# Introduction to the practical neuroimaging course

Matthew Brett

February 13 2015

# The idea

We want to make it easier to think by showing:

- how imaging analysis steps work;
- how analysis relates to other fields like engineering and statistics;
- how to reduce error;
- how to collaborate with other people on imaging analysis;
- how to work efficiently so you can think about the problem rather than the tools.

# Building

"What I cannot create, I do not understand"

Richard Feynman

# Cross-referencing

- Convolution;
- optimization;
- linear algebra;
- multiple regression.

# Raise standards for preclinical cancer research

**C. Glenn Begley** and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.

# It does not come out in the wash

**REPRODUCIBILITY OF RESEARCH FINDINGS**

Preclinical research generates many secondary publications, even when results cannot be reproduced.

| Journal impact factor | Number of articles | Mean number of citations of non-reproduced articles* | Mean number of citations of reproduced articles |
|---|---|---|---|
| >20 | 21 | 248 (range 3–800) | 231 (range 82–519) |
| 5–19 | 32 | 169 (range 6–1,909) | 13 (range 3–24) |

Results from ten-year retrospective analysis of experiments performed prospectively. The term 'non-reproduced' was assigned on the basis of findings not being sufficiently robust to drive a drug-development programme.

*Source of citations: Google Scholar, May 2011.

# Take no-one's word for it



(by kladcat under CC BY 2.0, via Wikimedia Commons)

# Ubiquity of error

> *The scientific method's central motivation is the ubiquity of error - the awareness that mistakes and self-delusion can creep in absolutely anywhere and that the scientist's effort is primarily expended in recognizing and rooting out error."*

Donoho, David L, et al. 2009. Reproducible research in computational harmonic analysis. *Computing in Science & Engineering* 11, 8–18.

# Don't do breezy

Computation–based science publication is currently a doubtful enterprise because there is not enough support for identifying and rooting out sources of error in computational work.

In my own experience, error is ubiquitous in scientific computing, and one needs to work very diligently and energetically to eliminate it. One needs a very clear idea of what has been done in order to know where to look for likely sources of error. I often cannot really be sure what a student or colleague has done from his/her own presentation, and in fact often his/her description does not agree with my own understanding of what has been done, once I look carefully at the scripts. Actually, I find that researchers quite generally forget what they have done and misrepresent their computations.

Computing results are now being presented in a very loose, "breezy" way—in journal articles, in conferences, and in books. All too often one simply takes computations at face value. This is spectacularly against the evidence of my own experience. I would much rather that at talks and in referee reports, the possibility of such error were seriously examined.

David Donoho - An invitation to reproducible computational research (2010)

# Don't sell the story

> *An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures"*

The wavelab front page

# Disbelief

Richard Feynman, What is Science? (1969)

> *Science alone of all the subjects contains within itself the lesson of the danger of belief in the infallibility of the greatest teachers in the preceding generation... Learn from science that you must doubt the experts ... Science is the belief in the ignorance of experts*

# Conversation

A scientist:

> *I analyzed these data with my favorite software; the analysis showed activation in the frontal lobe.*

Another scientist:

> *I don't believe you. Show me what you did in detail so I can check if you made a mistake.*

# Working efficiently

- Simple comes from Latin *simplex* - one fold. The opposite of *complex* - Latin for twisted together.
- Easy comes from old French *aiser* to Latin *adjacens* "lying close by".

Easy is what you are familiar with, not far from something you already know. Simple is when you have succeeded in breaking the problem into separate ideas.
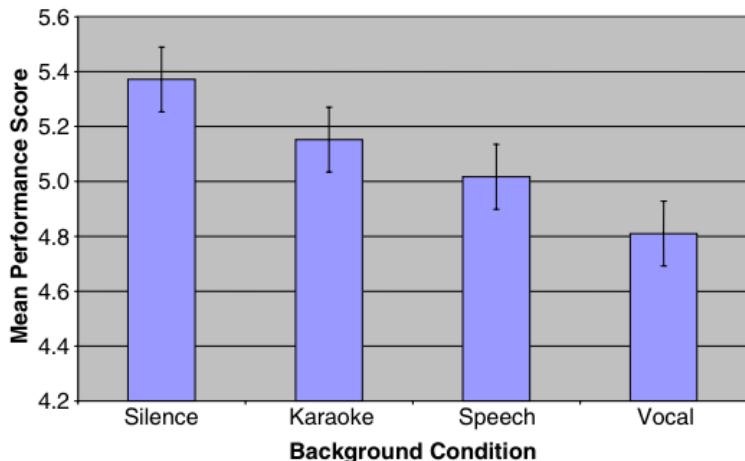
See: http://www.infoq.com/presentations/Simple-Made-Easy

Also: Oxford dictionary of Word origins edited by Julia Cresswell.

# It is worth working towards simple
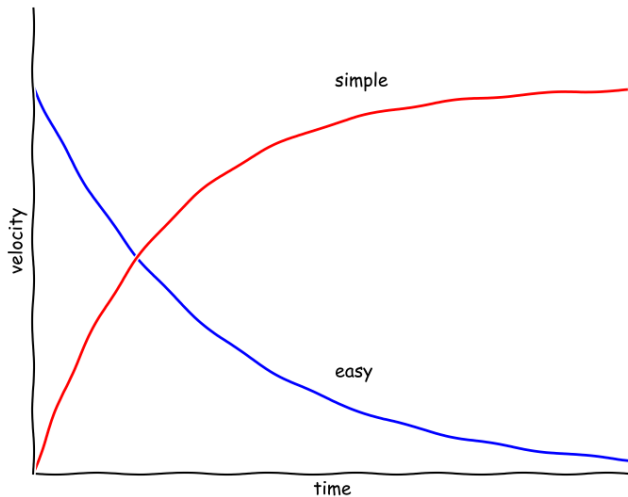
# Thinking two things



The Relative and Perceived Impact of Irrelevant Speech, Vocal Music and Non-vocal Music on Working Memory, Curr Psychol (2008) 27:277–289

# Why Python?

- Simple - and easy;
- fully free and open;
- culture of continuous testing;
- open means collaboration;
- rapid growth in science and elsewhere;
- mature stack of scientific packages;
- well adapted (via packages) to scientific problems.

# Why git?



popcon graph
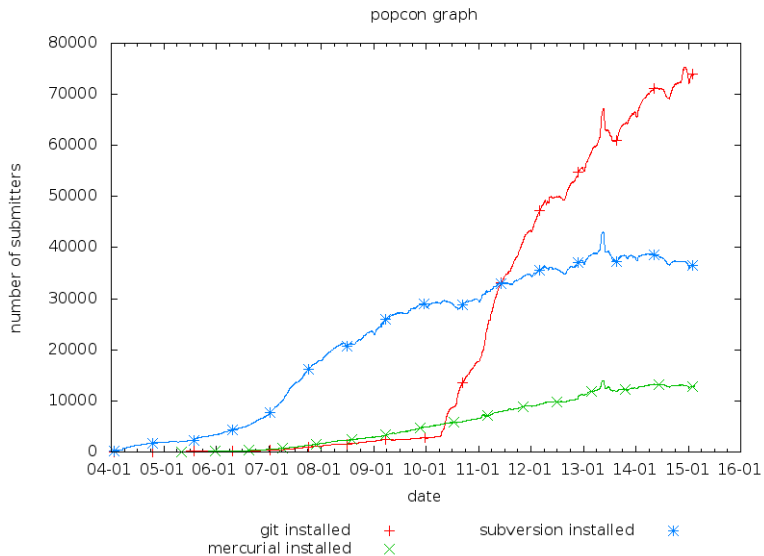
# Your text editor

*Use a Single Editor Well*

*The editor should be an extension of your hand; make sure your editor is configurable, extensible, and programmable.*

"The Pragmatic programmer: from journeyman to master" by Andrew Hunt, David Thomas

# Admin

- Around 30 minutes homework;
- Homework ready by end of day Monday for following Friday course;
- Class consists of:
  - 5 minute debrief from previous class;
  - 30 minute talk and introduction to problems;
  - 60 minutes working on problems;
  - 10 minutes reviewing.
- Office hours;

# Introduction to the exercises

- Open terminal;
- Make a directory "pna";
- Change directory to "pna";
- `git clone git://github.com/practical-neuroimaging/pna2015.git`
- `cd pna2015`
- `ls`
- `ipython notebook`
- select and open "what is an image" notebook;