Methodology of research

Mathieu Lagrange

. https://github.com/mathieulagrange/merec

Outline

Outline

Outline

Outline

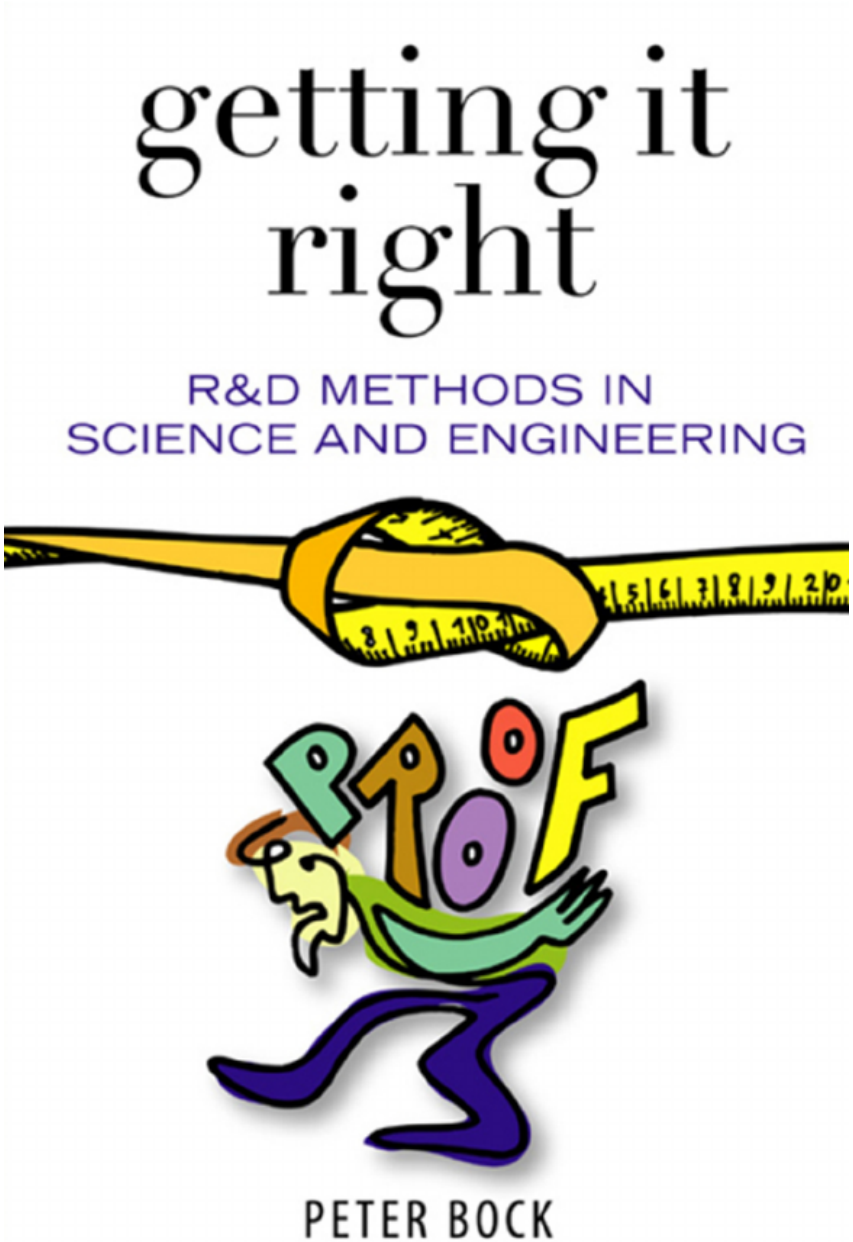
Outline

Outline

Outline

Acknowledgment

- Those slides are based on the beautiful book
- "getting it right" written by Peter Bock
- and `http://phdcomics.com`


getting it right
R&D METHODS IN SCIENCE AND ENGINEERING
PROOF
PETER BOCK

Terms

R&D

- **Research** is a process that acquires new knowledge
- **Development** is a process that applies knowledge to create new devices or effects

Research versus Development

- Historian / Politician
- Scientist / Engineer
- Psychologist / Therapist
- Physiologist / Doctor
- Linguist / Translator
- Economist / Investor

Getting it wrong

- Research results cannot be reproduced because of poor methodology and documentation.
- Speculations are not identified as such and are intermixed with supported conclusions.
- Knowledge is precarious, locked up in the heads of individuals.
- Data collection is haphazard and confounded with political issues.
- Experiment methods are chaotic, dominated by a "try this, try that" mentality.
- Experiment processes cannot be audited or reviewed due to a lack of logs and records.
- Statistical analysis of results is missing or naive.

Requirements to get it right

- a research notebook that contains a complete record of all your activities, updated faithfully and regularly
- a thorough understanding of descriptive statistics to avoid blind reliance on statistical packages
- thorough knowledge of all functions and capabilities of a powerful data processing program
- access to a wide variety of specialized software packages for data analysis and visualization

The notebook approach

- Put EVERYTHING in your notebook — use it like a diary and a scrapbook
- Keep your notebook with you at all times so that you are always ready to record your thoughts and ideas.
- Your notebook is your private record. No one else need ever see it, although you may choose to share items.
- Specify the sources of nonoriginal information.
- Do not erase major "mistakes." Cross them out and explain them. You may need to re-evaluate them later.

Where to write ?

You can use

- physical notebooks
- numeric ones
- using or not online tools

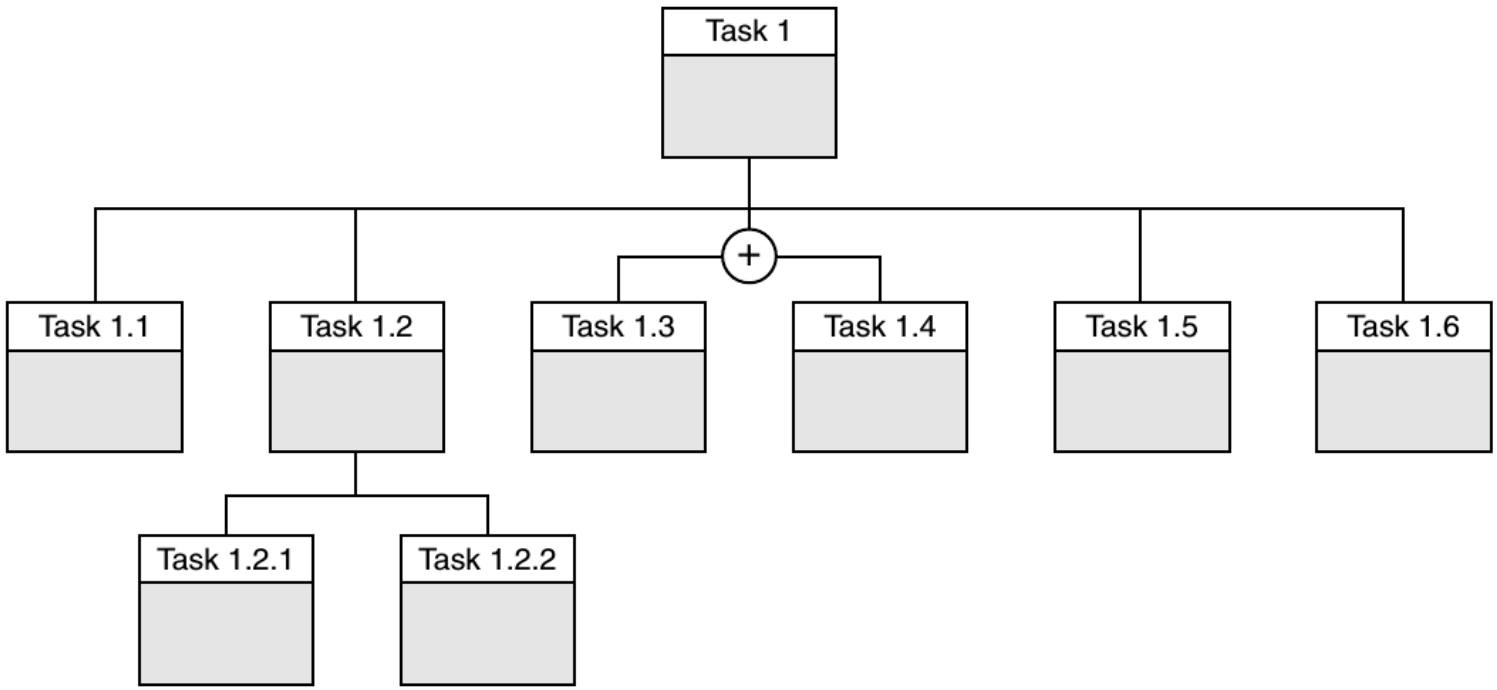
Tasking

The tasking approach

The tasking approach

- applies the "divide and conquer" approach
- to project managment
- viewed as a tree of dependencies
- viewed as timed milestone chart

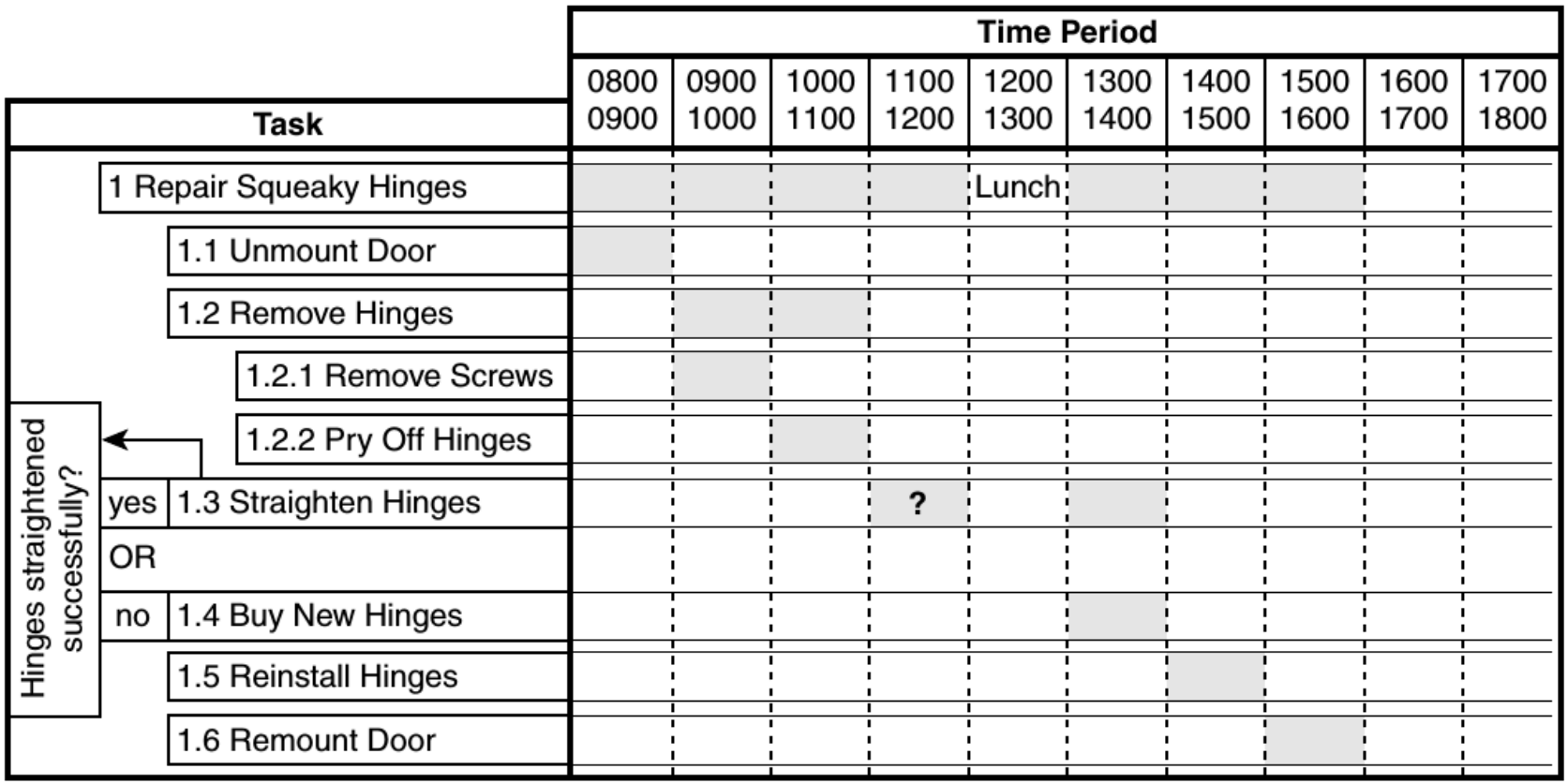
The task tree



The task tree

Primary Task	Descendant Tasks
1) Task 1	all 8 subtasks
2) Task 1.1	none
3) Task 1.2	Tasks 1.2.1 and 1.2.2
4) Task 1.3	none
5) Task 1.4	none
6) Task 1.5	none
7) Task 1.6	none
8) Task 1.2.1	none
9) Task 1.2.2	none

The milestone chart



The task

An R&D task applies

- a specified **method**
- to the **domain** of the task
- with the objective of obtaining a satisfactory result
- in the **range** of the task.

Task domain

The task domain

- comprises the **task** unit
- and the **resources** necessary
- to achieve the **task objective.**

Task objective

The task objective

- is a statement of what the task is intended to achieve,
- expressed as an infinitive phrase.

The task unit

The task unit

- is the set of objects or concepts
- that undergoes some required alteration
- before or during the task and measurement during the task.

Physical examples

- a car seat (whose comfort is being evaluated);
- a rock layer on an exposed cliff face (whose geological age is being estimated);
- a highway with badly congested traffic (whose speed is being measured);
- a cohort taking a new pharmaceutical drug (whose effectiveness is being measured);
- a sample of fruit (whose freshness is being evaluated).

Abstract examples

- the algebraic expression of a mathematical conjecture (that is to be proved);
- the content of journal articles (that are being summarized into a set of short abstracts);
- a software graphical user interface (GUI) (whose effectiveness is being evaluated);
- a linguistic principle (that explains structural differences among related languages);
- a new mechanism for government-citizen interaction (whose efficacy is being measured).

The task resources

- Inducer : an inducer is a device or mechanism that alters the task unit during or before the task
- Sensor : A sensor is a device that acquires the required data from the task unit
- Supervisor : the supervisor for a task is the set of human and automated agents that operates and monitors the task unit, the concurrent inducers, the sensors, and itself
- Channels : the channels in a task domain interconnect the resources of the task to provide the means for exchanging energy and information among them
- Domain knowledge

The task method

The task method

- comprises the solution specified to achieve the task objective
- and the experiments designed to determine the effectiveness of the solution :
- everything required to transform the task domain into the task range.

The task range

The task range

- comprises all the products of the task,
- including knowledge,
- devices,
- and effects.

Method

The modern scientific method

The Scientific Method comprises four sequential phases

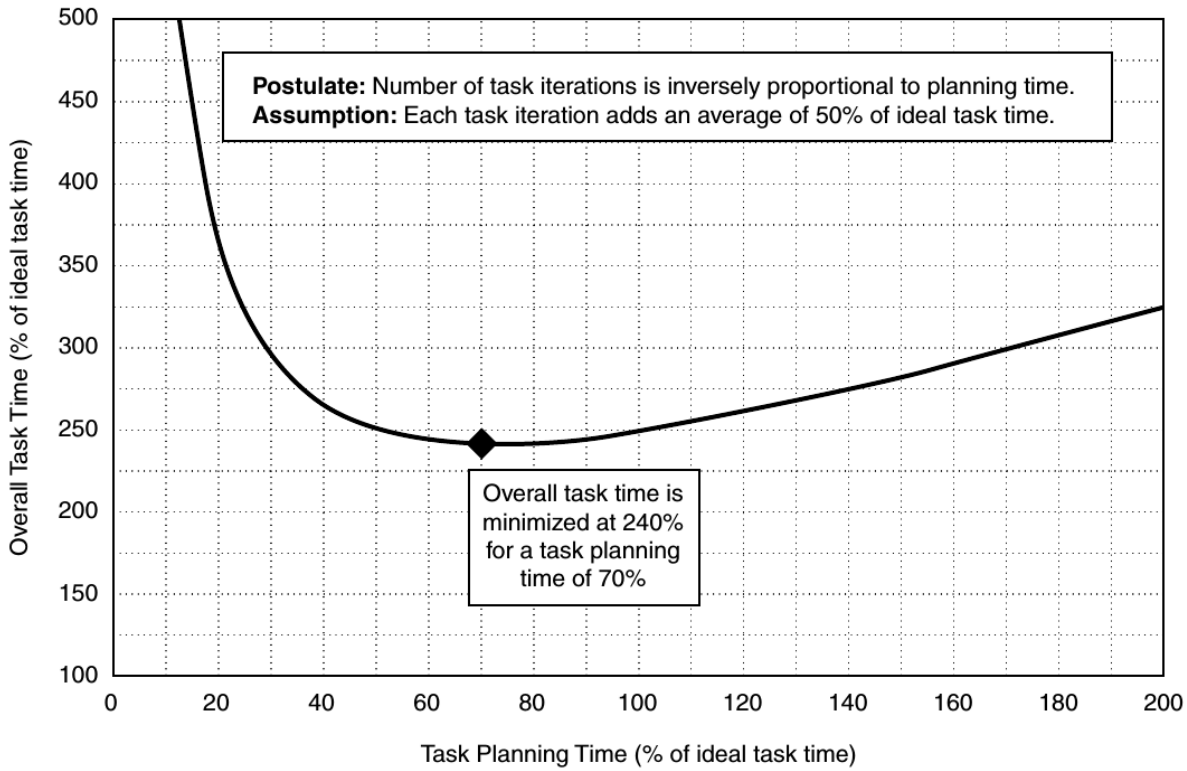
- Analysis,
- Hypothesis,
- Synthesis,
- and Validation

which are applied to a task iteratively and recursively to achieve the objective of the task.

Stepwise refinement

- **Analysis** : Describe Problem, Set Performance Criteria, Investigate Related Work, State Objective
- **Hypothesis** : Specify Solution, Set Goals, Define Factors, Postulate Performance Metrics
- **Synthesis** : Implement Solution, Design Experiments, Conduct Experiments, Reduce Results
- **Validation** : Compute Performance, Draw Conclusions, Prepare Documentation, Solicit Peer Review

Planning



In general, too little planning time incurs much heavier penalties in overall task time than too much planning time.

Feasability

Feasibility pilots are useful for exploring the feasibility of proposed task components.
Every pilot

- although perhaps only minimally planned
- and not usually reported in the final project documentation,
- must be fully recorded in the cognizant team members′ research note- books.
- Once a pilot has yielded a potentially useful result, it must be confirmed and validated by planning and executing a formal task.

Effort allocation

For **scientists and engineers** :

- Resist the urge to "make it up as you go along."
- Plan the project thoroughly in advance, but allow the plan to evolve as experience and pilots reveal the flaws and potential improvements in the plan.
- Be prepared to give your supervisors time and cost estimates whose accuracy they can trust.
- If you discover you will not be able to make a deadline, notify your manager immediately ; don't wait until the last moment.
- Understand that supervisors see and deal with a larger perspective than you do. Accept that the workplace is not a democracy.

Effort allocation

For **supervisors** :

- Support your scientists and engineers with the time and resources necessary to plan and stepwise refine these plans throughout the lifetimes of projects.
- Limit the work of each person to one or two concurrent projects.
- Don't steal time from project scientists and engineers for "fire drills" ; hire special staff members for that purpose. Don't micromanage. Don't nickel-and-dime.
- Try not to burden your scientists and engineers with bureaucratic tasks ; that's your job.

Analysis

The analysis phase

The objective of the Analysis Phase of the Scientific Method is :

- to gain a thorough understanding of the components of the problem domain,
- leading to the formulation
- of a single specific and reasonable task objective.

Describe problem

A problem statement is expressed as

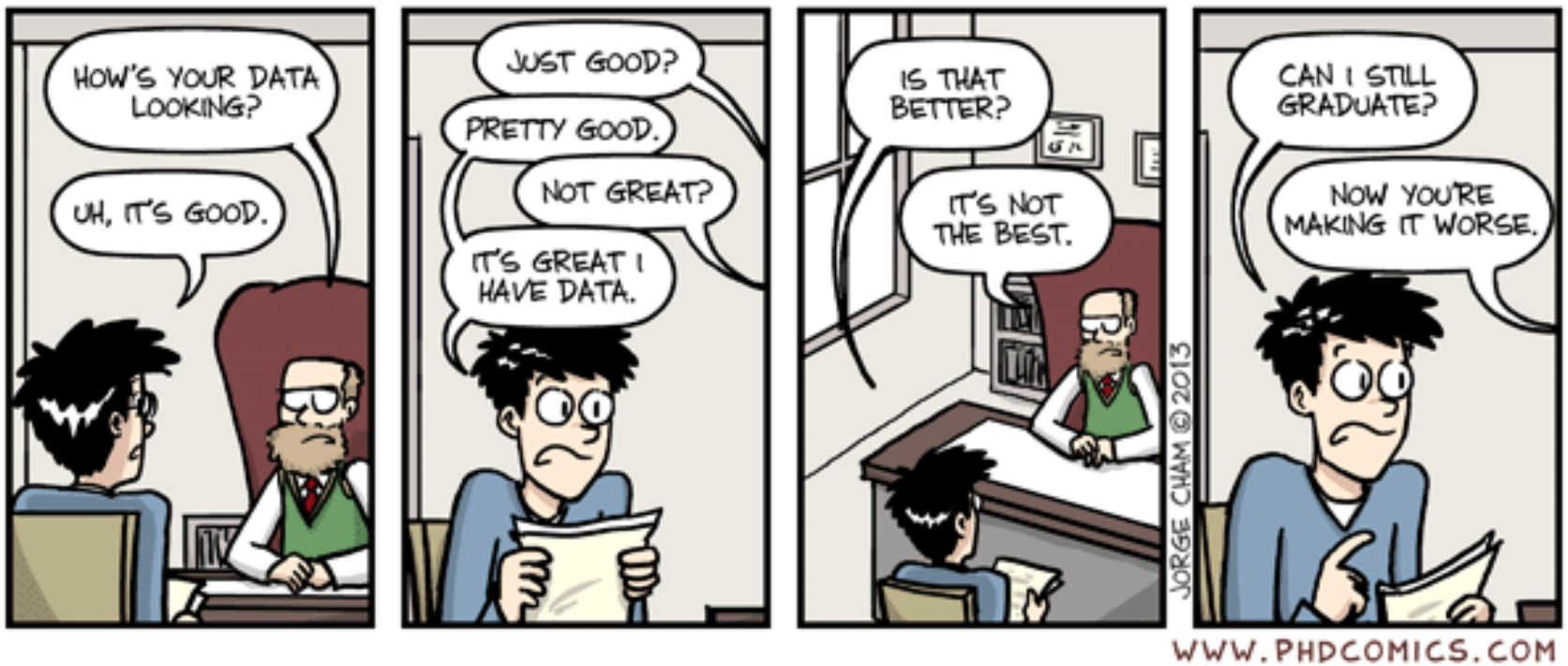
- an interrogative sentence,
- a declarative sentence,
- or an imperative sentence
- that summarizes a question, complaint, or requirement, respectively.

Describe problem

Set performance criteria

- Carefully choose the performance criteria
- Beware of conflicts
- Set boundaries

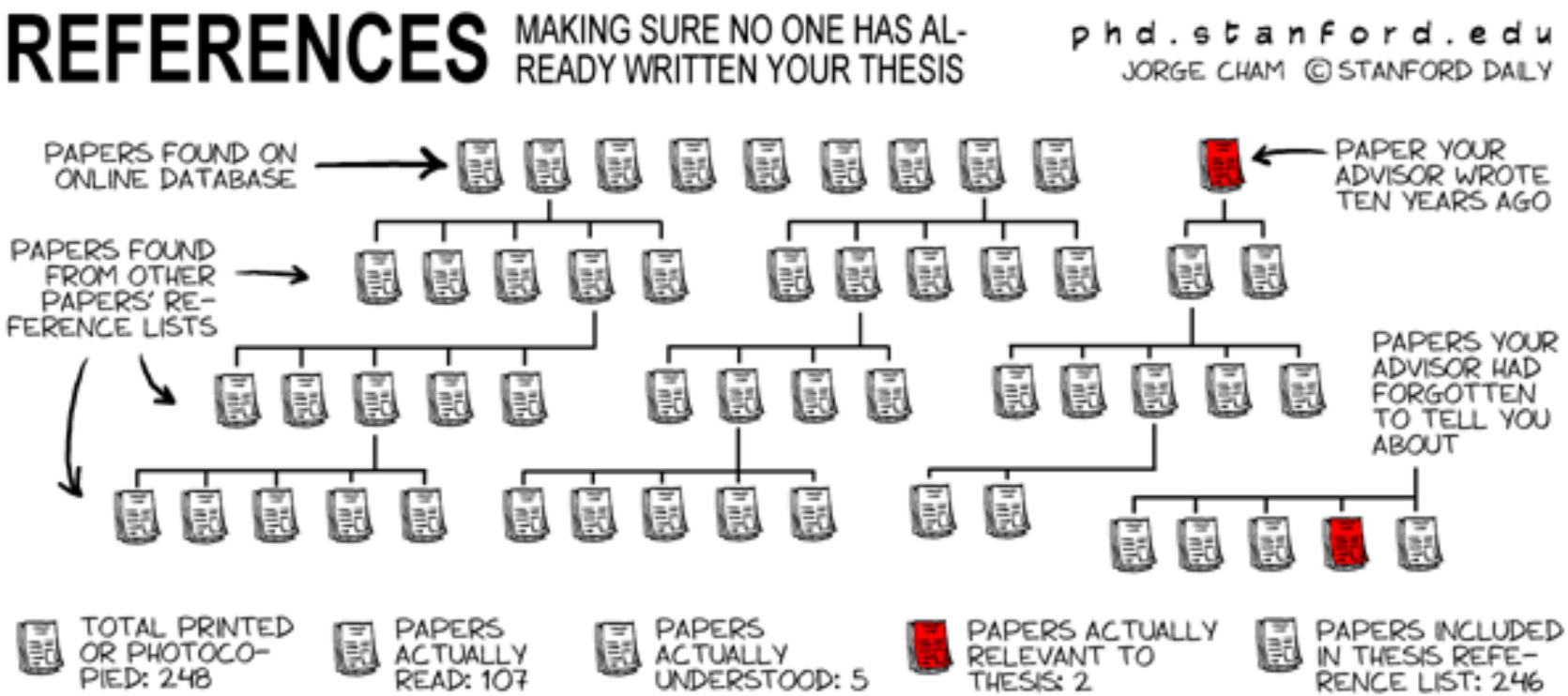
Set performance criteria



Investigate related work

- Don't drown
- read a lot, but pick only a few that you deeply understand
- target material size according to your project

Investigate related work

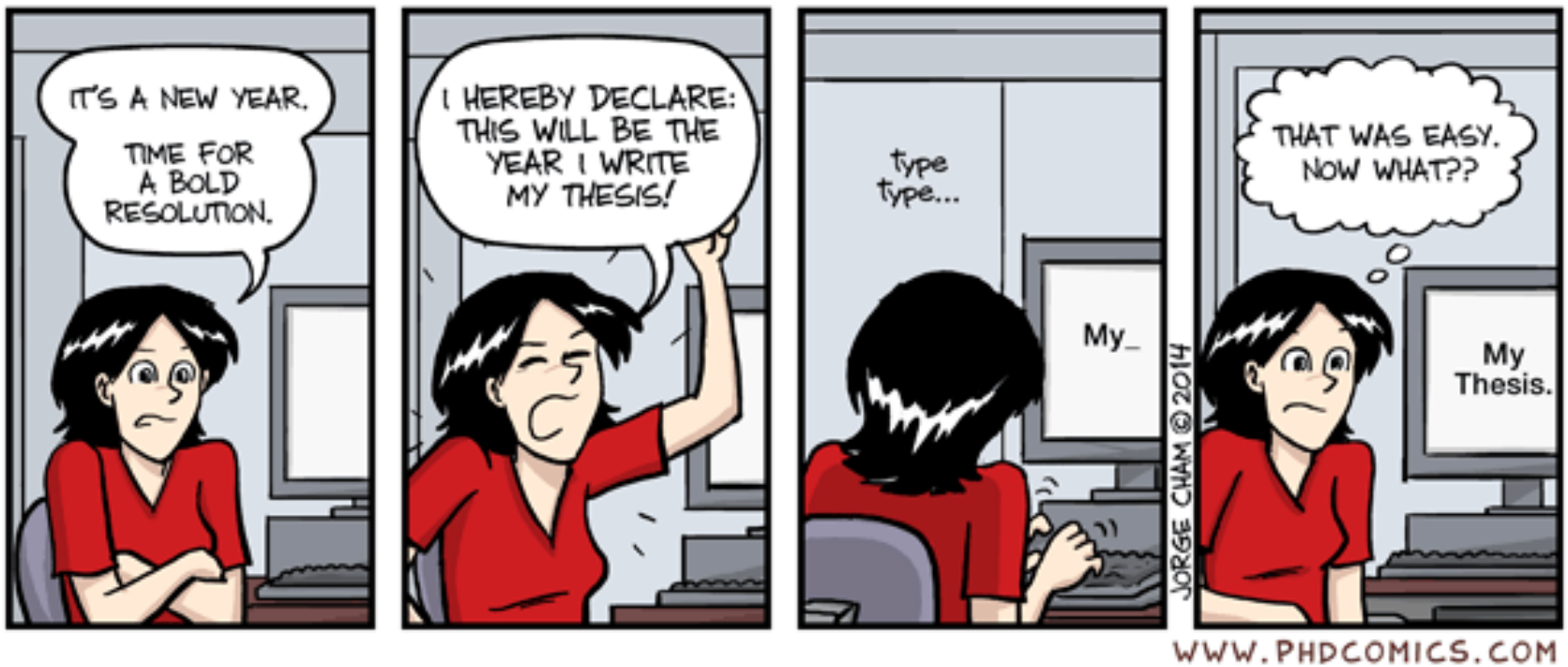


State objective

The task objective

- is a contract between the task team and management
- it often has a political component
- thus it shall be clear and objective

State objective



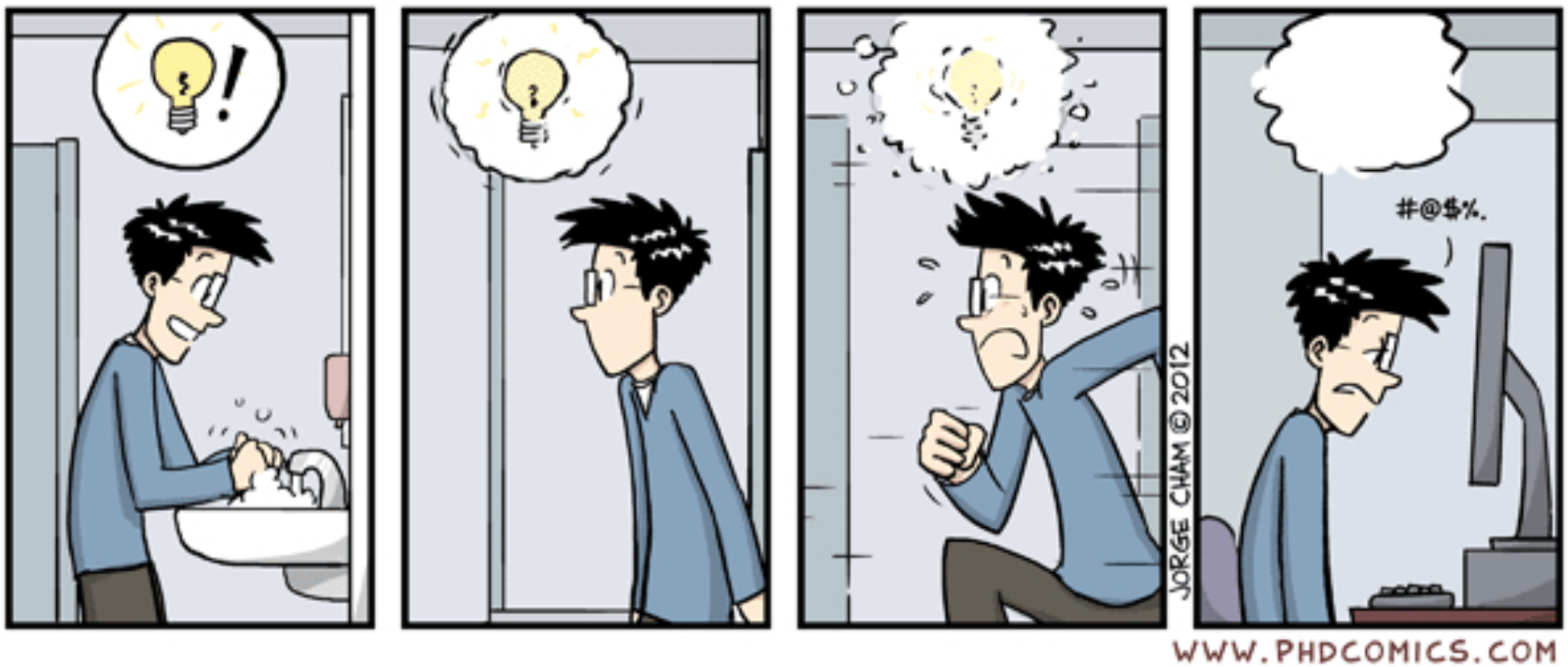
Hypothesis

The Hypothesis Phase

The objective of the Hypothesis Phase is to propose

- a solution to achieve the task objective,
- a set of goals and hypotheses for this solution,
- and the factors
- and performance metrics for testing the validity of the solution.

Specify Solution



Set Goals

Every task has at least one goal :

- to determine the response of the task unit
- to the application of the solution (or part of the solution),
- expressed as an infinitive phrase.

Set goals

Define factors

Use OP (operating-point) pilots

- to search for suitable values for fixed parameters and conditions
- and dynamic ranges and increments for factors.

Postulate Performance Metrics

A performance metric is a postulate that

- transforms the results of the task into measures of performance
- for drawing conclusions about the task objective.

Synthesis

The Synthesis Phase

The objective of the Synthesis Phase of the Scientific Method is to

- implement the task method (solution and experiments)
- to accomplish the goals
- and validate the hypotheses of the task.

Implement Solution

- If it meets the needs of the task, using an existing solution
- is almost always more cost-effective
- than manufacturing it yourself.
- if outsourcing, Arrange to be given complete, raw, and unpreprocessed data.
- if not, document your code!

Implement Solution

Design Experiments

An experiment acquires data

- to measure the performance of the solution
- under controlled conditions
- in a laboratory.

Design Experiments

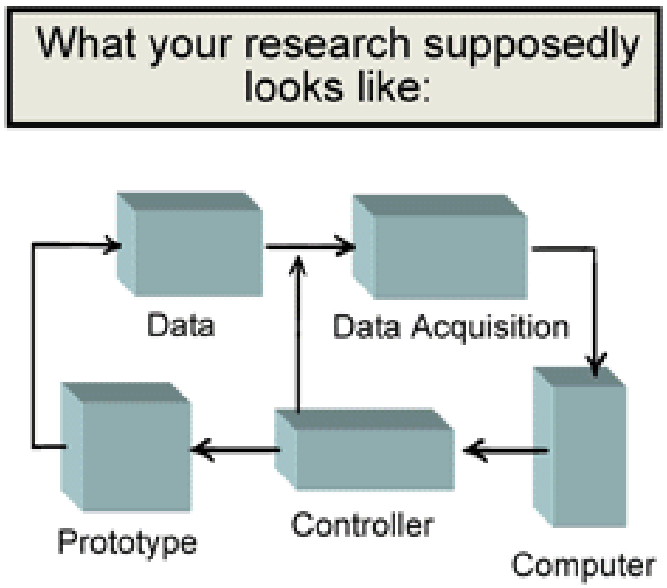


Figure 1. Experimental Diagram

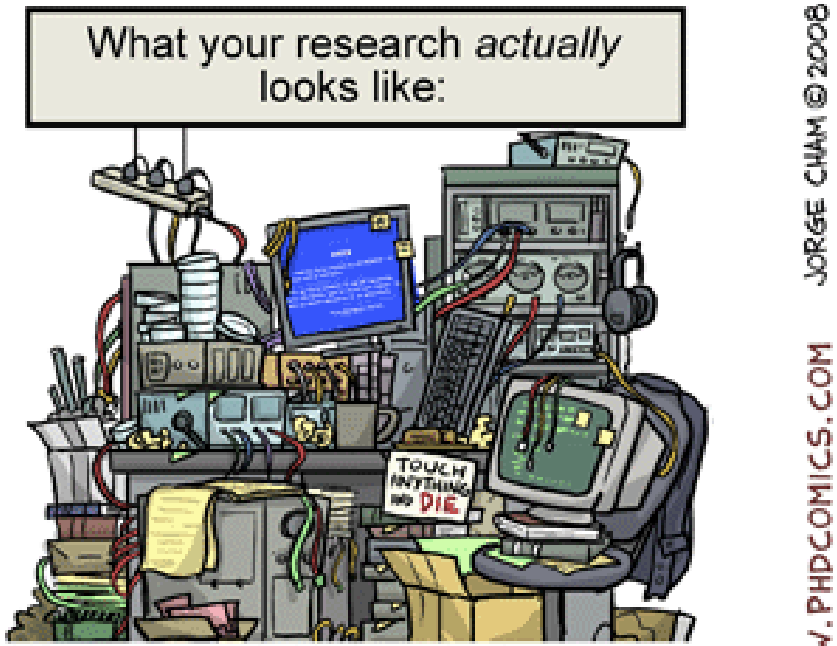


Figure 2. Experimental Mess

Definitions

- A **cohort** is a task component consisting of a set of plants or animals, all of which are assumed to be equivalent for the requirements of the task objective ; a member of a cohort is usually called a **subject**.
- A **sample** is a task component consisting of a set of nonliving objects or concepts, all of which are assumed to be equivalent for the requirements of the task objective ; a member of a sample may be called by several names : **instance ; response ; data point ; specimen ; item ; or case**.

Definitions

- An experiment **treatment** is a combination of one level or setting for every factor.
- An experiment **trial** is a complete set of treatments applied to a member of the task unit during the experiment.
- An experiment **block** is a set of trials that provides a cover of the factor space that is appropriate and adequate for achieving the task objective.
- A **control trial** measures the performance of one set of task components in the absence of another set of task components to isolate the effects of the included components on performance.

Conduct Experiments

- Do exactly what you stated
- Regard a failed experiment as a pilot

Reduce Data

- Because the data reduction methods and/or performance metrics may have to be modified or corrected,
- record both the raw and reduced results of experiments.

Validation

The Validation Phase

The objective of the Validation Phase of the Scientific Method

- is to decide whether the objective of the task has been achieved,
- based on formal conclusions about its goals
- and hypotheses and a rigorous peer review of the task methodology.

Compute performances

Conduct a pilot

- to analyze the raw or reduced results
- to detect any unusual artifacts or unexpected trends stemming from biases in the task methods
- and to confirm the appropriateness of the postulated performance metrics.

Draw Conclusions

Before abandoning formal rigor in the face of what seem to be obvious conclusions,

- check the distributions of the underlying performance values.
- Lacking estimates of the uncertainty of these values, when asked if their averages (or other location parameters) are significantly different,
- your answer must be "I don't know. I do not have enough information."

Prepare Documentation

- **Analysis** : Problem Statement (sentence), Performance Criteria (short phrases), Primary Objective (infinitive phrase)
- **Hypothesis** : Solution Description (list or diagram), Goals (tree or outline), Factors (table with ranges), Performance Metrics (formulas)
- **Synthesis** : System and Environment (diagram), Experiment Block Design (table)
- **Validation** : Performance (tables or charts), Conclusions (sentences), Informal Observations (sentences), Recommendations (list of phrases)

Prepare Documentation

Analysis :

- Problem Statement : sentence
- Performance Criteria : short phrases
- Primary Objective :infinitive phrase

Prepare Documentation

Hypothesis :

- Solution Description : list or diagram
- Goals : tree or outline
- Factors : table with ranges
- Performance Metrics : formulas

Prepare Documentation

Synthesis :

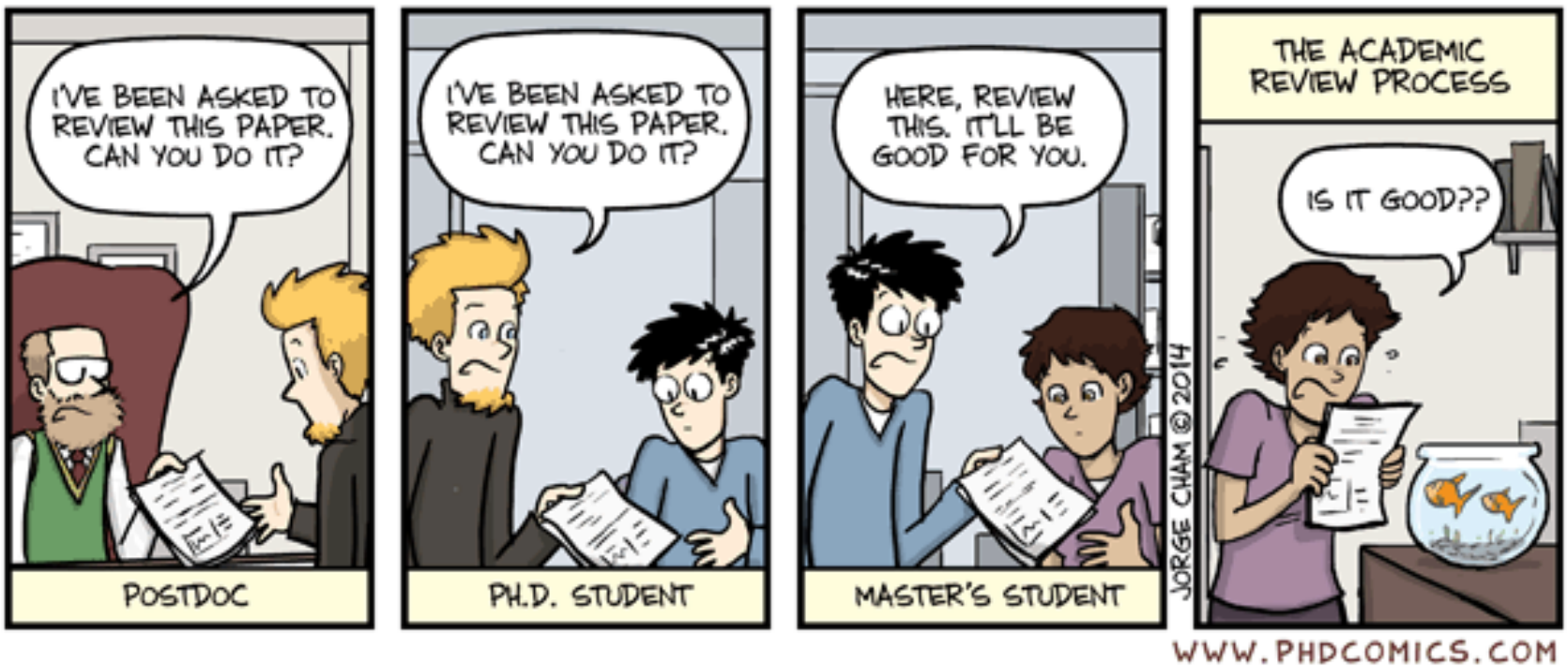
- System and Environment : diagram
- Experiment Block Design :table

Prepare Documentation

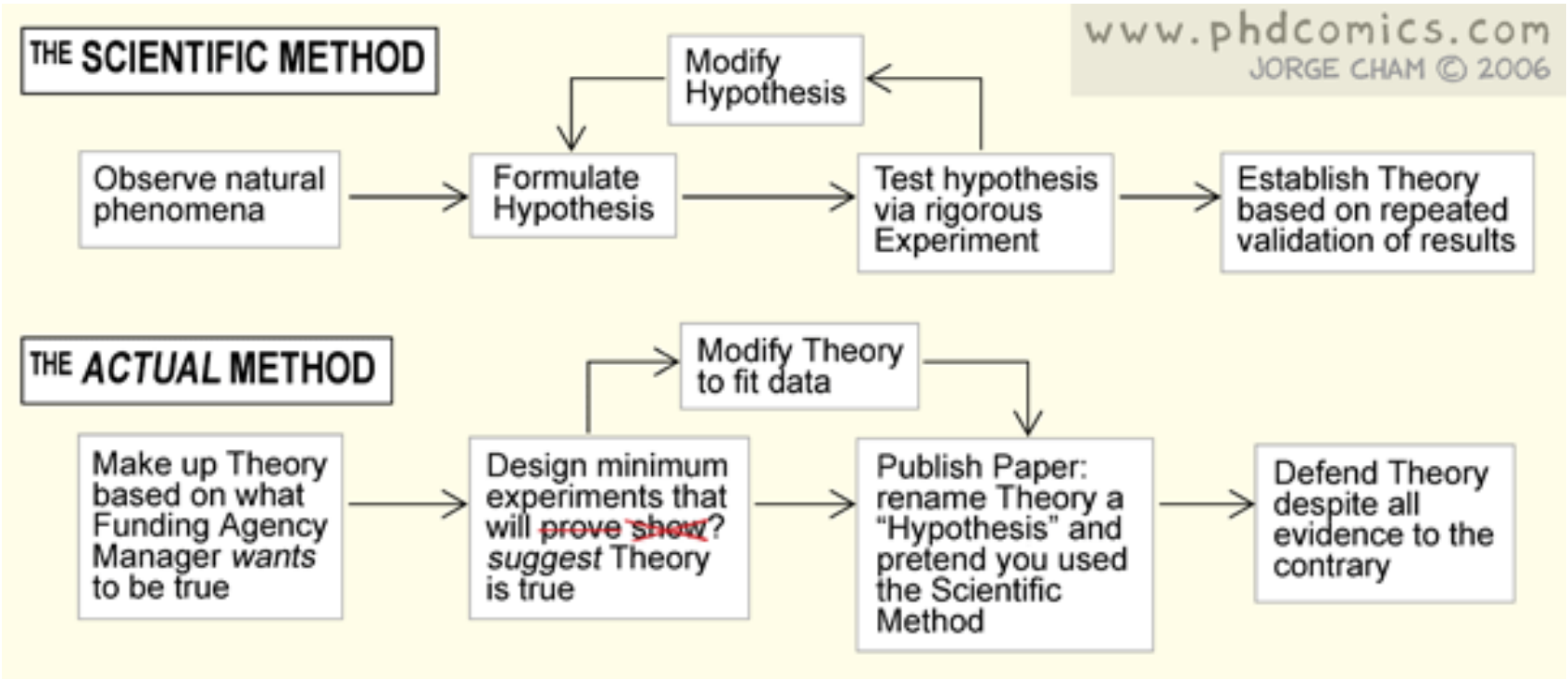
Validation :

- Performance : tables or charts
- Conclusions : sentences
- Informal Observations : sentences
- Recommendations : list of phrases

Sollicit Peer Review



Now, you know...


THE SCIENTIFIC METHOD
Modify Hypothesis
Observe natural phenomena
Formulate Hypothesis
Test hypothesis via rigorous Experiment
Establish Theory based on repeated validation of results
www.phdcomics.com
JORGE CHAM © 2006
THE ACTUAL METHOD
Modify Theory to fit data
Make up Theory based on what Funding Agency Manager wants to be true
Design minimum experiments that will prove show? suggest Theory is true
Publish Paper: rename Theory a "Hypothesis" and pretend you used the Scientific Method
Defend Theory despite all evidence to the contrary