Summer School BDA MDD 2026

How to conduct scientific experiments to enhance your research outputs

nicolas.travers@devinci.fr

DE VINCI RESEARCH CENTER

https://devinci.link/mdd26

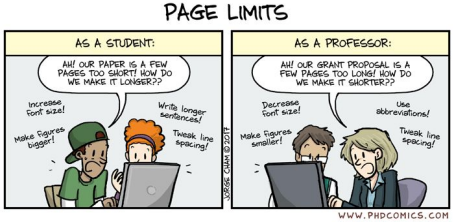
Research and Experiments

How to perform research?

Get "empirical" results

How to write research?

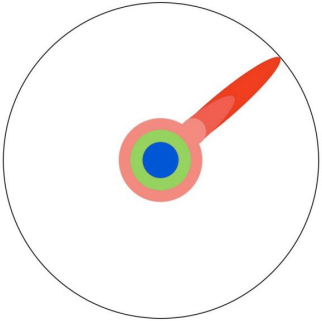
Get papers accepted



Sources and Thanks

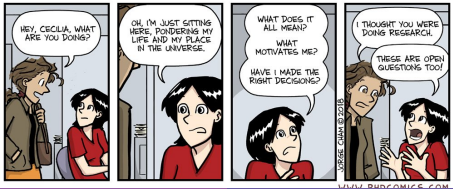
- *"Research and Methods in Computer Science"*
  S. Demeyer, University of Antwerp, 2018
- *"Case Study Research: Design and Methods"*
  R. K. Yin., 3rd Edition. SAGE Publications. California, 2009.
- *"Guidelines for Conducting and Reporting Case Study Research in Software Engineering"*
  P. Runeson, M. Höst. Empirical Softw. Eng. 14(2), 2009, 131-164.
- *"Steps in Conducting a Research Project or Experiment"*
  H. Zaleski, Experimental Design and Data Analysis Workshop, 2003.
- http://PhDcomics.com

What is Research?

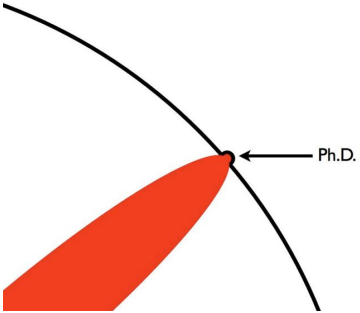


- Human knowledge
- Elementary School
- High School
- Bachelor
- Master
- Ph.D (∼first year)
- Ph.D (∼defense)

http://gizmodo.com/5613794/what-is-exactly-a-doctorate

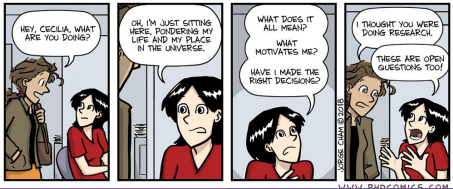


What is Research?

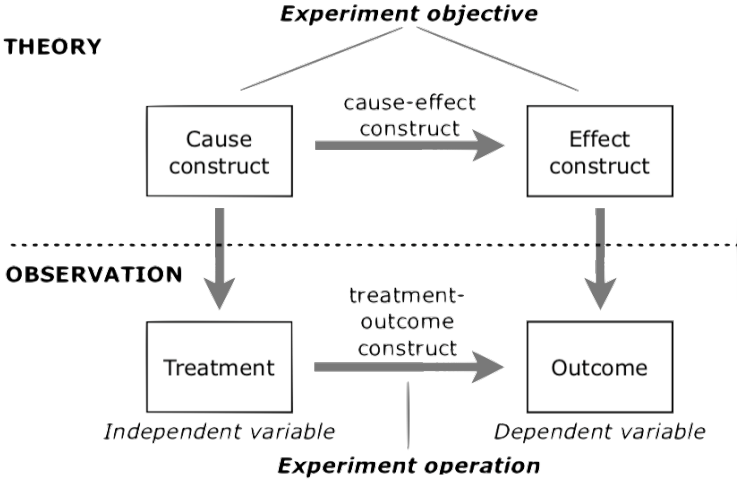


http://gizmodo.com/5613794/what-is-exactly-a-doctorate

- Human knowledge
- Elementary School
- High School
- Bachelor
- Master
- Ph.D (∼first year)
- Ph.D (∼defense)



Experiments Principle



C. Wohlin, et al. "*Experimentation in Software Engineering - An Introduction*", Kluwer Academic Publishers, 2000

Case studies

- Widely used in Computer Science
- ⚠"studying a case" vs "doing a case study" ⚠
- **Case studies** typology:

`Feasibility study`	*Is it possible?*
`Pilot case/Demonstrator`	*Is it appropriate?*
`Comparative study`	*Is it better?*
`Observational study`	*What is it?*
`Literature survey`	*What is known/unknown?*
`Formal model`	*Underlying concepts?*
`Simulation`	*What if?*

Feasibility Study

Here is a new idea, is it possible?

Ex: "CAP Theorem"

- Is it possible to solve a specific kind of problem... effectively?
  - Computer science perspective (P = NP, Turing test, etc.)
  - Engineering perspective (build efficiently; fast & small)
  - Economic perspective (cost effective & profitable)
- Is the technique new / novel / innovative?
  - Compare vs alternatives (*literature survey; comparative study*)
- Proof by construction
  - Build a prototype
  - Often by applying on a "`CASE`"
- Conclusions
  - Primarily qualitative: "lessons learned"
  - Quantitative: economic, engineering

Feasibility Study: Example

Three main properties manage database:

1. **C**onsistency
2. **A**vailability
3. **P**artition tolerance

CAP Theorem: Any database CAN ONLY HOLD two of the properties

Proved by construction[1].

[1] Eric A. Brewer, "*Towards robust distributed systems*", PODC, p.7, 2000

Pilot Case/Demonstrator

Here is an idea that has proven valuable; does it work for us?

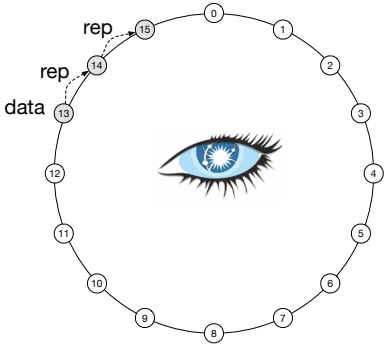
Ex: Apply consistency on Cassandra or MongoDB

- Proven valuable
  - Accepted merits (e.g. "lessons learned" from *feasibility study*)
  - Some implicit theory explains why the idea has merit
- Does it work for us?
  - Context is very important
- Demonstrated on a simple yet representative "`CASE`"
  - "*Pilot Case*" $\neq$ "*Pilot Study*"
- Proof by construction + prototyping + applied on a "`case`"
- Conclusions
  - Primarily qualitative; "lessons learned"
  - Quantitative; preferably with predefined criteria

  $\Rightarrow$ compare to context before applying the idea

Pilot Case/Demonstrator: Example

How can we handle both AVAILABILITY and CONSISTENCY on the distributed database *Cassandra*?



Consistency Level

ONE
TWO
THREE
QUORUM
ALL

Availability

Consistency

Conclusion: **CAP** Theorem → **PACELC** Theorem[2]

[2] D. Abadi, "Consistency Tradeoffs in Modern Distributed Database System Design: CAP is Only Part of the Story", Computer Journal 45(2), 2012, pp. 37–42

Comparative Study

Here are four/five techniques, which one is better?

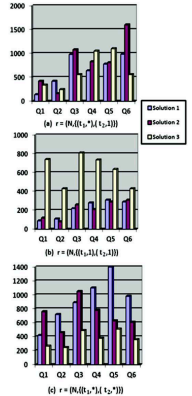
Ex: How denormalizing data models impacts NoSQL databases on a `case`?

- For a given purpose!
  - Not necessarily absolute ranking
- Where are the differences? What are the tradeoffs?
- Criteria check-list
  - Should not favor one technique
  - Qualitative (remain **unbiased**!) and Quantitative (what do you want to know?)
  - Criteria check-list should be complete and reusable!
    - If done well → most important contribution (*replication*!)
    - See `literature survey`
- Compare
  - Score criteria check-list (apply on a "`CASE`")

Comparative Study: Example

Benchmarking[3] NoSQL databases: Cassandra & MongoDB

NoSQL system	Relationship	Solution	Query	Time (s)
Cassandra	r = (N, {(t_1, ∗), (t_2, 1)})	Solution 1	Q1	140
			Q4	980
			...	
		Solution 2	Q1	830
			...	
		Solution 3	...	
			Q4	420
	r = (N, {(t_1, 1), (t_2, 1)})	Solution 1	...	
			Q5	310
		Solution 2	...	
			Q6	290
		Solution 3	...	
			Q4	735
	r = (N, {(t_1, ∗), (t_2, ∗)})	Solution 1	...	
			Q2	720
			Q5	1400
		Solution 2	...	
			Q3	1050
		Solution 3	...	
			Q5	510
			Q6	530
MongoDB	r = (N, {(t_1, ∗), (t_2, 1)})	Solution 1	...	
			Q4	4300
		Solution 2	...	
			Q6	6200
		Solution 3	Q5	1700
			Q6	1500
		Solution 4	Q2	870



Performance of data models wrt. NoSQL solutions

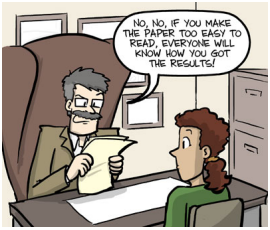
[3] F. Abdelhedi et al., "MDA-Based Approach for NoSQL Databases Modelling", Big Data Analytics and Knowledge Discovery, 2017

Observational Study

Understand phenomena through observations

Ex: Tourists' behavior on `Tripadvisor`

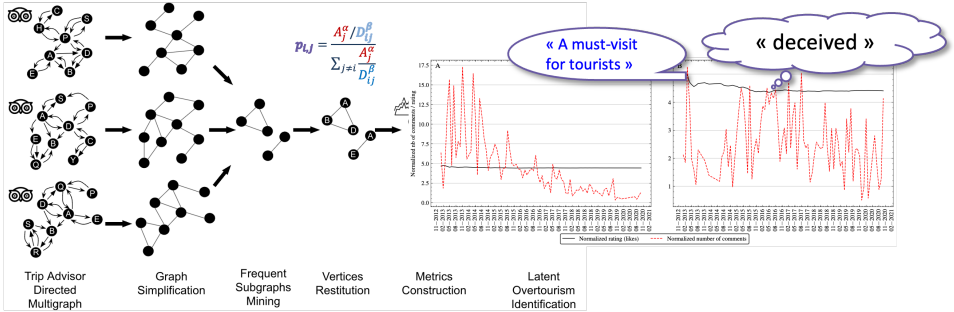
- Systematic collection of data derived from direct observation of the everyday life
- Phenomena is best understood in the fullest possible context
  - Observation & participation
  - Interviews & questionnaires
- Observing a `series of cases` "`CASE`"
- Conclusions
  - Primarily qualitative: classifications/observations/...



Observational Study: Example

Graph data model on Tourists' circulation: Propagation analysis

- Graph data model, aggregation operations, pattern mining, attractiveness model[4]



Latent Overtourism at Tampere, Finland

[4] H. Alatrista-Salas, G. Chareyron, S. Djebali, I. Ouled-Dlala, N. Travers, "From Flows to Graphs: Data-Driven Insights on Latent Overtourism with Frequent Pattern Mining", ADBIS, 2025

Literature Survey

What is known? What questions are still open?

Ex: How do recommendations systems work?

- Comprehensive: **Precise research question**[5]
- Define:
  - A search strategy: rigor, completeness, replication
  - A scope: criteria for inclusion and exclusion
- Specify information to be obtained
- The "CASES" = selected papers
- Organize conclusions:

classification	taxonomy	conceptual model
table	tree	frequency

[5] B. A. Kitchenham, "Procedures for Performing Systematic Reviews", Keele University Technical Report EBSE-2007-01, 2007

Literature Survey: Example

Survey on self-tuning database systems[6]:

((schema design OR physical design OR schema configuration OR schema self-tuning OR index OR materialized view OR partitioning OR clustering) AND (autonomic database OR relational database OR data warehouse OR OLAP))
OR
((modeling OR schema OR design OR tuning OR Model) AND (key-value OR document OR column OR graph) AND (NoSQL OR database OR data warehouse OR OLAP))

Survey Methodology

Database Type	Publication Type	Primary Studies	Years of Pub
SQL Databases	Journal	40	1974–2021
	Conference	64	1976–2021
	Book	3	1983–2020
	Dissertation or technical report	7	1970–2020
	Total	114	1970–2021
NoSQL Databases	Journal	21	2014–2021
	Conference	49	2010–2021
	Book	1	2019
	Dissertation or technical report	1	2018
	Total	72	2010–2021

Papers Analysis

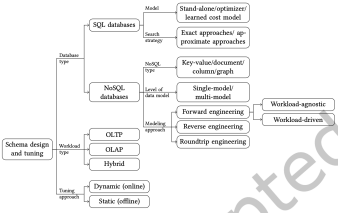


Fig. 5. Taxonomy of Schema Design and Tuning.

Taxonomy of Technics

Study	NoSQL Type	Level of Data Model	Workload	Modeling Approach			Tuning Approach	
				Forward	Reverse	Roundtrip	Static	Dynamic
[1]	Column, Document, Graph	Multi-model	OLTP	✓			✓	
[15]	Column, Document, Graph	Single-model	OLTP	✓			✓	
[45]	Document, Graph	Multi-model (Multistore)	OLTP	✓			✓	
[125]	Document, Graph	Single-model	OLTP	✓			✓	
[149], [144], [177]	Column	Single-model	OLTP	✓			✓	
[184]	Document	Single-model	OLTP	✓			✓	
[66], [5], [11]	Graph	Single-model	OLTP	✓			✓	
[72], [51], [31], [199]	Column	Single-model	OLAP	✓			✓	
[230], [35]	Column, Document	Single-model	OLAP	✓			✓	
[32], [54]	Document	Single-model	OLAP	✓			✓	
[209], [10]	Graph	Single-model	OLAP	✓			✓	
[73]	key-value	Single-model	OLAP	✓			✓	
[218], [29], [164], [91], [95]	Column	Single-model	OLTP	✓			✓	
[71], [117], [173], [116]	Document	Single-model	OLTP	✓			✓	
[18], [19]	key-value, Column, Document	Multi-model	OLTP	✓			✓	
[158]	key-value, Column, Document, Graph	Multi-model	OLTP	✓			✓	
[107]	Column	Single-model	OLTP	✓				✓
[89]	key-value, Column, Document, Graph	Multi-model (Polystore)	OLTP	✓			✓	

Review of Approaches

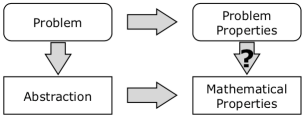
[6] Maryam Mozaffari, Anton Dignös, Johann Gamper, and Uta Störl. "Self-tuning Database Systems: A Systematic Literature Review of Automatic Database Schema Design and Tuning" (May 2024)

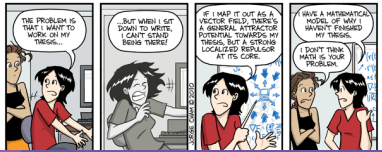
Formal Model

How can we understand/explain the world?

ex: How to find the best denormalized data models?

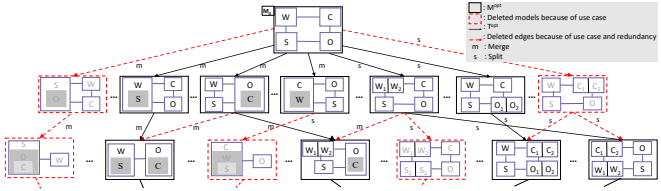
- Make a mathematical abstraction of a certain problem
  - Model: Analytical, stochastic, logical, etc.
  - Often explained using a "`CASE`"
- Prove some important characteristics
  - Based on inductive reasoning, axioms & lemma's...
- Motivate
  - Irrelevant vs relevant factors
  - Which properties are worthwhile (proven)?
    → See `literature survey`





Formal Model: Example

Formal model of data models denormalization[7]:



Theoretical graph of denormalized solutions

- Formal definition of data models' transformation rules;
- Proof of completeness;
- Complexity of the problem: $|\mathcal{M}^*| = (Fn_{|\mathcal{R}|} \times d) \times \prod_{k=1}^{|\mathcal{R}|} B_{|KeySet(r_k)|}$;
- Reduction of the complexity: $|\mathcal{M}^{opt}| = Fn_{|refs(\mathcal{Q})|} \times \prod_{k=1}^{|\mathcal{R}|} B_{|KeySet(\mathcal{Q}_k)|}$.

[7] Jihane Mali, Shohreh Ahvar, Faten Atigui, Ahmed Azough, Nicolas Travers, "DaMoOp: A global approach for optimizing denormalized schemas through a multidimensional cost model". Information Systems, 2026

Simulation

What would happen if ... ?

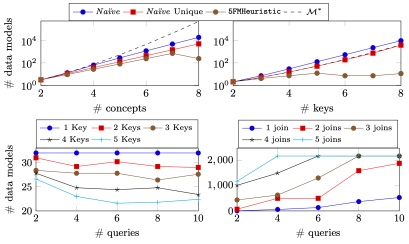
Ex: How to simulate the behavior of data models denormalization?

- Study circumstances of phenomena in detail
  - Real world too expensive; too slow or impossible
- Make prognoses about what can happen in certain situations
  - Test using real observations, typically obtained via a "`CASE`"
  - Heisenberg uncertainty principle[8]
- Motivate
  - Which circumstances are irrelevant (excluded) and which are not (included)?
  - Which properties are worthwhile (to be observed/predicted) ?
    → See `literature survey`

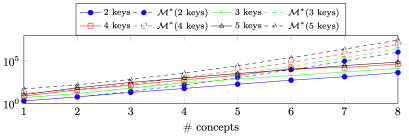
[8] Hazewinkel, Michiel, ed. (2001), "Uncertainty principle", Encyclopedia of Mathematics, Springer Science+Business Media B.V. / Kluwer

Simulation: Example

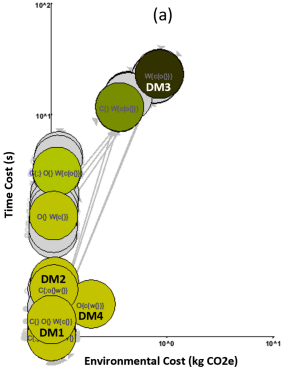
Simulation of the impact of denormalization on use cases[9]:

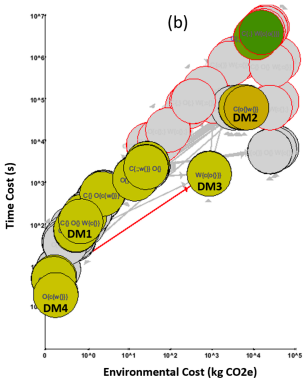


Denormalization impact on queries



Complete use case's impact





Simulation of denormalization impact of environment *vs* time

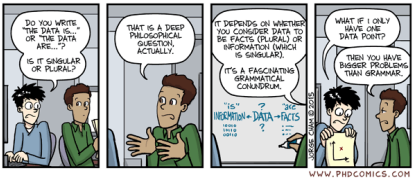
[9] Jihane Mali, Shohreh Ahvar, Faten Atigui, Ahmed Azough, Nicolas Travers, "DaMoOp: A global approach for optimizing denormalized schemas through a multidimensional cost model". Information Systems, 2026

Case studies Revisited

`Feasibility study`	*Proof by construction; often by applying on a "CASE"*
`Demonstrator`	*Demonstrated on a simple yet representative "CASE"*
`Comparative study`	*Score criteria check-list; often by applying on a "CASE"*
`Observational study`	*Observing a series of "CASES"*
`Literature survey`	*"CASES" = selected papers*
`Formal model`	*Often explained using a "CASE"*
`Simulation`	*Test prognoses with real observations obtained via a "CASE"*

Spectrum of cases

- **Toy Example**
  - Created for explanation (*foo/bar, Alice/Bob*)
  - Simple model; illustrates differences
- **Exemplar**
  - Accepted teaching vehicle ("*textbook* example")
  - Simple but illustrates relevant issues
- **Case**
  - *Real-life* example
  - Context is difficult to grasp
- **Community Case**
  - *Competition*; approved by community, comparing, Kaggle
- **Benchmark**
  - Approved by community (TPC-H, TPC-C, YCSB)
  - Known context
  - "Planted" issues



Case study

A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between the phenomenon and context are not clearly evident

[Robert K. Yin. “Case Study Research: Design and Methods”, p. 13]

Experiment

Case Study

⚠ Counter-Examples ⇒ Formal generalization is overvalued

Steps to Perform Experiments

1. Review relevant literature
2. Specify the case study
   - Define objectives/hypotheses
   - Specify the population
   - Evaluate the testing feasibility
3. Define research procedure (see next slide)

Research Procedure

Planning experiments/Case study

- Treatment design
- Sampling/XP design/#replicates
- Measurements choice
- Units of observation
- Border effects / adjacent units
- Expected results
- Outline of analyses to do
- Measuring instruments

Experimenting/Studying

- Install experiment
- Collect Data: each step!
- Complete analysis of data
- Finally, prepare a complete, correct, and readable report.

Need to restart? ⇒ Refine the "planning"

The 3 R's of experimentation

1. **Repeat**
   - Provide a measure of variation / error term;

   ⇒ Validity of conclusions.
2. **Randomize**
   - Treatments in a purely random manner;

   ⇒ Prevents bias.
3. **Request Help**
   - Not everyone is a statistician;

   ⇒ Do it when planning, not after.

Protocol Example

- For each **Research Question**
- **Treatment Design**: independent vs dependent variables
- **Experimental steps**: *input/output* for each step, quality *measure*, expected results

What is an *Electronic Laboratory Notebook* (ELN)?

- A digital platform to record, organize, and manage experimental data
- Replaces traditional paper lab notebooks
- Enables structured documentation of:
  - Experimental protocols
  - Observations and results
  - Data files and metadata
- Accessible via web interface (local server or cloud)
- Supports collaboration among researchers

How ELNs Support Scientific Experimentation

- **Reproducibility:** Detailed and standardized experiment records
- **Traceability:** Automatic timestamps and version control
- **Collaboration:** Shared access across teams and institutions
- **Data Integrity:** Secure storage and audit trails
- **Efficiency:** Searchable content and centralized information
- **Compliance:** Supports regulatory standards (e.g., GDPR, FAIR principles)

Criteria	eLabFTW	SciNote	OpenBIS	JupyterLab
Deployment Ease	++++	++++	+++	+++
User Interface	++++	++++	+++	+++
Traceability/Compliance	+++++	++++	++++	++
Life Sciences Fit	++++	++++	+++++	+++
Community Activity	++++	++++	+++	+++++

Some Practice on Use Case and XP (today)

The goal is to plan some experiments on expected results

- **My** work (**Phase 0**):
  - Collected papers from related scientific domains;
  - Removed experiments, title and authors (blinded);
  - Associated with *Use Case* categories (sometime several ones);
  - Put them **HERE**.

Some Practice on Experimenting

Phase 1 - This morning (until noon)

- **Your** work:
  - Choose an article: 1 article = group of **3/4 persons**.
    The choice form can be modified here
    Everybody must specify is name
  - Read the article (without experiments);
  - ⚠DO NOT SEARCH THE ARTICLE ON THE WEB/LLM⚠;
  - Propose a **title** and **experiments** (protocol and expected results);
  - Write and explain *your* XP protocol in a document.
    ⚠An LLM can be used for writing the protocol (not reading) - if so, you MUST specify it on your document;
  - Connect to `Easychair` and if necessary **create your account**: easychair
  - Submit the original article AND your "XP" on `Easychair`.
    ⚠ **Add all your "co-authors" during submission**
    Deadline: Before NOON.

Some Practice on Reviewing

Phase 2

- Your work:
  - **Congratulations, you are in the Program Committee!** (I need your *Easychair* accounts!!!!)
  - Do your biddings on Easychair (available Today until 7pm);
  - You will receive your assignment from Easychair, review it with proposed XP protocol (blinded), detailed review form:

Title	Help text	Kind	has text	text required	text visible to authors	has scores	score required	score visible to authors	Scores
Proposed title	The proposed title is appealing	text/score	✔	✔	✔	✔	✔	✔	none; simple (code of the article...); Not really appealing; Appealing; Attractive – I wish to read this article!
XP adequacy with the article	The detailed provided in the XP are related to the model / contributions provided by the article. How much adequate are the XP and helps to enhance the contributions?	text/score	✔	✔	✔	✔	✔	✔	(none); Irrelevant; Somehow adequate; Good for some points: datasets / competitors / measurements / predicted resuls; All points are good
Detailed protocol	Comment the quality of the experimental protocol. How much the experimental protocol is detailed?	text/score	✔	✔	✔	✔	✔	✔	None; Just a list of items; Few details; Argues for choices; Very detailed XP
Article difficulty	How much the complexity of this article helped the designing of the experiments?	text/score	✔		✔	✔	✔	✔	Easy. Too much details were given in the article; Some details details guided; Few details. Most experiments are new; No details. great job!

  - Deadline: tomorrow **12pm**.