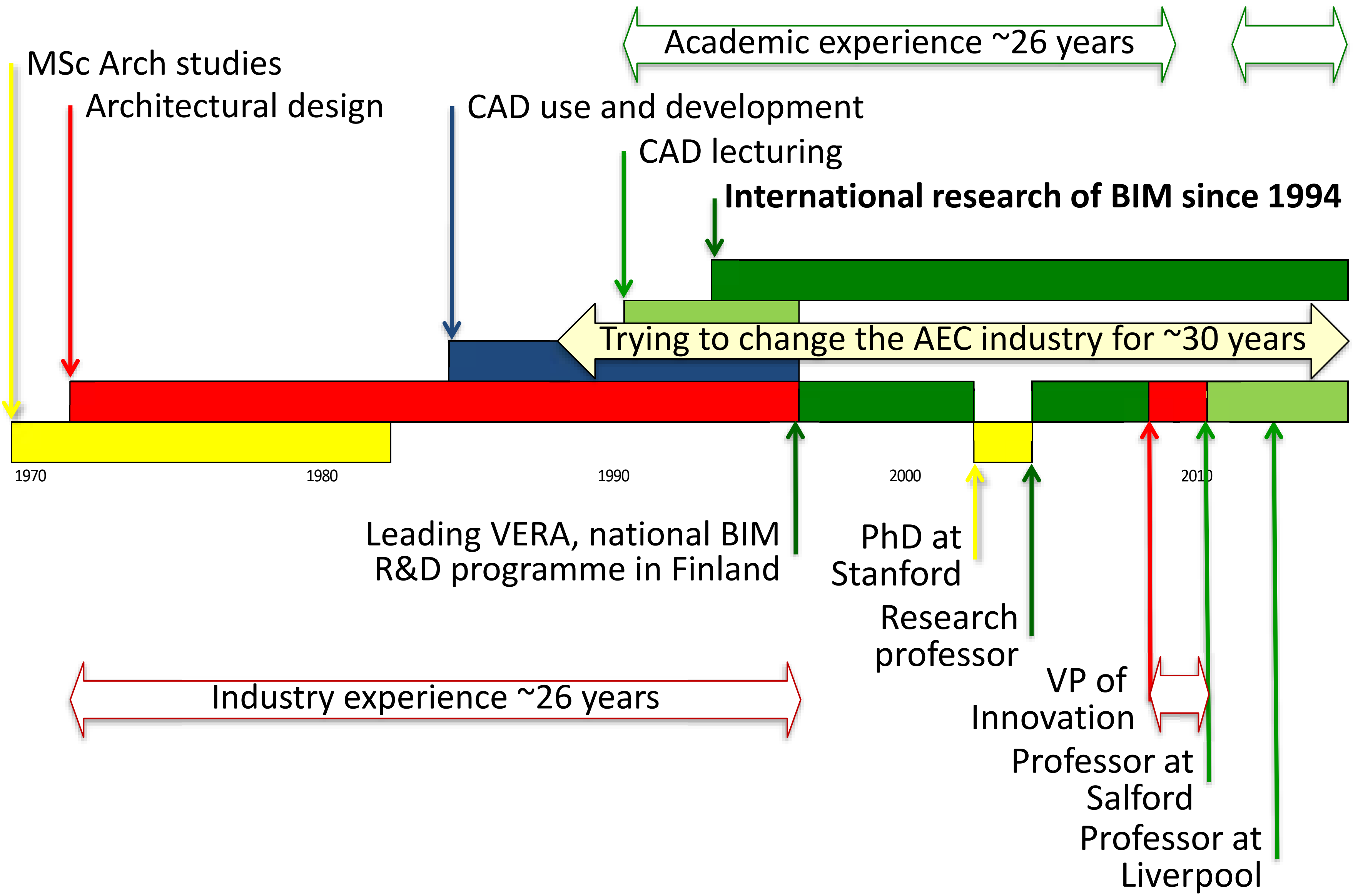
Some thoughts about PhD research related to the Built Environment…

Arto Kiviniemi
Professor of Digital Architectural Design
University of Liverpool

Arto Kiviniemi, personal background



Extensive experience in architectural design 1972-1996

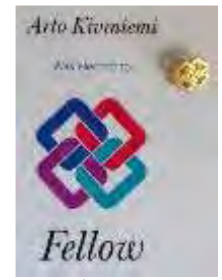
- Architect Office Arto Sipinen 1973…91: **12 first prizes** and several other prizes in architectural competitions, 25 completed projects
  - Cultural centres, concert halls, town halls, university buildings…
- My own office Studio Kivi 1990...1996
  - Started as a design office, from 1991 focus increasingly in software development and ICT consultancy and finally a software vendor
- Partner in Architect Office Arcadia 1991-1996

Architect Office Arto Sipinen 1973-91: Raisio Town Hall, Mikkeli Concert Hall, Tapiola Cultural Centre, Jyväskylä University Chemistry and Physics Laboratories, Helsinki Töölönlahti City Centre

Global visibility since 1996...

One of the globally leading experts of integrated BIM...

- 78 keynotes, 92 invited presentations, 16 journal papers, 40+ refereed conference papers and several working papers, technical reports and chapters in books
- 57 memberships in scientific or organising committees of various international conferences and seminars
- Associate Professor in Ecole de technologies superior (ETS), Montreal, Canada, 2013-2016
- Several leading roles in buildingSMART International (former IAI):
  - Founding member and 1st Chair of BuildingSMART Nordic Chapter 1996-1998
  - International Council and Excel: Chair 1998-2000, Deputy Chair 2000-2002
  - International Technical Management Committee: Chair 2005-2007
  - Technical Advisory Group: Member since 2005
  - buildingSMART Fellow since 2017



- FIATECH, member of several committees 2010-2013: Academic Committee, European Advisory Committee, Interoperability Committee, Conference Planning Committee; Fiatech CETI (Celebration of Engineering & Technology Innovation) Outstanding Researcher Award for the international merits in developing integrated BIM, 2009
- Czech BIM Council, honorary member since 2011
- CIB IDDS (Integrated Design and Delivery Solutions), member of the Core Group since 2010
- ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers): member of the BIM Steering Committee since 2009
- Advisory Board in Politecnico di Milano: Member since 2018
- Steering Committee of Salford Centre for Research and Innovation, Chair 2002-2009
- Scientific Committee in BuildingEnvelopes.org project at the Center of Design Informatics of Harvard University: Member 2001-2004
- Industrial Advisory Board and Technical Advisory Board of CIFE at Stanford University: Member 1999-2005

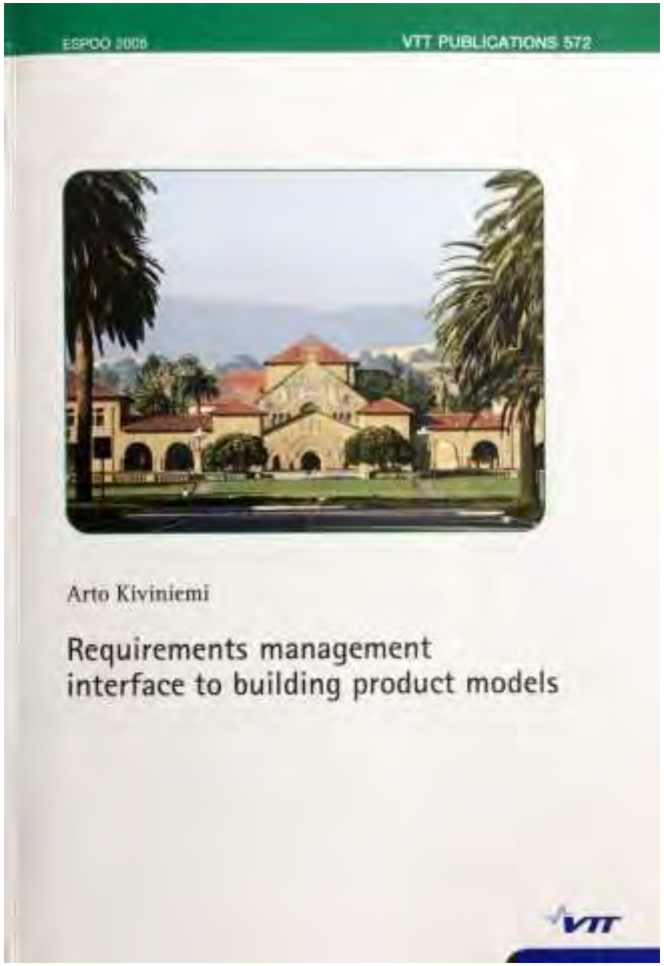
When asking questions, please remember: I am actually hearing disabled…

Basic vs. applied research

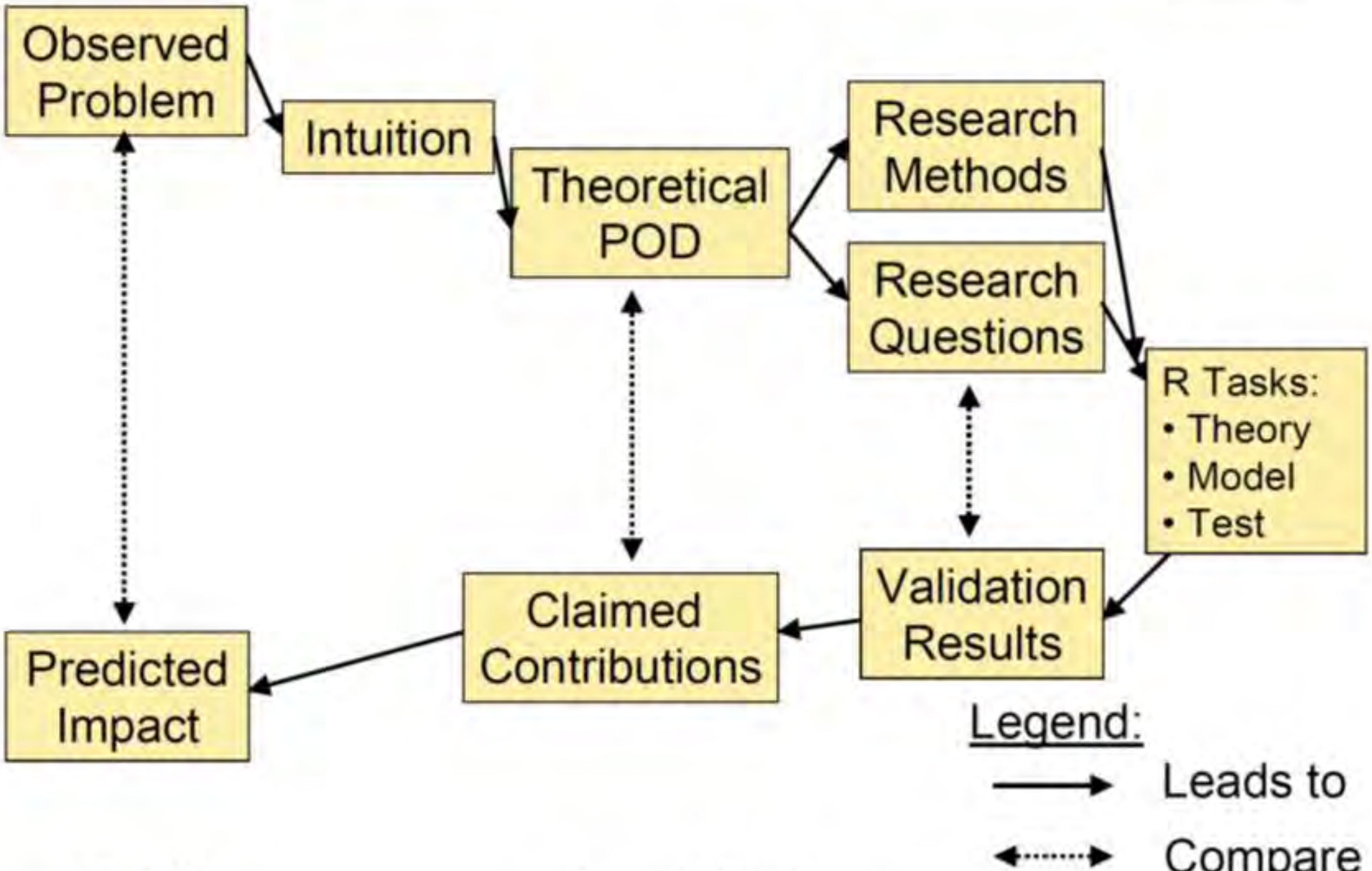
- **Basic research** fills in the knowledge we do not have; it tries to learn things that are usually not directly applicable or immediately useful.

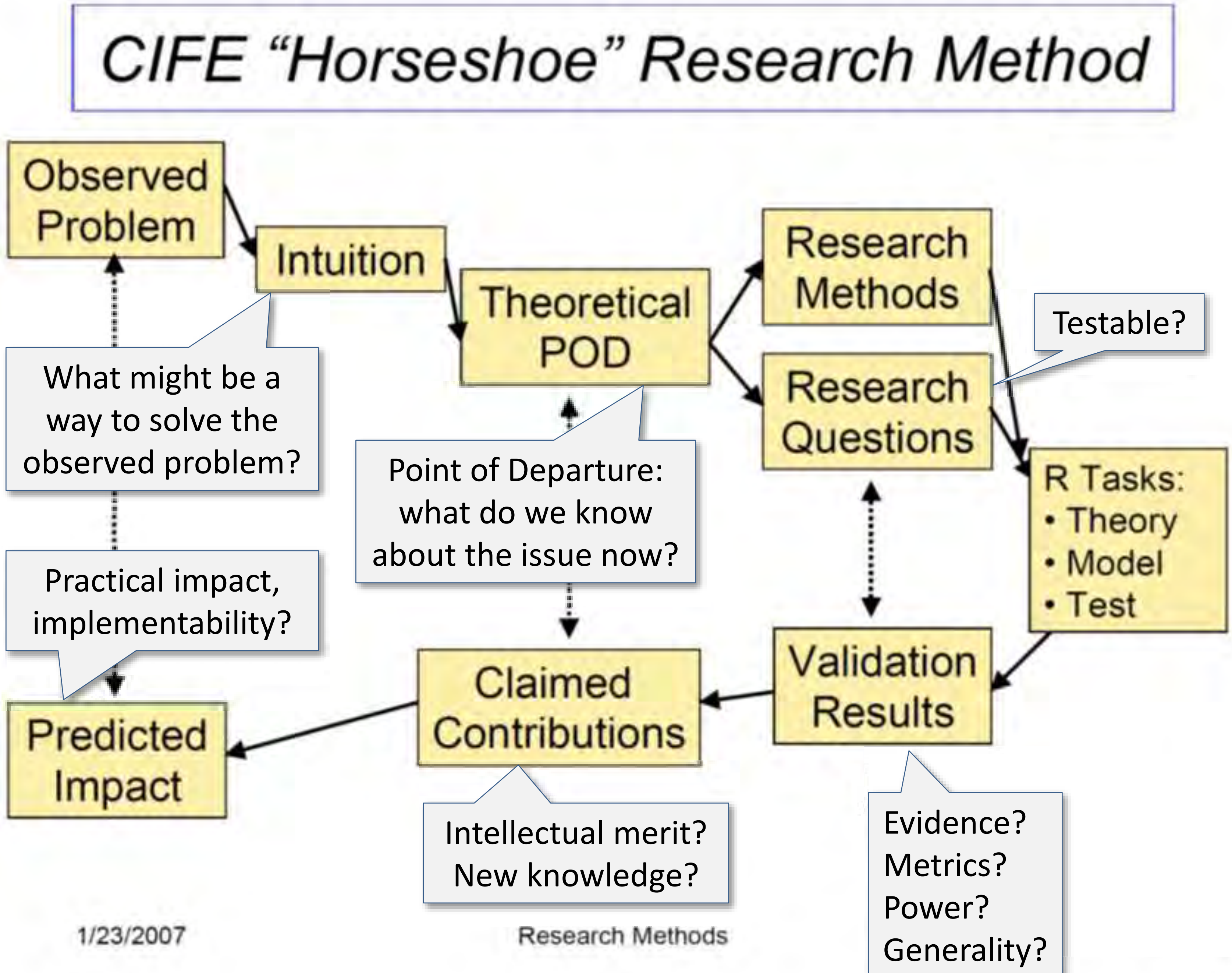
Basic vs. applied research

- **Applied research** seeks answers to questions in the real world and tries to solve problems
  - However, if we already have the necessary knowledge to solve the problem, it does not need research.
  - Problem solving is not an appropriate PhD topic, because PhD research must expand our current knowledge.
- Research related to the Built Environment is by definition always applied research.



CIFE "Horseshoe" Research Method





CIFE "Horseshoe" Research Method
Observed Problem
Intuition
Theoretical POD
Research Methods
Research Questions
Testable?
R Tasks:
• Theory
• Model
• Test
Validation Results
Claimed Contributions
Predicted Impact
What might be a way to solve the observed problem?
Point of Departure: what do we know about the issue now?
Practical impact, implementability?
Intellectual merit? New knowledge?
Evidence? Metrics? Power? Generality?
1/23/2007
Research Methods

Some observations of CIFE horseshoe

- Not really a research method, but a framework for applied research
- Actual research methods and tasks must be selected/ defined based on the research questions!
- Useful in helping to understand the research process and dependencies between different parts of it

Data collection –
interviews or questionnaires?

Interviews vs. questionnaires

- Interviews vs. questionnaires – depth vs. width?
  - How do you select the participants?
  - What do you ask?
    - Interviews are "self-correcting"; the discussion can lead to additional and more meaningful questions. Similar interaction is missing from questionnaires.
    - However, the questions should not imply the "expected" answer, e.g. you should not ask about "benefits" but "impacts"
  - Can you get meaningful number of answers to your questionnaire; statistical relevance?
    - In addition, the people who respond are usually the ones interested in the topic, e.g. not average industry people

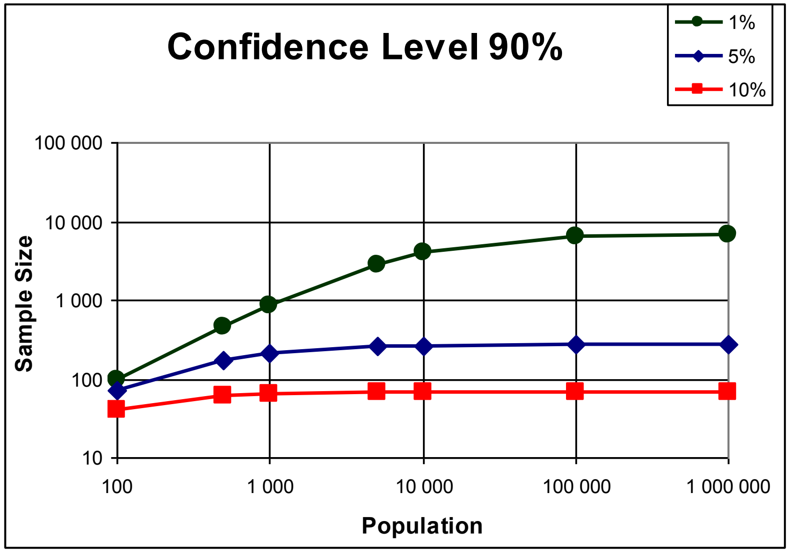
Challenges in questionnaires

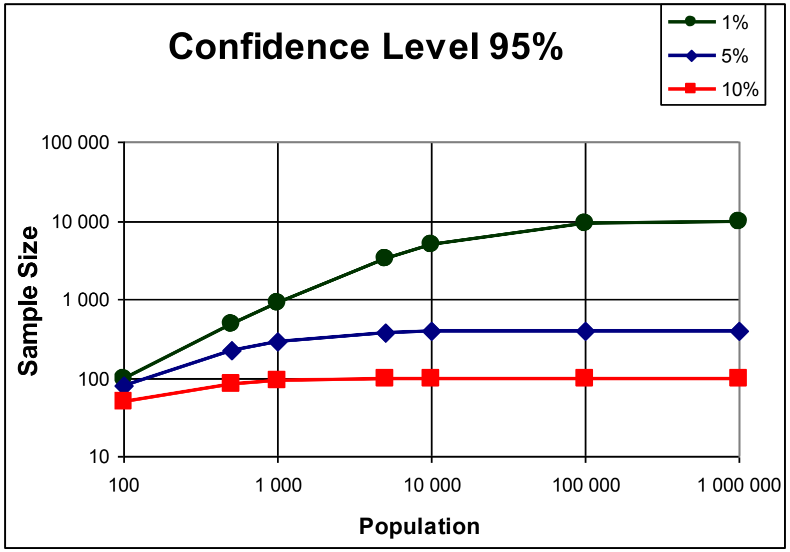
- Good questionnaires are really difficult to make:
  - Can you get the answers you need for your research?
  - Would a particular question yield the expected result?
    - "Did you participate in the design meeting last week?" vs. "How many times have you been to the design meetings during the project?"
  - Will everyone understand the questions in the same way? For example "Are you using BIM in your projects?"
    - What does "using BIM" mean; using Revit to produce drawings, sharing data, integrated project delivery, … ?
    - In how many projects you must use BIM to answer "yes"; 1, 10, all?
  - If you have multiple choices in answers, are **ALL** relevant options listed?
    - Very often the answers do not include some possible options

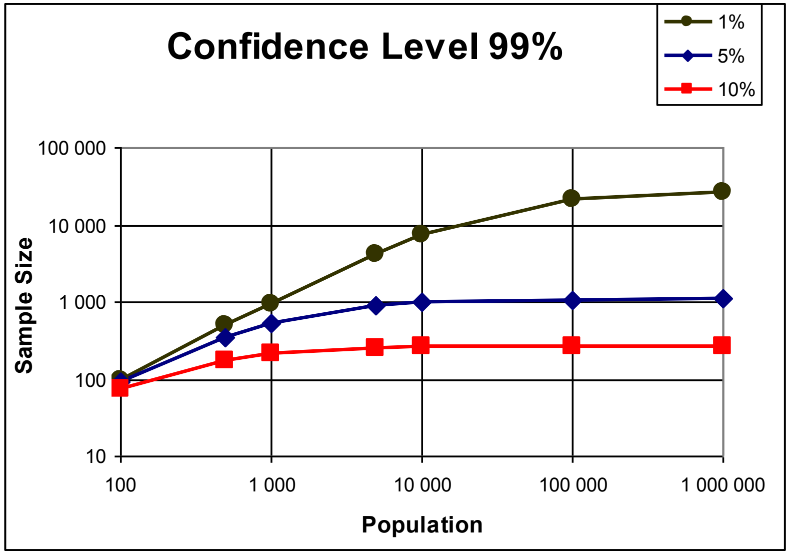
Sample size and randomisation

Confidence Level 99%

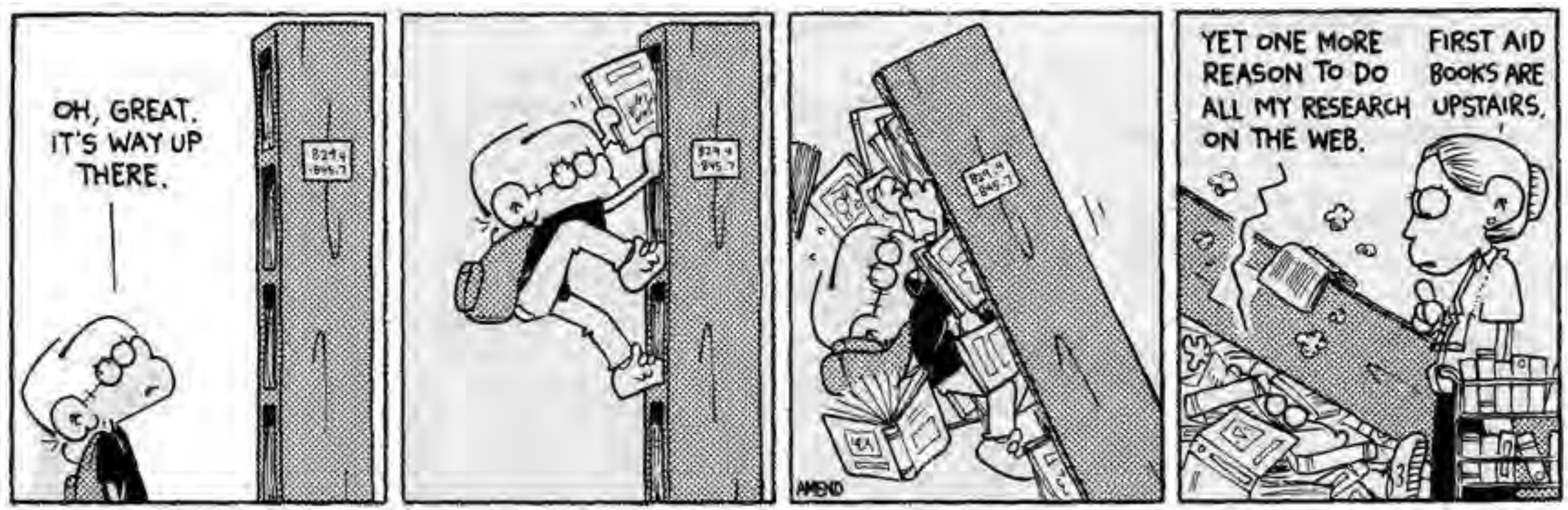
Population	Sample size: 10 % error	5 % error	1 % error
100	73	92	100
500	176	343	491
1 000	214	521	965
5 000	258	893	4223
10 000	265	980	7310
100 000	271	1075	21368
1 000 000	272	1086	26455







Sample size calculators



Some sample size calculators in the web *(checked 19/06/2018):*

- http://www.surveysystem.com/sscalc.htm
- https://www.surveymonkey.com/mp/sample-size-calculator/
- http://www.raosoft.com/samplesize.html
- http://www.nss.gov.au/nss/home.nsf/pages/Sample+size+calculator

Margin of Error and Confidence Levels

- Surveying has been likened to taste-testing soup – a few spoonfuls tell what the whole pot tastes like.
- The key to the validity of any survey is **randomness**.
  - Just as the soup must be stirred in order for the few spoonfuls to represent the whole pot, when sampling a population, the group "must be stirred" before respondents are selected.

Source: https://www.isixsigma.com/tools-templates/sampling-data/margin-error-and-confidence-levels-made-simple/

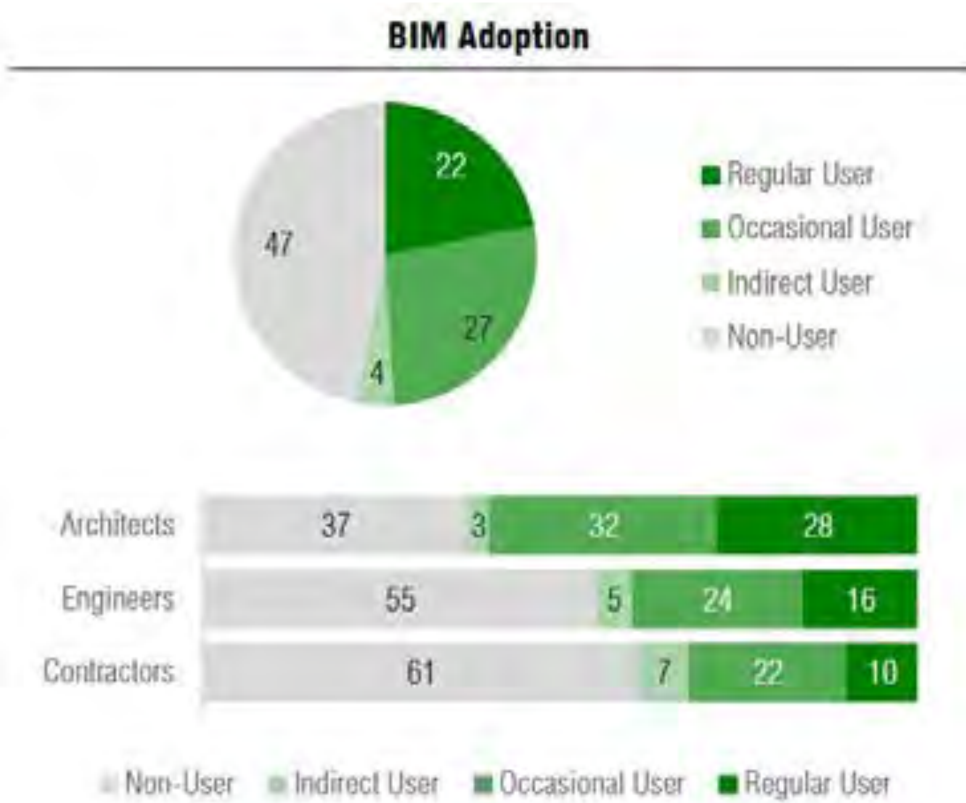
Margin of Error and Confidence Levels

- How well the sample represents the population is gauged by two important statistics – the survey's margin of error and confidence level.
  - For example, a survey may have a margin of error of ± 3 at a 95% level of confidence.
  - These terms simply mean that if the survey were conducted 100 times, the data would be within ± 3 percentage points above or below the percentage reported in 95 of the 100 surveys.
  - In other words, if 50% of the respondents say Company X's customer service is "very good." This means that if the survey were conducted 100 times, the percentage who say service is "very good" will range between 47% and 53% in 95 of the surveys.

Source: https://www.isixsigma.com/tools-templates/sampling-data/margin-error-and-confidence-levels-made-simple/

Also the response rate is crucial…

- Even if you have selected a valid, randomised questionnaire sample, a poor response rate can invalidate the results!



"One specific number is quite telling: out of 50,000 invitations sent, only 1,338 surveys were completed. This represents a measly 2.6% response rate..."

You cannot make **any** conclusions of the BIM adoption in Chile based on this research!

Source: 2016 National BIM Survey Chile

Why you cannot trust the Chilean research?

Calculate Your Sample Size:

Population Size: 50000

Confidence Level (%): 99

Margin of Error (%): 5

CALCULATE

SurveyMonkey

Sample Size:

657

50,000 invitations

1,338 responses

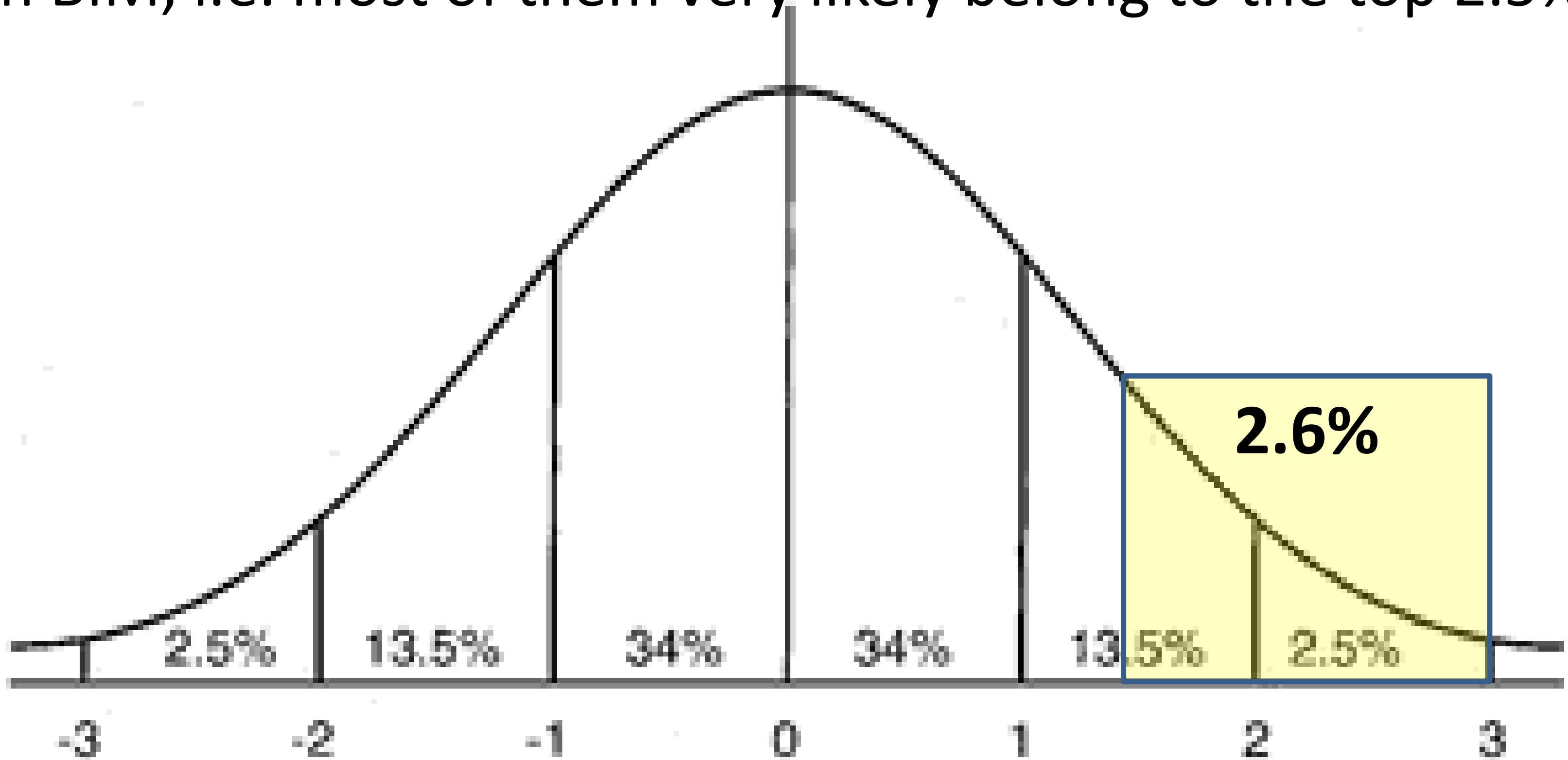
~ 2x

What is wrong?

Why you cannot trust the Chilean research?

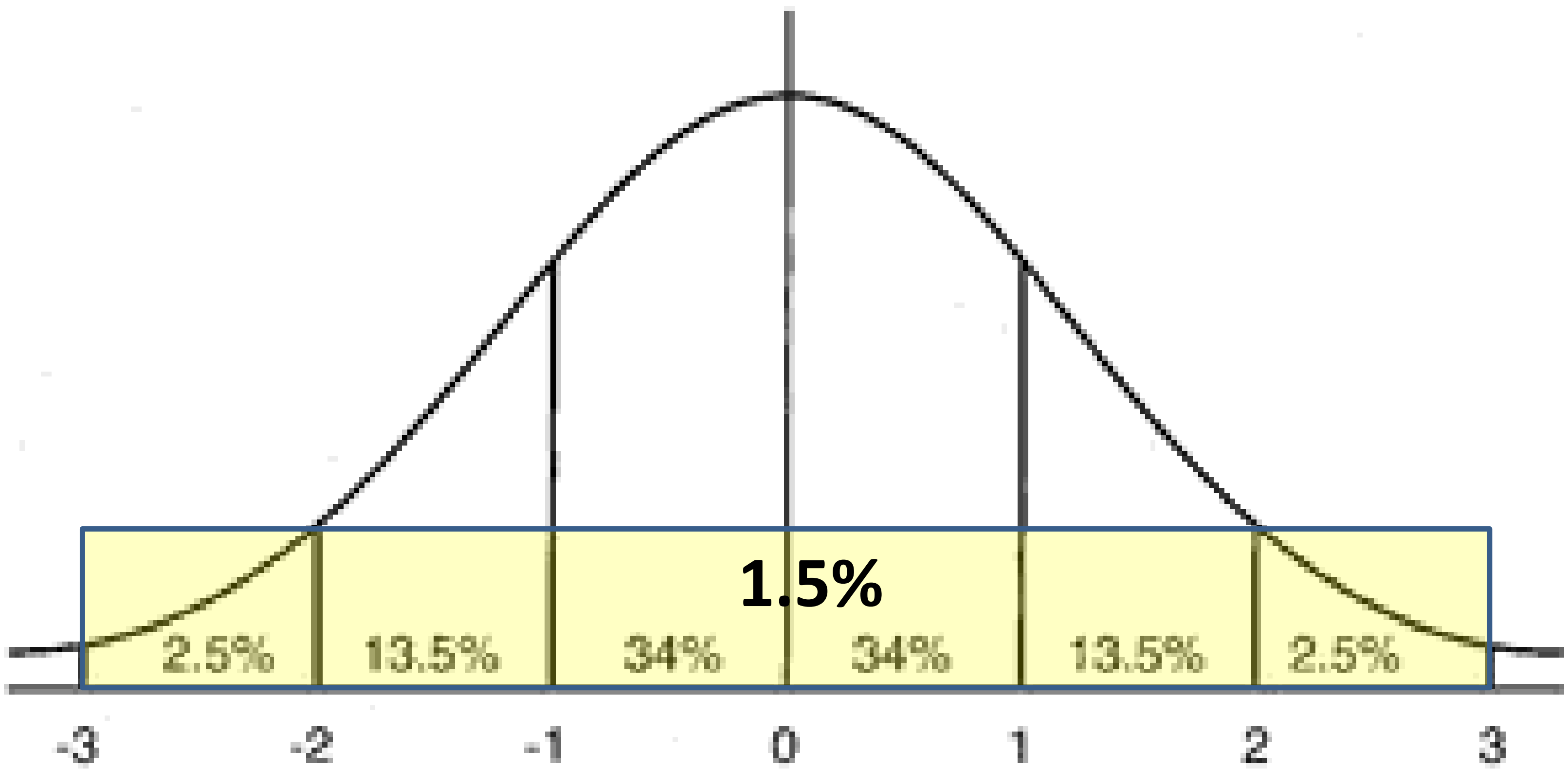
People tend to answer **only** if they are interested in topic!

We can **assume** that most of the respondents were interested in BIM, i.e. most of them very likely belong to the top 2.5%



Why you cannot trust the Chilean research?

If the population is 50,000 and we send 750 questionnaires to **randomised** people/companies, and almost all reply, then the results would be trustworthy!



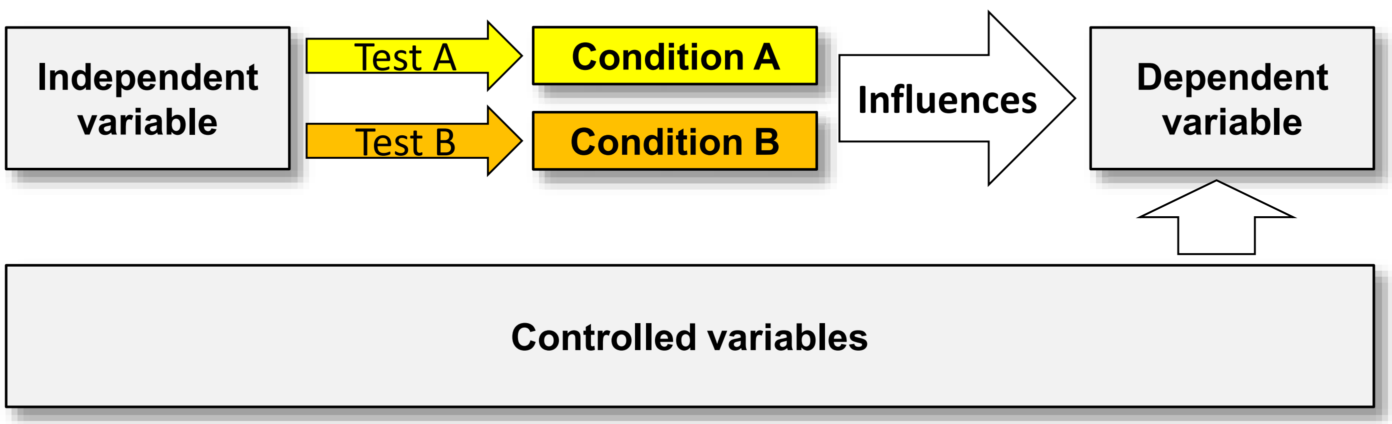
If you cannot find information of the population, sample size, selection method and response rate, do not trust to the survey results too much!

How about National BIM Surveys in the UK?

What is required for valid research arrangements?

What you need to measure the impact?

- **Controlled variables**: These are the things that are kept the same throughout your experiment/research.
- **Independent variable**: The **one** variable that you purposely change and test.
- **Dependent variable**: The measured change observed because of the independent variable. It is crucial to decide how you are going to measure the change.



Manufacturing industry is a controlled environment

Why it is impossible to measure benefits on project level?

- We have no controlled variables which could be kept the same
  - Each project is different
  - Each project team is different
  - Each site is different
  - Weather can be different…
- Even if you would rebuild the same building with the same team, the learning from the previous project would affect the results

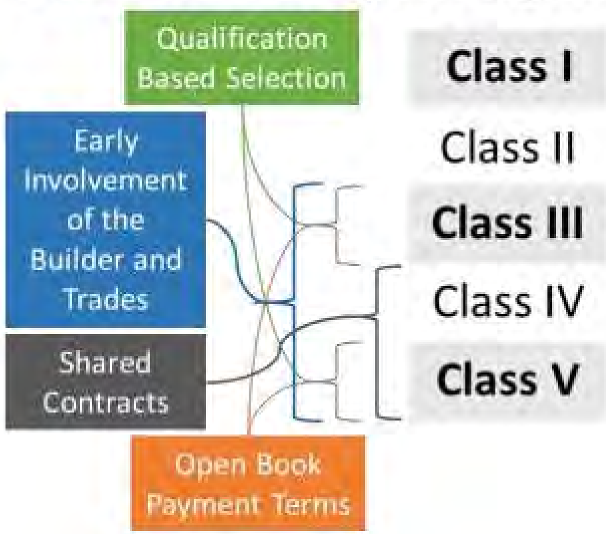
Impact of Delivery Strategy

204 projects

Public	127	(62%)
Private	77	(38%)

Completed 2008 - 2013

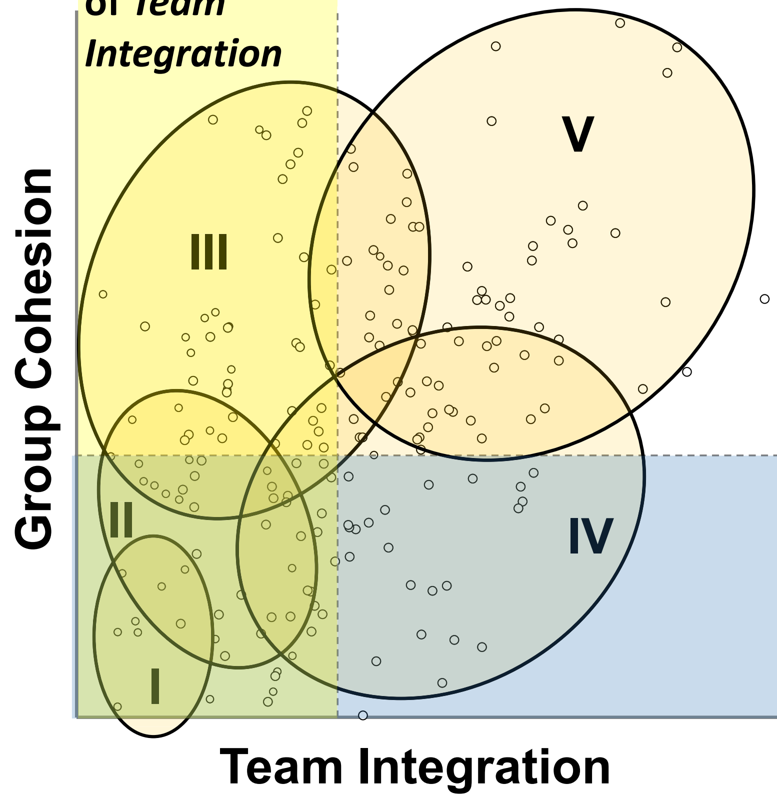
Different Delivery Strategies



70% of projects delivered late had below average levels of *Team Integration*

Group Cohesion:

- Reduced ***cost growth***
- Improved ***turnover experience***
- Higher ***system quality***



Team Integration:

- Reduced ***schedule growth***
- Enabled more ***intense schedules***
- Led to ***more group cohesion***

58% of over budget projects had below average levels of *Group Cohesion*

Courtesy Prof Robert M. Leicht, Penn State University, 2018

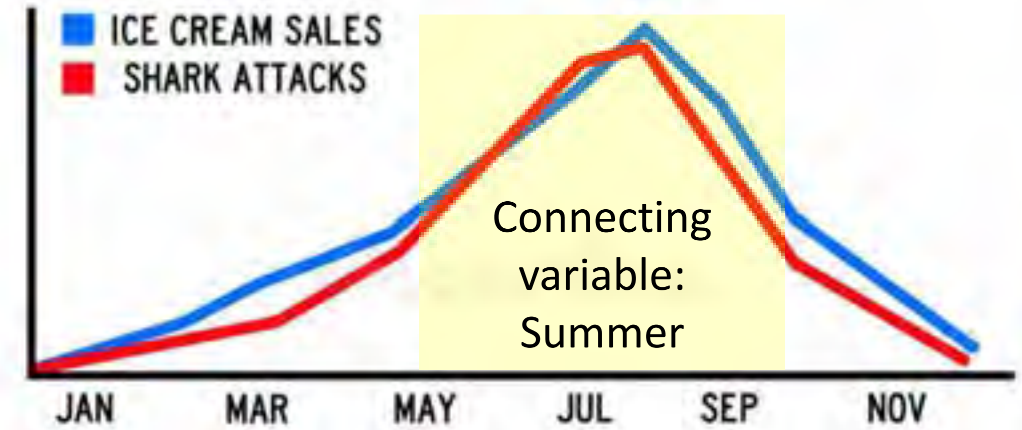
Internal and external validity

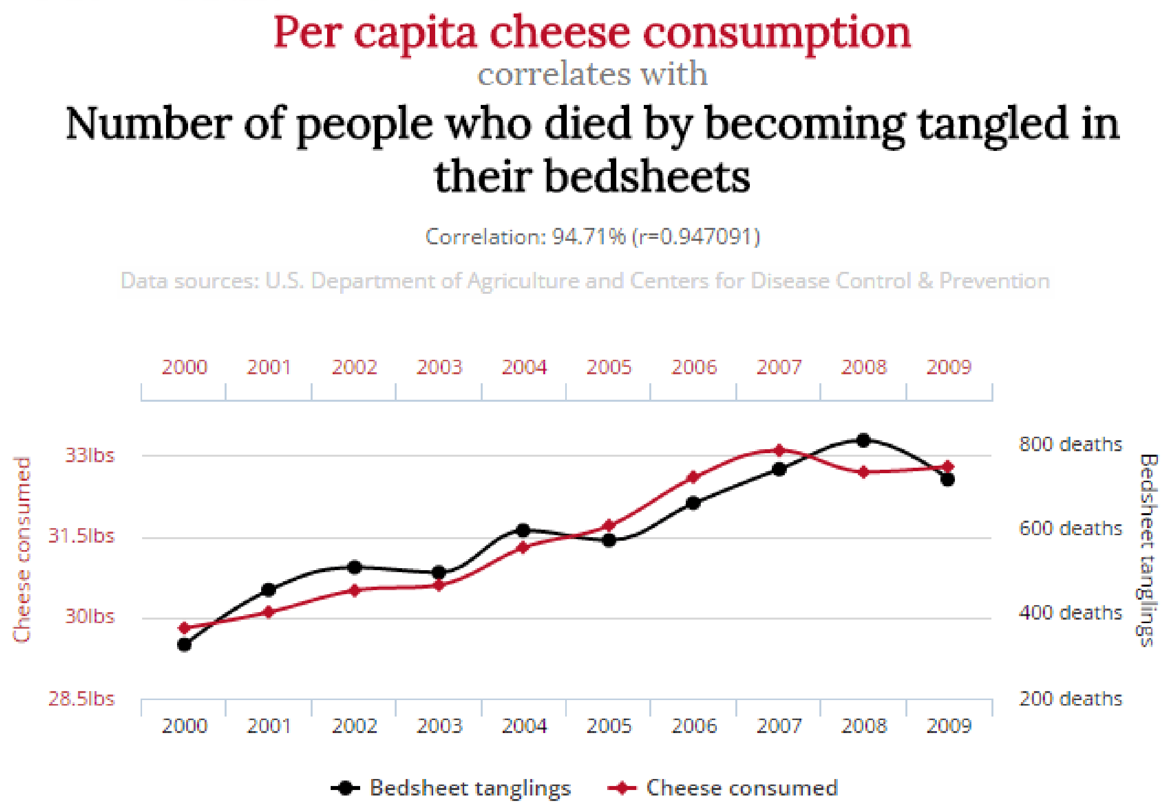
- **Internal Validity**
  - the degree to which the results are attributable to the independent variable and not some other rival explanation
- **External Validity**
  - the extent to which the results of a study can be generalised

Source:
https://pdfs.semanticscholar.org/presentation/7efd/428c7a81fc9839f17f01f28150b049a758b8.pdf

Interpreting and validating the results?

Correlation does not mean causation!





No connection, but almost 95% correlation!

Validation of the results can be challenging

- The test group should **not** consist of people who know what you try to prove – especially if they are your friends and want to support you
  - People have a strong tendency to reply as they think is expected. Sometimes this means that you must mislead the participants about the purpose of your research to get honest reactions. Deception can range from relatively minor omissions, such as not telling people the full story, to outright falsehood about the nature of the research.
- The results can be highly dependent on the skills in the test group
  - For example, you will get very different results if you test the usability of a software with software developers, savvy users or unexperienced users…

Experimental and Quasi-Experimental Designs for Research

Donald T. Campbell and Julian C. Stanley 1963

https://wagner.nyu.edu/files/doctoral/Campbell_and_Stanley_Chapter_5.pdf

Focus area

- Application area - relevance to us?
- Pre-, true and quasi-experimental designs
- Internal and external validity
  - Sources of invalidity in different designs
- Different experimental designs
  - 3 Pre-Experimental Designs
  - 3 True Experimental Designs
  - 10 Quasi-Experimental Designs

Relevance for us?

- Examples in the book are from the research of education and learning
  - **Experiment designs:**
    - Randomisation, treatment, observation
  - **Internal validity:**
    - History, maturation, instrumentation, regression, selection, mortality, interaction of selection, maturation, etc.
  - **External validity** (X = Treatment)**:**
    - Interaction of testing and X, interaction of selecting and X, reactive arrangements, multiple-X interference
- How does this relate to our research?

Difference / when to use ?

- True experiment
  - Full control of all aspects
  - Should be used when ever possible
- Quasi-experiment
  - Control over data collection procedures
    - the *when* and *to whom of* measurement
  - Lack of full control over scheduling of experimental stimuli
    - the *when* and *to whom* of exposure
    - the ability to randomise exposures

Which is better?

- Different research questions and situations demand different designs
  - Full control is not always possible
  - Randomisation is not always possible
- Each design has its risks *(and most have also some benefits)*
  - Important to understand
    - how to analyse
    - which are the risks

Internal validity risks

1. **History**
2. Maturation
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

1. The events between the measurements in addition to the experimental variable

Internal validity risks

1. History
2. **Maturation**
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

2. Process within the respondents happening because of the time per se (growing older, more experienced, more tired, etc.)

Internal validity risks

1. History
2. Maturation
3. **Testing**
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

3. The effects of previous tests to the following tests

Internal validity risks

1. History
2. Maturation
3. Testing
4. **Instrumentation**
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

4. Changes in the measurements caused by the changes in the measuring instruments, observers or scorers

Internal validity risks

1. History
2. Maturation
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

5. Operating where groups have been selected on the basis of their extreme scores

Internal validity risks

1. History
2. Maturation
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

6. Biases because of the differential selection of respondents in comparison groups

Internal validity risks

1. History
2. Maturation
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

7. Differential loss of respondents in comparison groups

Internal validity risks

1. History
2. Maturation
3. Testing
4. Instrumentation
5. Regression
6. Selection
7. Mortality
8. Interaction of selection, maturation, etc.

8. Effects of interaction between the different factors can be interpreted as the effect of the experimental variable - *high risk in certain multiple-group quasi-experiments!*

External validity risks

9. Interaction of Testing and X
10. Interaction of Selection and X
11. Reactive Arrangements
12. Multiple -X Interference

9. Pretest might affect to the respondent's sensitivity to the experimental variable

External validity risks

9. Interaction of Testing and X
10. Interaction of Selection and X
11. Reactive Arrangements
12. Multiple -X Interference

10. The interaction between the selection biases and the experimental variable

External validity risks

9. Interaction of Testing and X
10. Interaction of Selection and X
11. Reactive Arrangements
12. Multiple -X Interference

11. Exposure to the experimental variable in non-experimental settings prevents generalisation about its effect

External validity risks

9. Interaction of Testing and X
10. Interaction of Selection and X
11. Reactive Arrangements
12. Multiple -X Interference

12. Effect of prior treatments; likely to occur whenever multiple treatments are applied to the same respondents

Pre- and True Experimental Designs

+ Controlled factor
- Definite weakness
? Possible source of concern
Not relevant

Design	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
Pre-Experimental Designs												
1. One-Shot Case Study X O	-	-				-	-			-		
2. One-Group Pretest-Posttest Design O_1 X O_2	-	-	-	-	?	+	+	-	-	-	?	
3. Static Group Comparison X O_1 .... O_2	+	?	+	+	+	-	-	-		-		
True Experimental Designs												
4. Pretest-Posttest Control Group Design R O_1 X O_2 R O_3 O_4	+	+	+	+	+	+	+	+	-	?	?	
5. Solomon Four-Group Design R O_1 X O_2 R O_3 O_4 R X O_5 R O_6	+	+	+	+	+	+	+	+	+	?	?	
6. Posttest-Only Control Group Design R X O_1 R O_2	+	+	+	+	+	+	+	+	+	?	?	

Pre- and True Experimental Designs

- **+** Controlled factor
- **-** Definite weakness
- **?** Possible source of concern
- Not relevant

	Internal: History	Internal: Maturatio	Internal: Testing	Internal: Instrumen	Internal: Regressio	Internal: Selection	Internal: Mortality	Internal: Interactio maturatio	External: Interactio and X	External: Interactio and X	External: Reactive Arrangem	External: Multiple – Interference
Pre-Experimental Designs												
1. One-Shot Case Study X O	-	-				-	-			-		
2. One-Group Pretest-Posttest Design O_1 X O_2												
3. Static Group Comparison X O_1 O_2												
True Experimental Designs												
4. Pretest-Posttest Control Group Desig R O_1 X O_2 R O_3 O_4												
5. Solomon Four-Group Design R O_1 X O_2 R O_3 O_4 R X O_5 R O_6												
6. Posttest-Only Control Group Design R X O_1 R O_2	+	+	+	+	+	+	+	+	+	?	?	

Sources of Invalidity

X	= Treatment
X_s	= Treatments in temporal order
O	= Observation
O_S	= Observations in temporal order
R	= Random assignment
- - - -	= Comparison groups **not** equated by randomisation
M_s	= Material (only Design 9)

One-Shot Case Study

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Pre-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
1. One-Shot Case Study X O	-	-				-	-			-		

What is positive and/or negative in this design?
Any specific risks?

- Total absence of control ⇨ no scientific value
- Valid comparisons not possible
- Most weaknesses of all designs
- Often used - but should not be!

One-Group Pretest-Posttest Design

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
Pre-Experimental Designs												
2. One-Group Pretest-Posttest Design O_1 X O_2	-	-	-	-	?	+	+	-	-	-	?	

What is positive and/or negative in this design?
Any specific risks?

- Can be used where ”nothing better can be done”
- Many possible rival explanation for change: history, maturation, pretest, instrumentation, and interaction of
- Statistical regression also possible source of invalidity
- Why ”-” in the Interaction column?

Static Group Comparison

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Pre-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
3. Static Group Comparison X O_1 .... O_2	+	?	+	+	+	-	-	-		-		

What is positive and/or negative in this design?
Any specific risks?
Examples of possible usage?

- No randomisation; differences in the groups can cause the difference in the observations
- Mortality (drop-out persons) in the groups can be different
- Possible ”guinea-pig” effect in those who have the treatment

Pretest-Posttest Control Group Design

- Randomised groups; equal (if correctly selected)
- Good control of internal validity; history, maturation, testing, regression, selection
- However, mortality can cause differences if it is not carefully considered
- Pretest can effect to the results
- Can we totally rule out the ”guinea-pig” effect as a potential source of differences? Not really discussed in the book...
- Can O3 and O4 effect differently than O1 and O2? Two tests without any treatment in between could make people frustrated.

4. Pretest-Posttest Control Group Design R O_1 X O_2 R O_3 O_4	+	+	+	+	+	+	+	+	-	?	?

What is positive and/or negative in this design?
Any specific risks?

Solomon Four-Group Design

+ Controlled factor
- Definite weakness
? Possible source of concern
Not relevant

- Otherwise the same as Design 4, but two additional groups will check the possible effects of the pretest
- Improved external validity

Design	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
4. Pretest-Posttest Control Group Design R O_1 X O_2 R O_3 O_4	+	+	+	+	+	+	+	+	-	?	?	
5. Solomon Four-Group Design R O_1 X O_2 R O_3 O_4 R X O_5 R O_6	+	+	+	+	+	+	+	+	+	?	?	

What is the difference and benefits compared to Design 4?

Posttest-Only Control Group Design

		History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
+	**Controlled factor**												
-	**Definite weakness**												
?	**Possible source of concern**												
	Not relevant												
6. Posttest-Only Control Group Design R X O_1 R O_2		+	+	+	+	+	+	+	+	+	?	?	

- **Statistical tests available for Design 4 are more powerful**
- **However:**
  - **pretests effects or difficulties to arrange it can justify the use of this design**
  - **this fits better for new issues, where pretest ‘would be ”awkward”’**

Obviously simpler than Design 4 and better external validity.
So are there any reasons to use Design 4 instead of this?

Quasi-Experimental Designs 1/4

+ Controlled factor
- Definite weakness
? Possible source of concern
Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
7. Time Series O O O O X O O O O	-	+	+	?	+	+	+	+	-	?	?	
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O etc	+	+	+	+	+	+	+	+	-	?	-	-
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O etc	+	+	+	+	+	+	+	-	-	?	?	-
10. Nonequivalent Control Group Design O X O .... O O	+	+	+	+	?	+	+	-	-	?	?	
11. Counterbalanced Designs X_1O X_2O X_3O X_4O .... X_2O X_4O X_1O X_3O .... X_3O X_1O X_4O X_2O .... X_4O X_3O X_2O X_1O	+	+	+	+	+	+	+	?	?	?	?	-

Time Series

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
7. Time Series O O O O X O O O O	-	+	+	?	+	+	+	+	-	?	?	

What is positive and/or negative in this design?
Any specific risks?

- Useful if experimental variable is transient or reversibleMultiple X interference is a special hazard for external validity
- Best use for physiological research; repeated stimulus for one animal
- Also possible for observation of an external factor; like the effect of parent-observers in the classroom behavior

Equivalent Time Samples Design

- **+** Controlled factor
- **-** Definite weakness
- **?** Possible source of concern
- Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
8. Equivalent Time Samples Design X_1O X_0O X_1O X_0O etc	+	+	+	+	+	+	+	+	-	?	-	-

What is positive and/or negative in this design?
Any specific risks?

- Useful if experimental variable is transient or reversible
- Multiple X interference is a special hazard for external validity
- Best use for physiological research; repeated stimulus for one animal
- Also possible for observation of an external factor; like the effect of parent-observers in the classroom behavior

Equivalent Materials Samples Design

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
9. Equivalent Materials Samples Design M_aX_1O M_bX_0O M_cX_1O M_dX_0O etc	+	+	+	+	+	+	+	-	-	?	?	-

What is positive and/or negative in this design?
Any specific risks?

- Multiple X interference is a special hazard for external validity
- Internal validity similar as in Design 8
- Reactive arrangements less involved than in Design 8

Nonequivalent Control Group Design

- **+** Controlled factor
- **-** Definite weakness
- **?** Possible source of concern
- Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
10. Nonequivalent Control Group Design O X O .... O O	+	+	+	+	?	+	+	-	-	?	?	
4. Pretest-Posttest Control Group Design R O_1 X O_2 R O_3 O_4	+	+	+	+	+	+	+	+	-	?	?	

What is positive and/or negative in this design?
Any specific risks?

- Same as Design 4 (Pretest-Posttest Control Group) but without randomisation -
- However, results less uncertain than in Design 2 (One-Group Pretest-Posttest)

Counterbalanced Design

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
11. Counterbalanced Designs X_1O X_2O X_3O X_4O X_2O X_4O X_1O X_3O X_3O X_1O X_4O X_2O X_4O X_3O X_2O X_1O	+	+	+	+	+	+	+	?	?	?	?	-

What is positive and/or negative in this design?
Any specific risks?

- All groups get all treatments but in different order; results can be compared by Xs in columns or rows
- Possible to turn into true experiment by sufficient replication

- Multiple X interference is a special hazard for external validity

Separate-Sample Pretest-Posttest Designs

- **+** Controlled factor
- **-** Definite weakness
- **?** Possible source of concern
- Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
12. Separate-Sample Pretest-Posttest Design R O (X) R X O	-	-	+	?	+	+	-	-	+	+	+	
12a Separate-Sample Pretest-Posttest Design R O (X) R X O R O (X) R X O	+	-	+	?	+	+	-	+	+	+	+	
12b Separate-Sample Pretest-Posttest Design R O_1 (X) R O_2 (X) R X O_3	-	+	+	?	+	+	-	?	+	+	+	
12c Separate-Sample Pretest-Posttest Design R O_1 X O_2 R X O_3	-	-	+	?	+	+	+	-	+	+	+	

What are the main differences in these designs?

Separate-Sample Pretest-Posttest Control Group Designs

+ Controlled factor
- Definite weakness
? Possible source of concern
Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
13. Separate-Sample Pretest-Posttest Control Group Design R O (X) R X O .... R O R O	+	+	+	+	+	+	+	-	+	+	+	

Compare 13 to 12a:
What is the different; why is the internal validity better?
(except Interaction of selection, maturation, etc.)

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
12a Separate-Sample Pretest-Posttest Design R O (X) R X O .... R O (X) R X O	+	-	+	?	+	+	-	+	+	+	+	

Separate-Sample Pretest-Posttest Control Group Designs

Obviously excellent, but expensive design.

According to the book never used (I don't know if still true?)

13a Separate-Sample Pretest-Posttest Control Group Design

```
      ┌ R   O   (X)
      │ R        X   O
      │ ................
      │ R   O   (X)
R    <  R        X   O
      │ ................
      │ R   O   (X)
      └ R        X   O
      ┌ R   O
      │ R            O
      │ ................
R    <  R   O
      │ R            O
      │ ................
      │ R   O
      └ R            O
```

+	+	+	+	+	+	+	+	+	+	+

Quasi-Experimental Designs 4/4

+	Controlled factor
-	Definite weakness
?	Possible source of concern
	Not relevant

Quasi-Experimental Designs

14. Multiple Time-Series

O O O O X O O O O

O O O O O O O O

15. Institutional Cycle Design

Class A	X	O_1		
Class B_1		RO_2 X	O_3	
Class B_2		R X	O_4	
Class C			O_5	X

General Population Controls for Class B O_6

General Population Controls for Class C O_7

16. Regression Discontinuity

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
14. Multiple Time-Series	+	+	+	+	+	+	+	+	-	-	?	
15. Institutional Cycle Design: $O_1 > O_2$, $O_4 > O_5$	+	-	+	+	?	-	?		+	?	+	
$O_2 > O_3$	-	-	-	?	?	+	+		-	?	+	
$O_2 > O_4$	-	-	+	?	?	+	?		+	?	?	
$O_6 = O_7$, $O_{2y} = O_{2o}$		+						-				
16. Regression Discontinuity	+	+	+	?	+	+	?	+	+	-	+	+

Multiple Time-Series

+	**Controlled factor**
-	**Definite weakness**
?	**Possible source of concern**
	Not relevant

Quasi-Experimental Designs	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
14. Multiple Time-Series O O O O X O O O O O O O O O O O O	+	+	+	+	+	+	+	+	-	-	?	

Compare 14 to 7 and 10:
What is the different; why is the validity better?

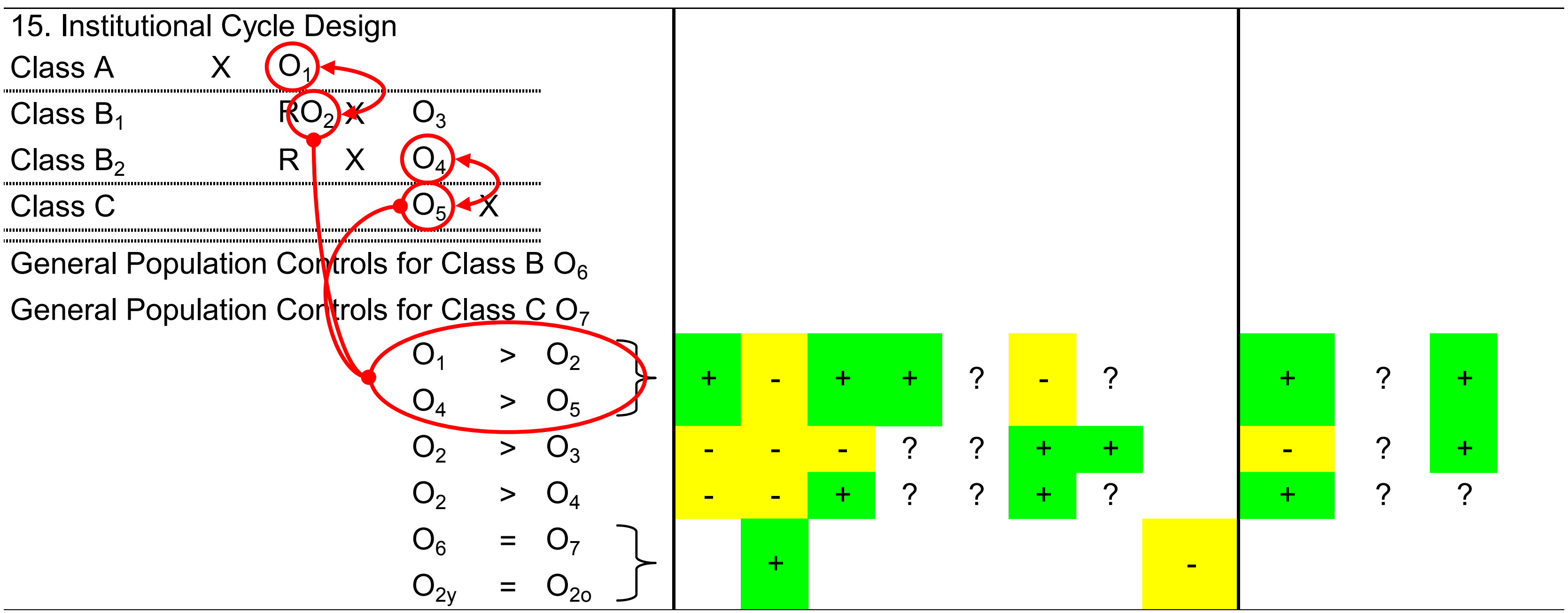
Design	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of selection, maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements	Multiple –X Interference
7. Time Series O O O O X O O O O	-	+	+	?	+	+	+	+	-	?	?	
10. Nonequivalent Control Group Design O X O .... O O	+	+	+	+	?	+	+	-	-	?	?	

- **“In general, this is an excellent quasi-experimental design, perhaps the best of the more feasible designs.”**

Institutional Cycle Design

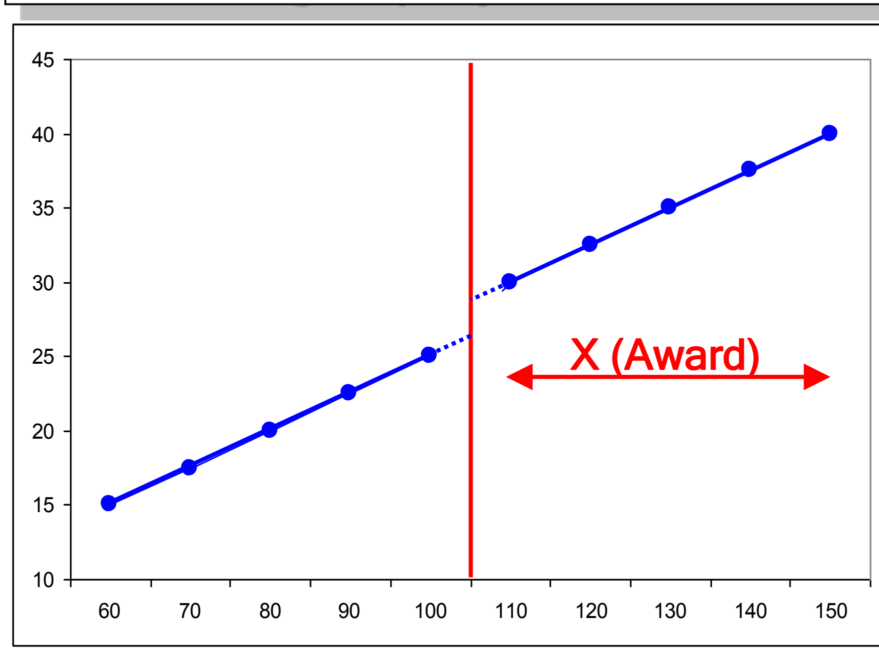
”Patched-up” design; strategy for field research which starts with inadequate design and then adds specific features to better control; example of the Air Force cadet training research.

What does ”O1 > O2” & ”O4 > O5” mean?



Regression Discontinuity

- Limited set of applications; mainly for educational settings
- Example of studying expected relation of pre-award ability to later achievement:
    - Two groups; just below and above the cutting point



- Correlation does not necessarily mean causation
- However, relatively inexpensive correlational approach can provide preliminary survey for hypothesis to identify the potential cases for more expensive studies

16. Regression Discontinuity	+	+	+	?	+	+	?	+	+	-	+	+