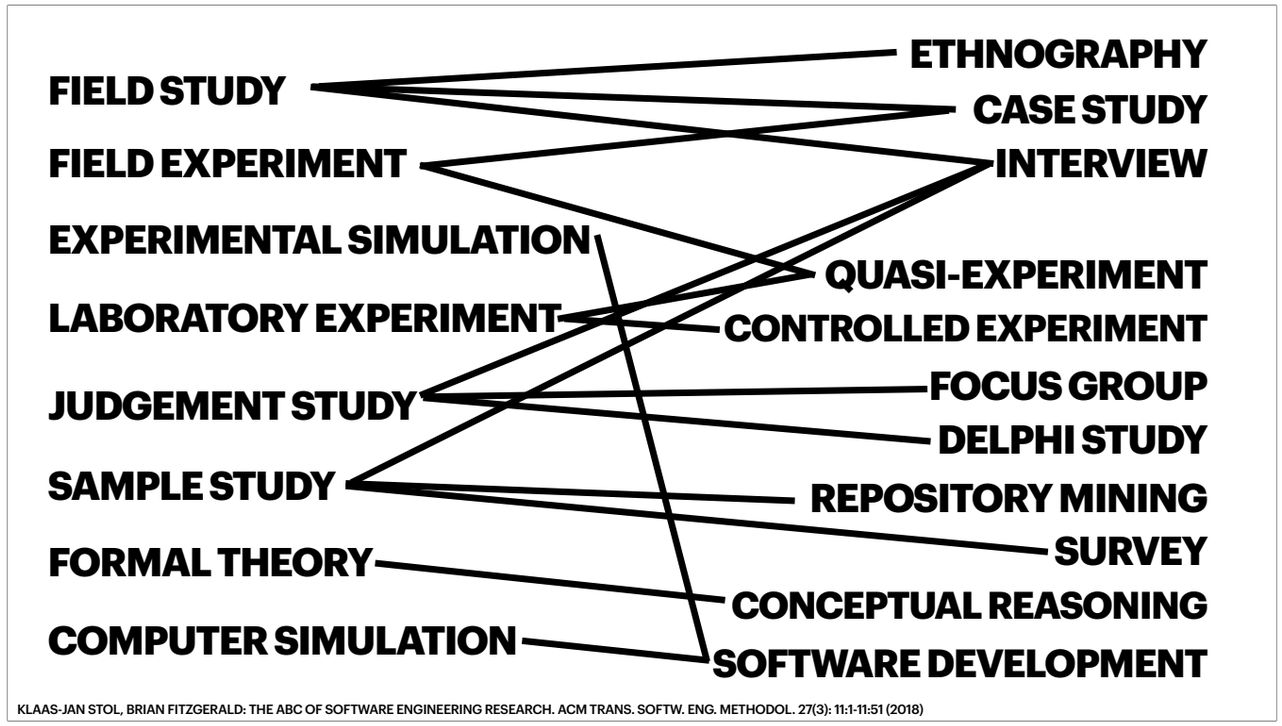
STATEGIES VS METHODS?

Created by monkik
from Noun Project



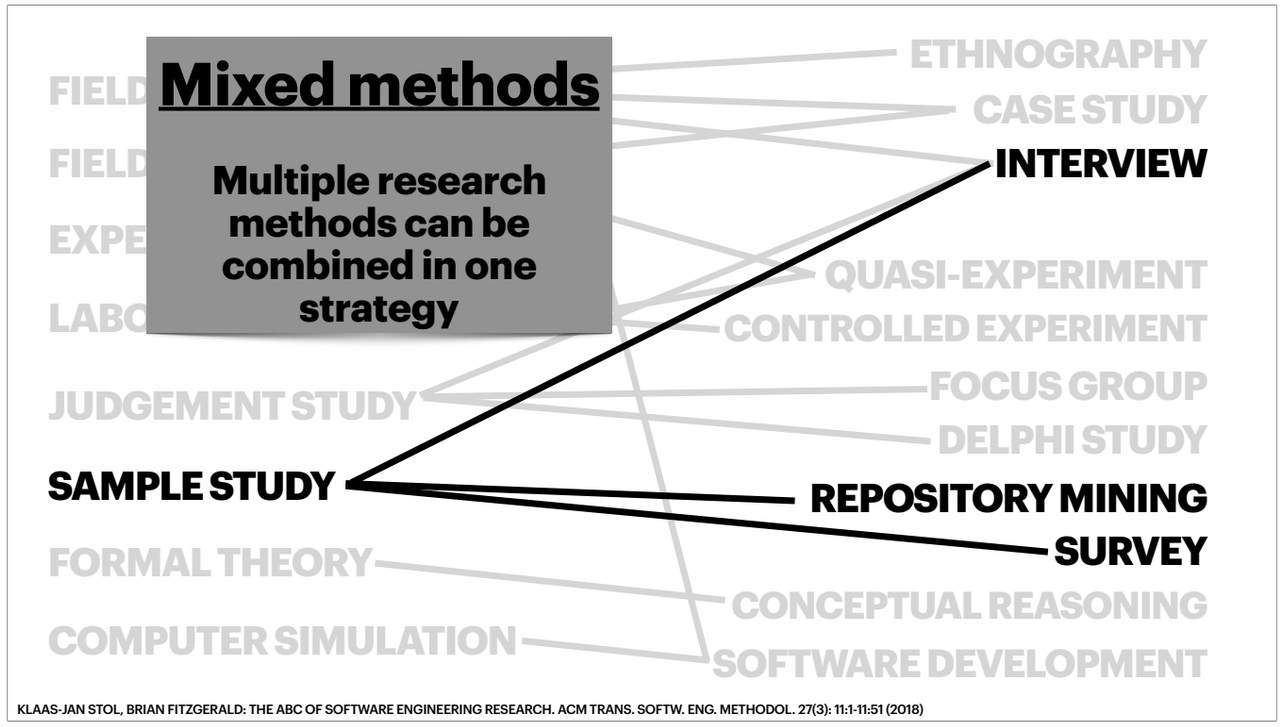
Research strategies on the left, research methods on the right

A **quasi-experiment** is an empirical interventional study used to estimate the causal impact of an intervention on target population without random assignment. Quasi-experimental research shares similarities with the traditional experimental design or randomized controlled trial, but it specifically lacks the element of random assignment to treatment or control.



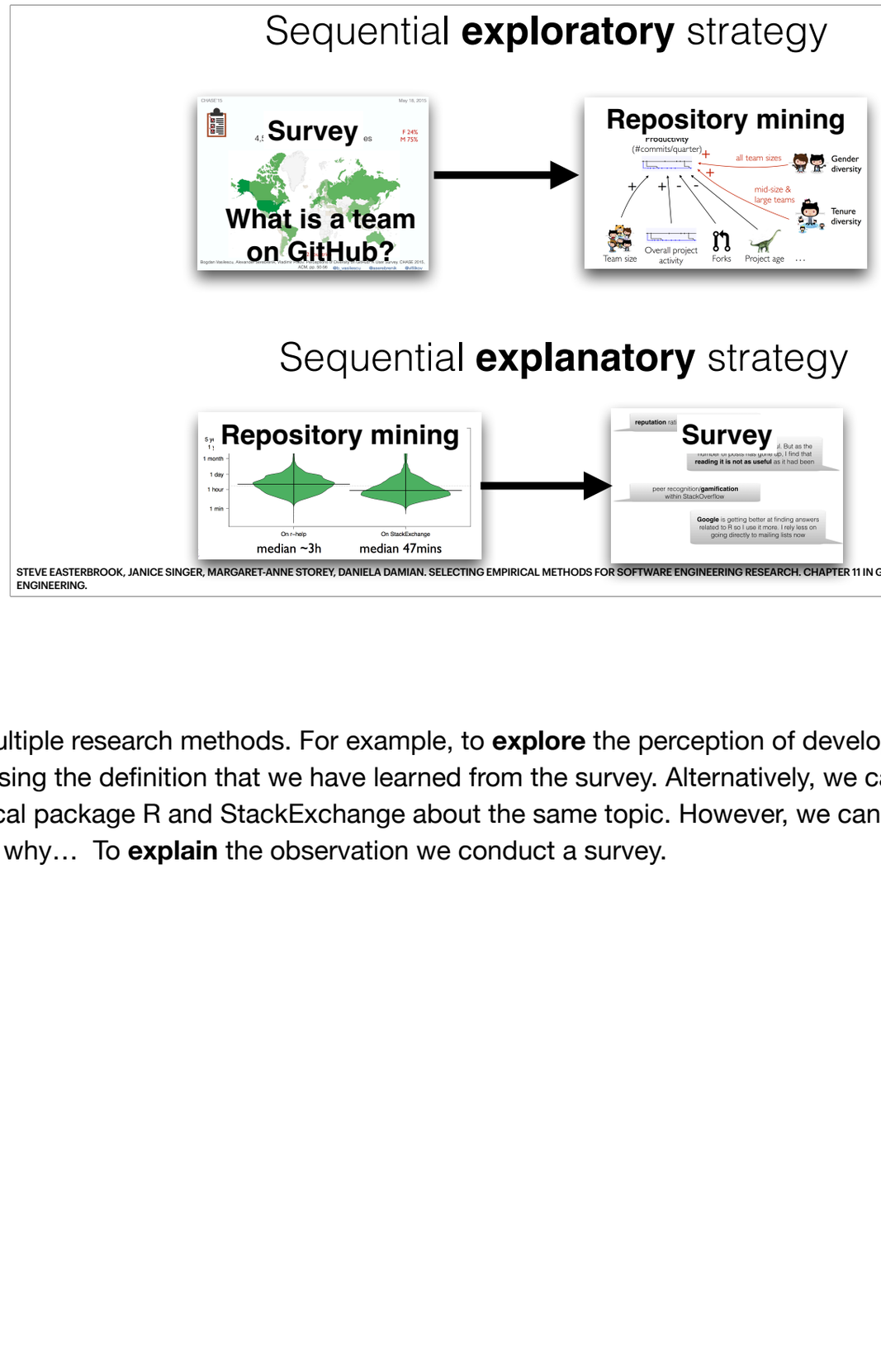
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

The links merely indicate *typical* methods. Software development in case of simulation means development of simulation environments. Of course, repository mining and controlled experiments often require building software tools: repository mining needs analysis tools and controlled experiments need experimental environments. In the following lectures we will consider many of these research methods and will discuss in context of what research strategies they are useful.



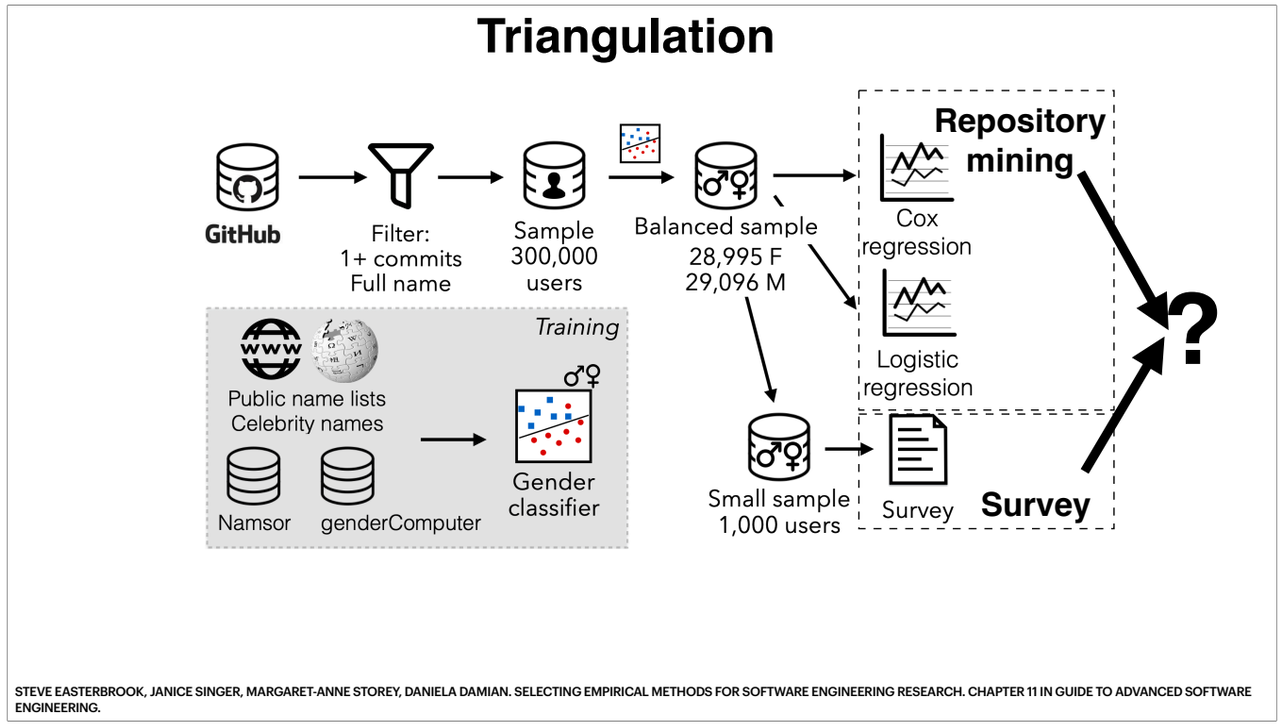
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

Moreover, sometimes multiple research methods can be combined as part of one and the same strategy.



STEVE EASTERBROOK, JANICE SINGER, MARGARET-ANNE STOREY, DANIELA DAMIAN. SELECTING EMPIRICAL METHODS FOR SOFTWARE ENGINEERING RESEARCH. CHAPTER 11 IN GUIDE TO ADVANCED SOFTWARE ENGINEERING.

There are several ways of combining multiple research methods. For example, to **explore** the perception of developers “what is a team on GitHub?” we conduct a survey and then to analyse data from GitHub using the definition that we have learned from the survey. Alternatively, we can first analyse data - here we compared response times on a mailing list about the statistical package R and StackExchange about the same topic. However, we can only observe that the questions are answered faster on StackExchange but we do not know why… To **explain** the observation we conduct a survey.

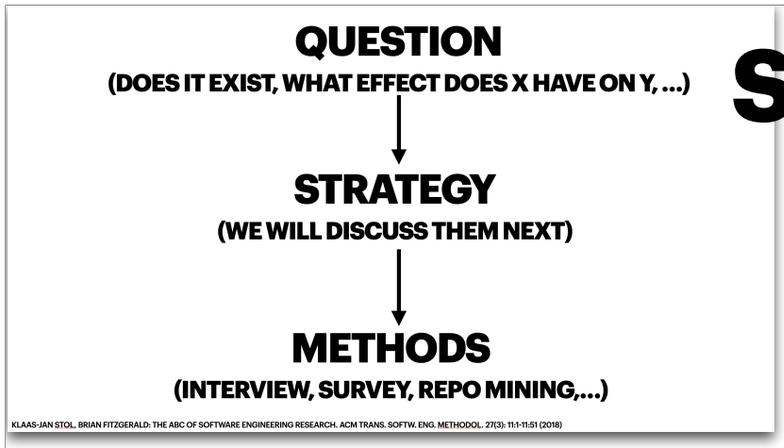


However, sometimes different research methods do not inform each other but are executed in parallel such that we can compare results obtained in different ways. This is called triangulation.

COMING LECTURES

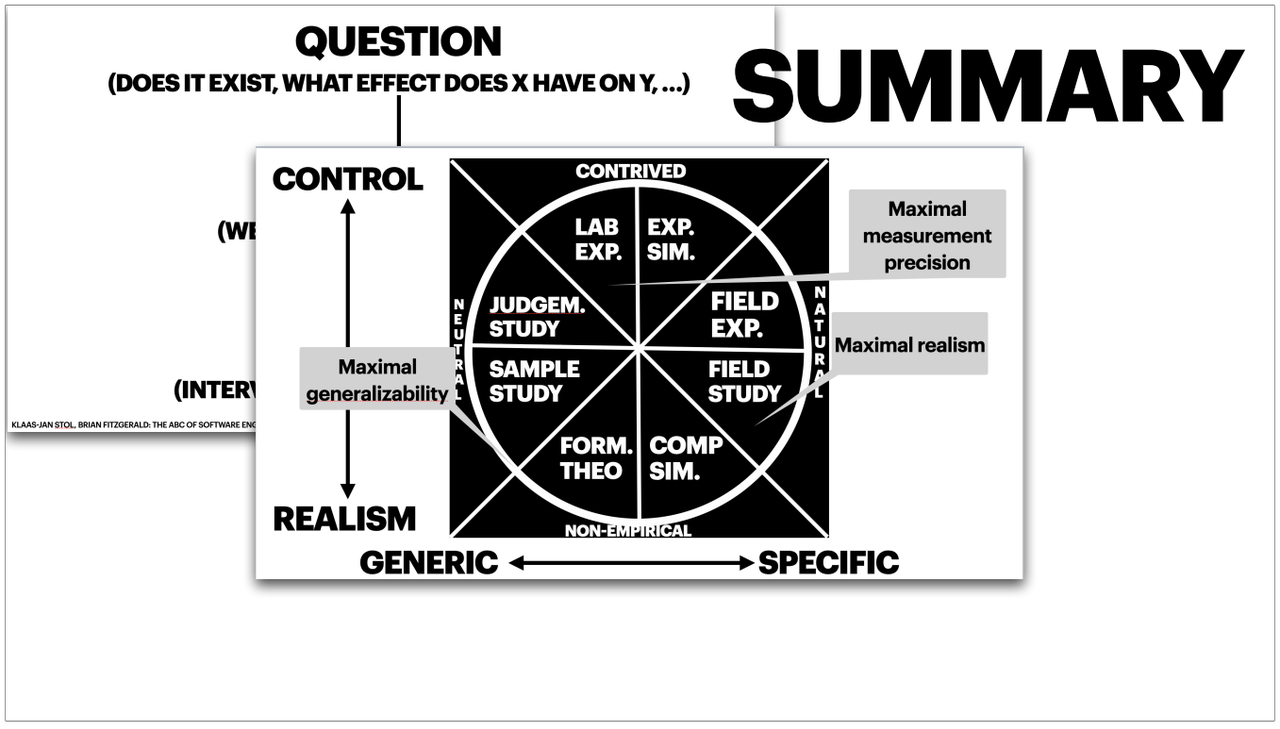
ETHNOGRAPHY
CASE STUDY
INTERVIEW

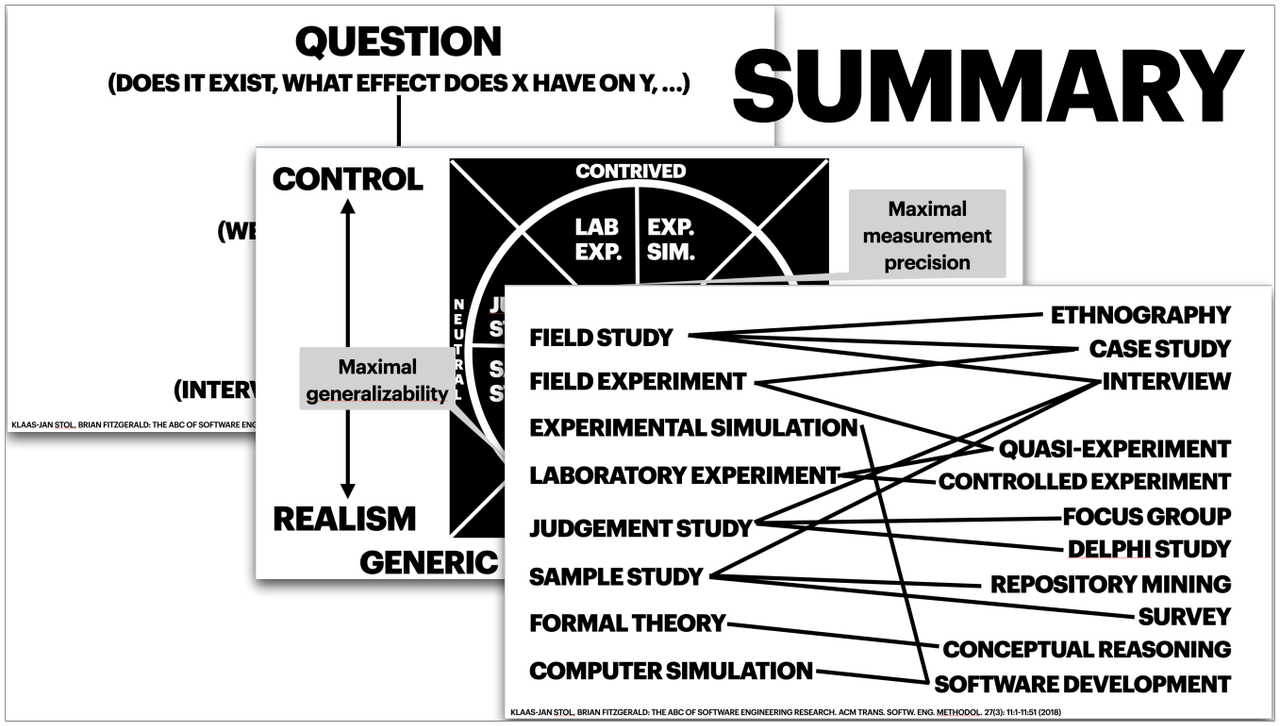
QUASI-EXPERIMENT
CONTROLLED EXPERIMENT
FOCUS GROUP
DELPHI STUDY
REPOSITORY MINING
SURVEY
CONCEPTUAL REASONING
SOFTWARE DEVELOPMENT


QUESTION
(DOES IT EXIST, WHAT EFFECT DOES X HAVE ON Y, ...)
STRATEGY
(WE WILL DISCUSS THEM NEXT)
METHODS
(INTERVIEW, SURVEY, REPO MINING,...)
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)

SUMMARY

SUMMARY


QUESTION
(DOES IT EXIST, WHAT EFFECT DOES X HAVE ON Y, ...)
CONTROL
(WE
(INTERV
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENG
REALISM
CONTRIVED
LAB EXP.
EXP. SIM.
JUDGEM. STUDY
FIELD EXP.
SAMPLE STUDY
FIELD STUDY
FORM. THEO
COMP SIM.
NON-EMPIRICAL
NEUTRAL
NATURAL
Maximal generalizability
Maximal measurement precision
Maximal realism
GENERIC
SPECIFIC


QUESTION
(DOES IT EXIST, WHAT EFFECT DOES X HAVE ON Y, ...)
SUMMARY
CONTROL
(WE
(INTERV
CONTRIVED
LAB EXP.
EXP. SIM.
NEUTRAL
Maximal measurement precision
Maximal generalizability
REALISM
GENERIC
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENG
FIELD STUDY
FIELD EXPERIMENT
EXPERIMENTAL SIMULATION
LABORATORY EXPERIMENT
JUDGEMENT STUDY
SAMPLE STUDY
FORMAL THEORY
COMPUTER SIMULATION
ETHNOGRAPHY
CASE STUDY
INTERVIEW
QUASI-EXPERIMENT
CONTROLLED EXPERIMENT
FOCUS GROUP
DELPHI STUDY
REPOSITORY MINING
SURVEY
CONCEPTUAL REASONING
SOFTWARE DEVELOPMENT
KLAAS-JAN STOL, BRIAN FITZGERALD: THE ABC OF SOFTWARE ENGINEERING RESEARCH. ACM TRANS. SOFTW. ENG. METHODOL. 27(3): 11:1-11:51 (2018)