

The Science(?) of Documentation



Paul Anzel

Data Scientist, Metromile Inc.

<https://github.com/anzelpwj> (← slides available here)

Let's go to the GOOG



software documentation

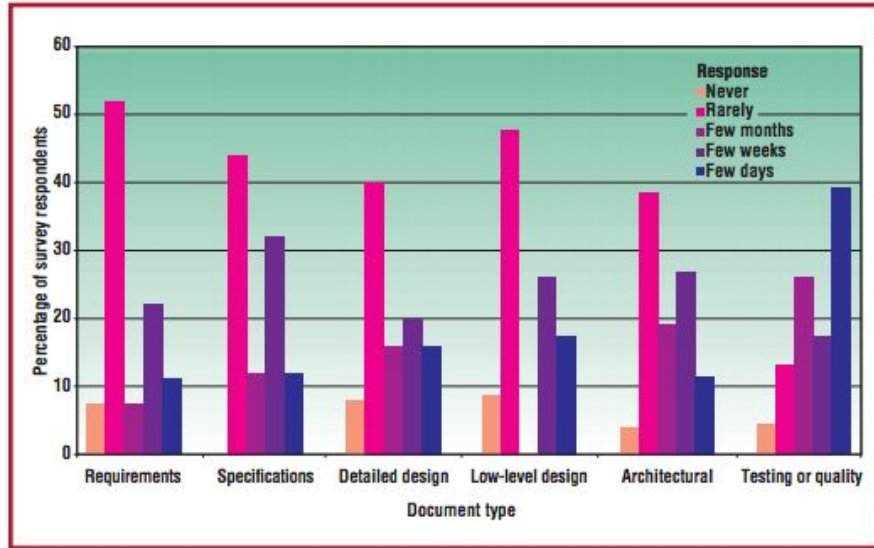


☒ Articles (☒ include patents) ☐ Case law

Stand on the shoulders of giants

Most people think documentation is out of date, and go to source code

Figure 1. The time between system changes and documentation updates for different documentation types.



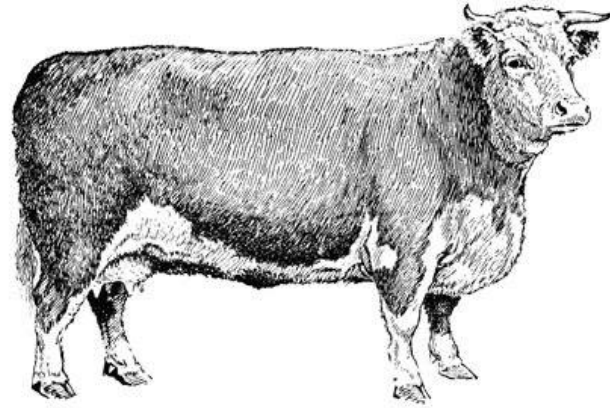
Lethbridge TC, Singer J, Forward A. How software engineers use documentation: The state of the practice. IEEE software. 2003 Nov;20(6):35-9.

Forward A, Lethbridge TC. The relevance of software documentation, tools and technologies: a survey. In Proceedings of the 2002 ACM symposium on Document engineering 2002 Nov 8 (pp. 26-33). ACM.

Roehm T, Tiarks R, Koschke R, Maalej W. How do professional developers comprehend software?. In Proceedings of the 34th International Conference on Software Engineering 2012 Jun 2 (pp. 255-265). IEEE Press.

:(

No comments, no documentation but 20 tickets



The Guy Who
Wrote This Is Gone

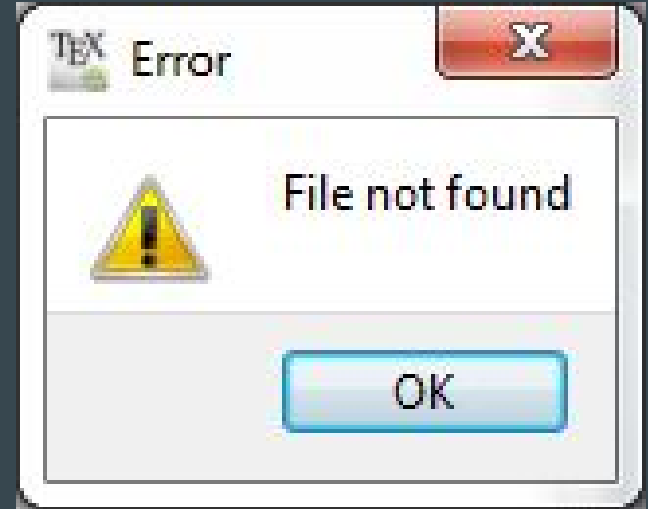
It's running everywhere

O RLY?

FML

Anything else “software/technical documentation”?

- Do different development methodologies (Agile, TDD, etc) improve documentation?
- Is there anything else “cultural” that’s led to documentation improvements?
- Are auto-doc tools a good thing or a bad thing?
- Hour-for-hour, what are the most important documents to write?
 - Any sort of cost-benefit analysis?
- Bueller? Bueller?



YOUR SCIENCE IS BAD



AND YOU SHOULD FEEL BAD



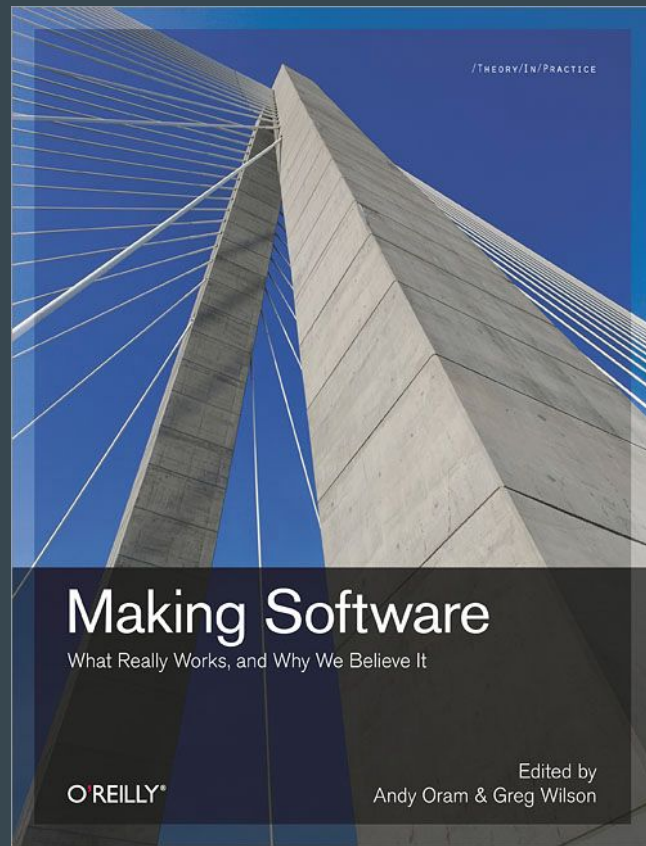
Your developer peers aren't that much farther along

“Computer Science Strong Opinion”

- Greg Wilson

C-style syntax as hard to learn as a
randomly generated language

- Andreas Stefik



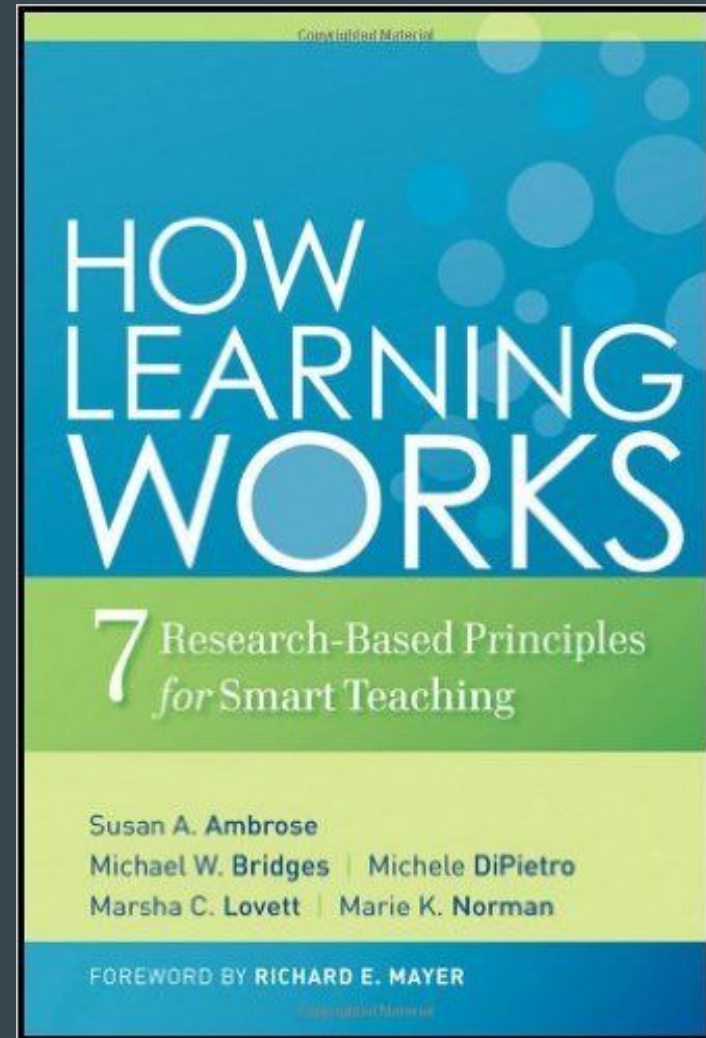


You can do science!



There is a rich literature of educational psych to draw on

- We build knowledge on prior models
- It's helpful for students to see mistakes being made (“expert blind-spot”)
- Practice works better for facts, worked examples for skills
- Subgoals improve performance



Look at study quality

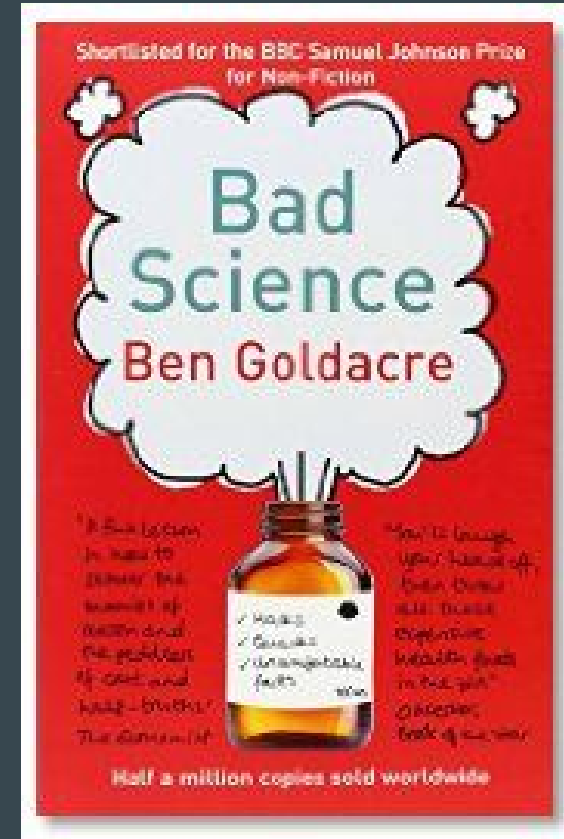
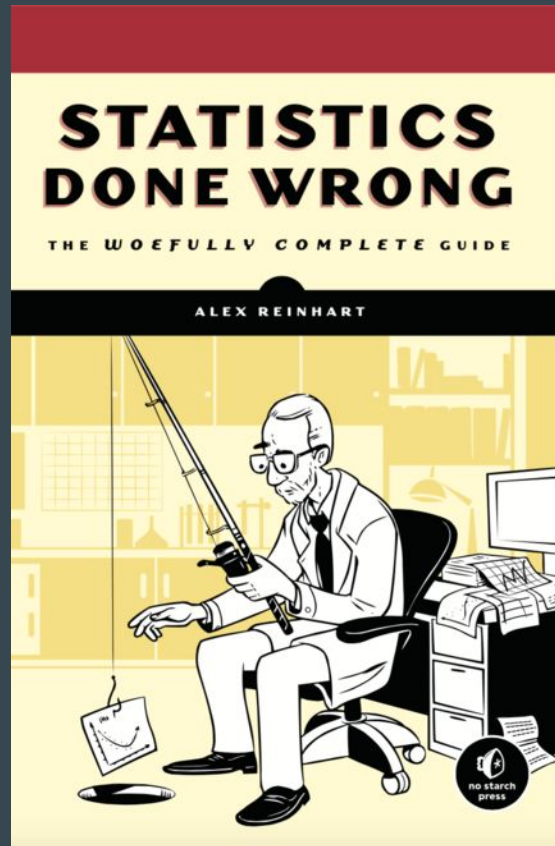
- Be wary of small N
- Be wary of no error-bars
- Is it an important effect?
- Evidence hierarchy:

Meta-analysis

Randomized controlled trial

Survey

Case study



Really, I don't want to be a scold...

Articles in the [NEJM] Journal that were covered by the [New York] Times received a disproportionate number of scientific citations in each of the 10 years after the Journal articles appeared. The effect was strongest in the first year after publication, *when Journal articles publicized by the Times received 72.8 percent more scientific citations than control articles*. This effect was not present for articles published during the strike; articles covered by the Times during this period were no more likely to be cited than those not covered.

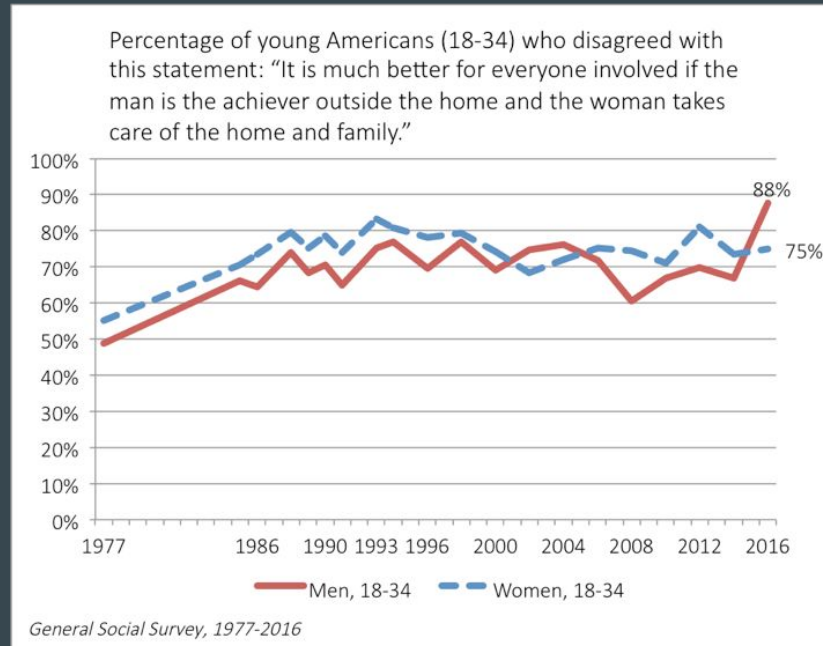
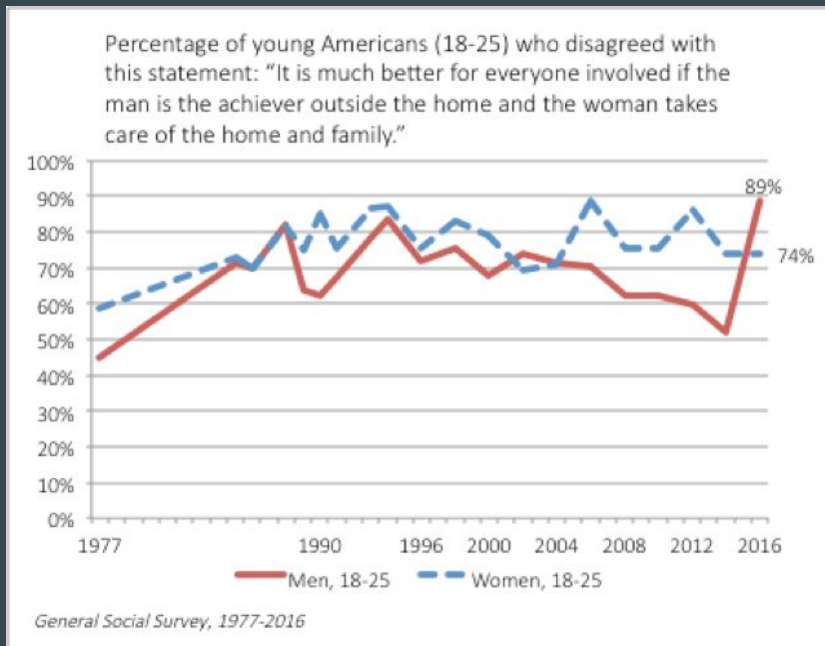
DP Phillips, et al. New England Journal of Medicine, 1991 “Importance of the lay press in the transmission of medical knowledge to the scientific community”, DOI: 10.1056/NEJM199110173251620

In conclusion

- READ ALL THE BOOKS!
- When someone makes a statement about whether something is good or bad, ask for their evidence.
- There *is* value in documentation, and I think there is real opportunity to make our field better.



Look at study quality



<https://thesocietypages.org/socimages/2017/04/03/adventures-in-garbage-millennial-confirmation-bias/>