

# Scientific Writing

Oliver Hohlfeld  
RWTH Aachen University

# How to write a paper?

# Why Care?

- Become a better researcher
- Write better papers
- Publish at higher quality conferences
- Increase scientific impact

# Focus

1. Reading
2. Writing
3. Publishing
4. Reviewing
5. Case Study

# Reading

# How to read a paper

- Top-down. Avoid drowning into details
- **1st pass** (5-10 min):
  - Read: Title, abstract, intro, headings, conclusion, refs
  - Answer: Category, context, correctness, contributions, clarity
- **2nd pass** (<1 hour)
  - Read sections, look at figures but ignore proofs etc.
- **3rd pass** (1-5 hours)
  - Understand paper completely / virtually re-implement it

# Writing

# 1: Every paper tells a story

- What is the “elevator pitch” of your story?
  - Summary that is short enough to give during an elevator ride
- The story is ***not*** what you did, but rather
  - what you show: new ideas, new insights
  - why interesting, important?
- Why is the story of interest to others?
  - universal truths, hot topic, surprises or unexpected results?
- **know your story!**



# 2: Write top down

- computer scientists (and most human beings) think this way!
- state broad themes/ideas first, then go into detail
  - context, context, context
- even when going into detail ... write top down!

# 2: Write top down

- Point first, then explanation
  - We evaluated systems A and B on Google Glass and ImageNet. Figure 3 shows the measured frames per second for both systems. When the frame resolution is low, A achieves similar frames per second as B. As the resolution increases, A gradually outperforms B; it outperforms B by 20% with the maximum resolution.
  - Our evaluation shows system A outperforms system B by up to 20%. The evaluation is based on Google Glass and ImageNet. Figure 3 shows the measured frames per second for both systems. When the frame resolution is low, A achieves similar frames per second as B. As the resolution increases, A gradually outperforms B; it outperforms B by 20% with the maximum resolution.

# 3: Introduction

- If reader not excited by intro, paper is lost
- Recipe:
  - para. 1: motivation: broadly, what is problem area, why important?
  - para. 2: narrow down: what is problem you specifically consider
  - para. 3: “In the paper, we ....”: most crucial paragraph, tell your elevator pitch
  - para. 4: how different/better/relates to other work
  - para. 5: “The remainder of this paper is structured as follows”

# 4: Master the basics of organized writing

- paragraph = ordered set of topically-related sentences
- lead sentence
  - sets context for paragraph
  - might tie to previous paragraph
- sentences in paragraph should have logical narrative flow, relating to theme/topic
- don't mix tenses in descriptive text
- one sentence paragraph: warning!

# 5: Put yourself in place of the reader

- less is more:
  - “I would have sent you less if I had had time”
  - take the time to write less
- readers shouldn't have to work
  - won't “dig” to get story, understand context, results
  - need textual signposts to know where ‘story’ is going, context to know where they are
    - good: “e.g., Having seen that ... let us next develop a model for .... Let Z be ....”
    - bad: “Let Z be”
- what does reader know/not know, want/not want?
  - write for reader, not for yourself

# 6: Put yourself in place of the reader

- page upon page of dense text is **no fun** to read
  - avoid cramped feeling of tiny fonts, small margins
  - create openness with white space: figures, lists
- enough context/information for reader to understand what you write?
  - no one has as much background/content as you
  - no one can read your mind
  - all terms/notation defined?

# 6: Be Concise

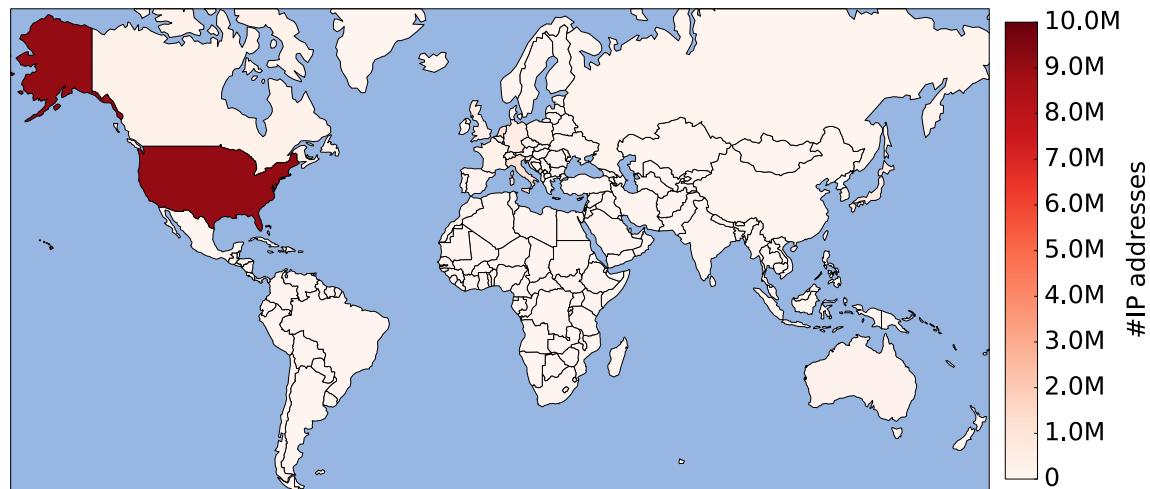
- Focus
  - Remove words, paragraphs, sections, figures, and concepts if they are not absolutely necessary
- No brain dump
- **Don't** have to **tell** readers **everything you know**.  
You tell them enough to appreciate your work

# 7: No one (not even your mother) is as interested in this topic as you

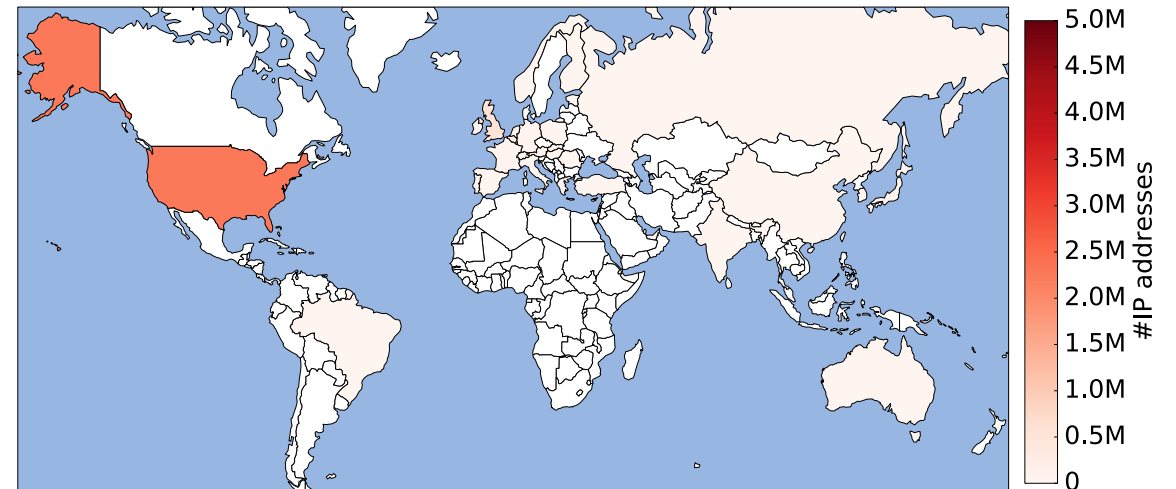
- so you had better be (or appear) interested
- tell readers why they should be interested in your “story”
- don’t overload reader with 40 graphs:
  - think about main points you want to convey with graphs
  - can’t explore entire parameter space
- don’t overload reader with pages of equations
  - put long derivations/proofs in appendix, provide sketch in body of paper



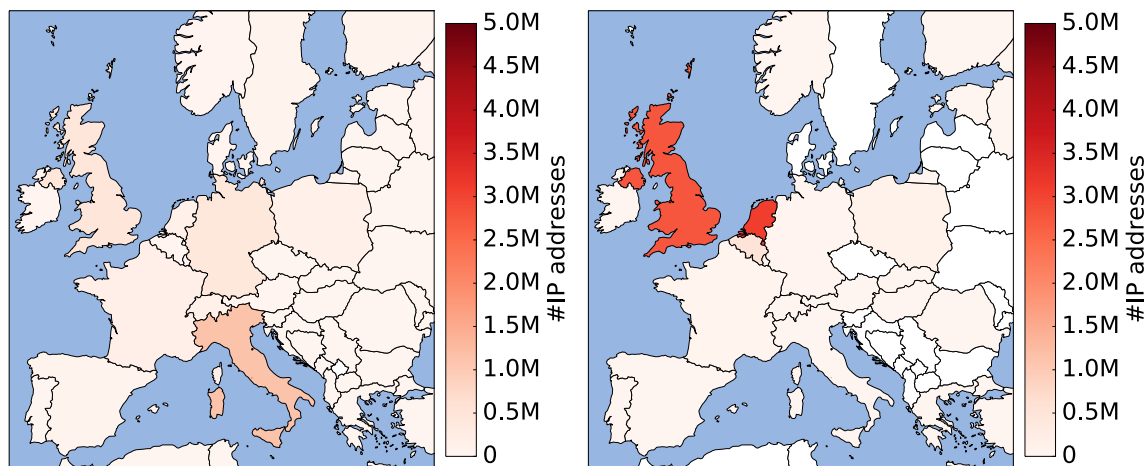
# Plot Bloat



**Figure 6.13** Country distribution of the analyzed hosts' geolocation. The top three countries by number of IP addresses are the United States, Italy, and the Netherlands.

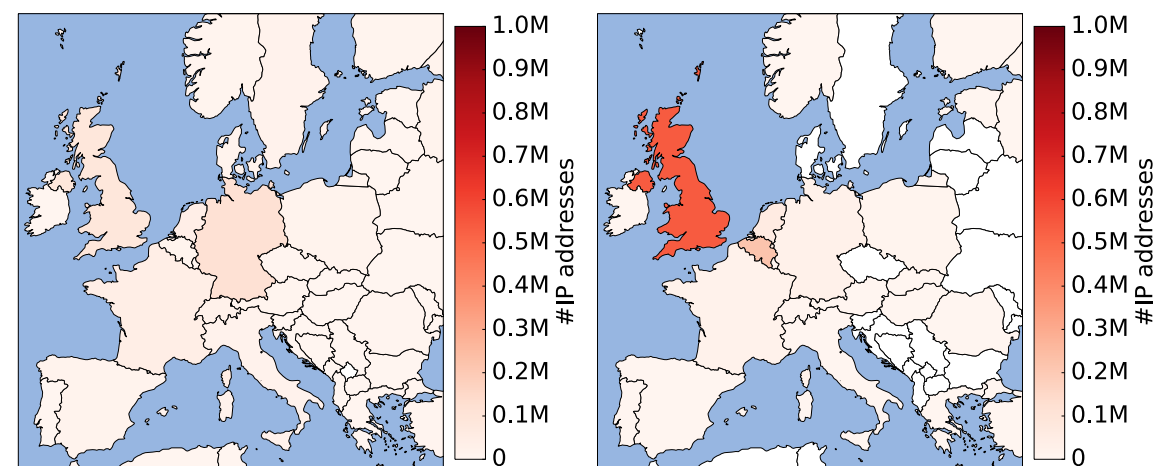


**Figure 6.15** Distribution of the issuer countries per unique certificate. Here, the top three countries are the United States, the United Kingdom, and Belgium.



**(a)** Subject country. Most common here are Italy with about 1.1 million, the United Kingdom with 460,000, and Germany with 440,000 hosts, respectively. **(b)** Issuer country. Most common here are the Netherlands with about 3.1 million, the United Kingdom with 2.8 million, and Belgium with 520,000 hosts, respectively.

**Figure 6.17** Subject and issuer country distribution per host in Europe in Figure 6.17a and Figure 6.17b, respectively. Again, we note that for about 4 million out of 15 million IP addresses no geolocation country could be derived.



**(a)** Subject country. Most common here are Germany with about 120,000 certificates and the United Kingdom with 86,000 certificates. **(b)** Issuer country. Most common here are the United Kingdom with about 540,000 certificates and Belgium with about 220,000 certificates.

**Figure 6.16** Subject and issuer country distribution per certificate in Europe in Figure 6.16a and Figure 6.16b, respectively.

# Unbloat your papers

- For every word, paragraph, plot: ask if it is needed and how the paper would change if you delete it
  - If nothing changes, delete it
- Creating meaningful plots is time-consuming!

# 8: State the results carefully

- clearly state assumptions (see overstate/understate your results)
- experiment/simulation description: enough info to nearly recreate experiment/description
- simulation/measurements:
  - statistical properties of your results (e.g., confidence intervals)
- are results presented representative?
  - or just a corner case that makes the point you want to make

# 9: Don't overstate/understate your results

- overstatement mistake:
  - “We show that X is prevalent in the Internet”
  - “We show that X is better than Y”  
when only actually shown for one/small/limited cases
- understatement mistake: fail to consider broader implications of your work
  - if your result is small, interest will be small
  - “rock the world”

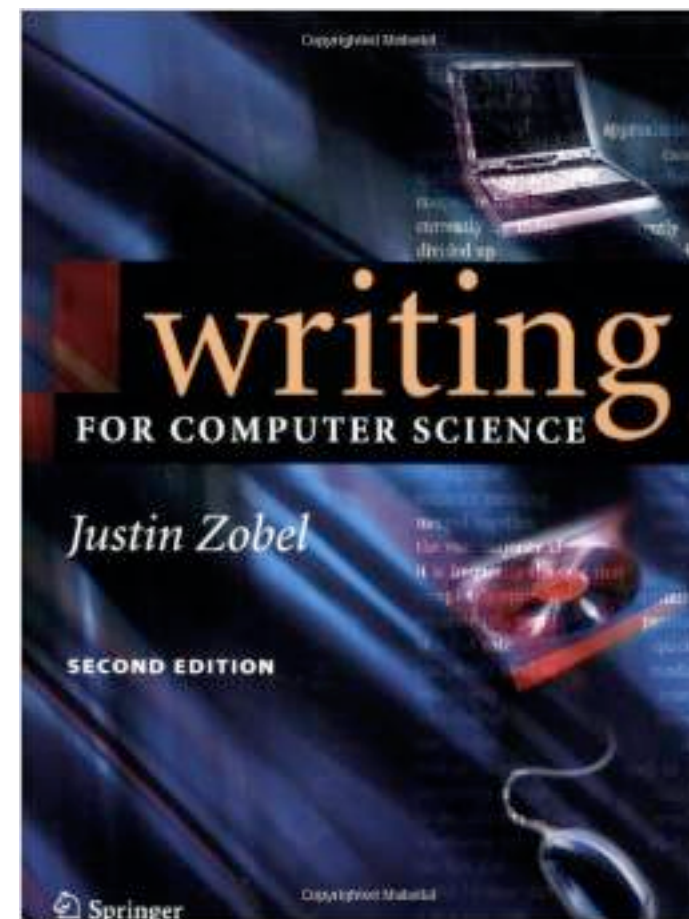
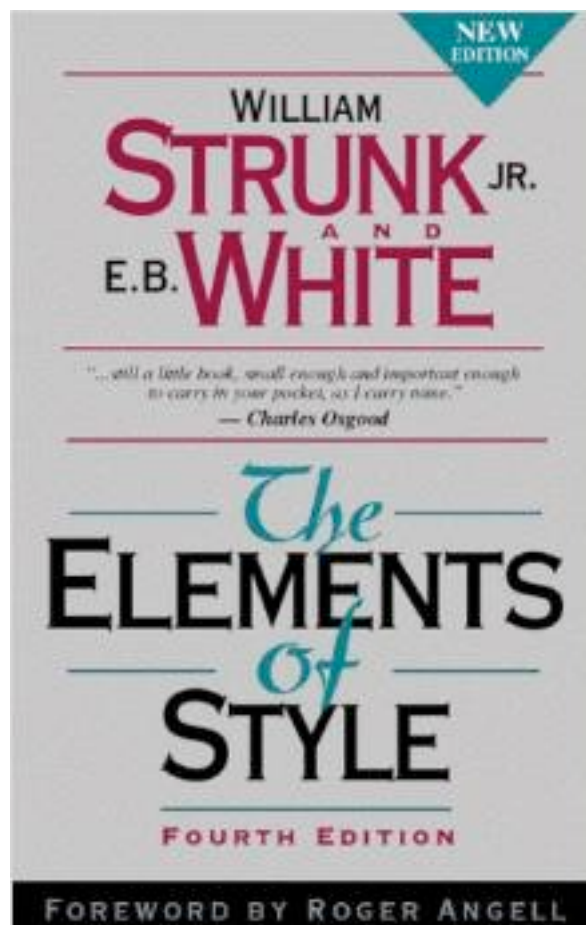
# 10: Study the art of writing

- writing well gives you an “unfair advantage”
- **writing well matters in getting your work published in top venues**
- highly recommended:
  - The Elements of Style, W. Strunk, E.B. White, Macmillan Publishing, 1979
  - Writing for Computer Science: The Art of Effective Communication, Justin Sobel, Springer 1997.
- who do **you** think are the best writers in your area: **study their style**

# 11: Good writing takes times

- give yourself time to reflect, write, review, refine
- give others a chance to read/review and provide feedback
  - get a reader's point of view
  - find a good writer/editor to critique your writing
- starting a paper three days before the deadline, while results are still being generated, is a non-starter

# Further Reading



# Loose emotional connection

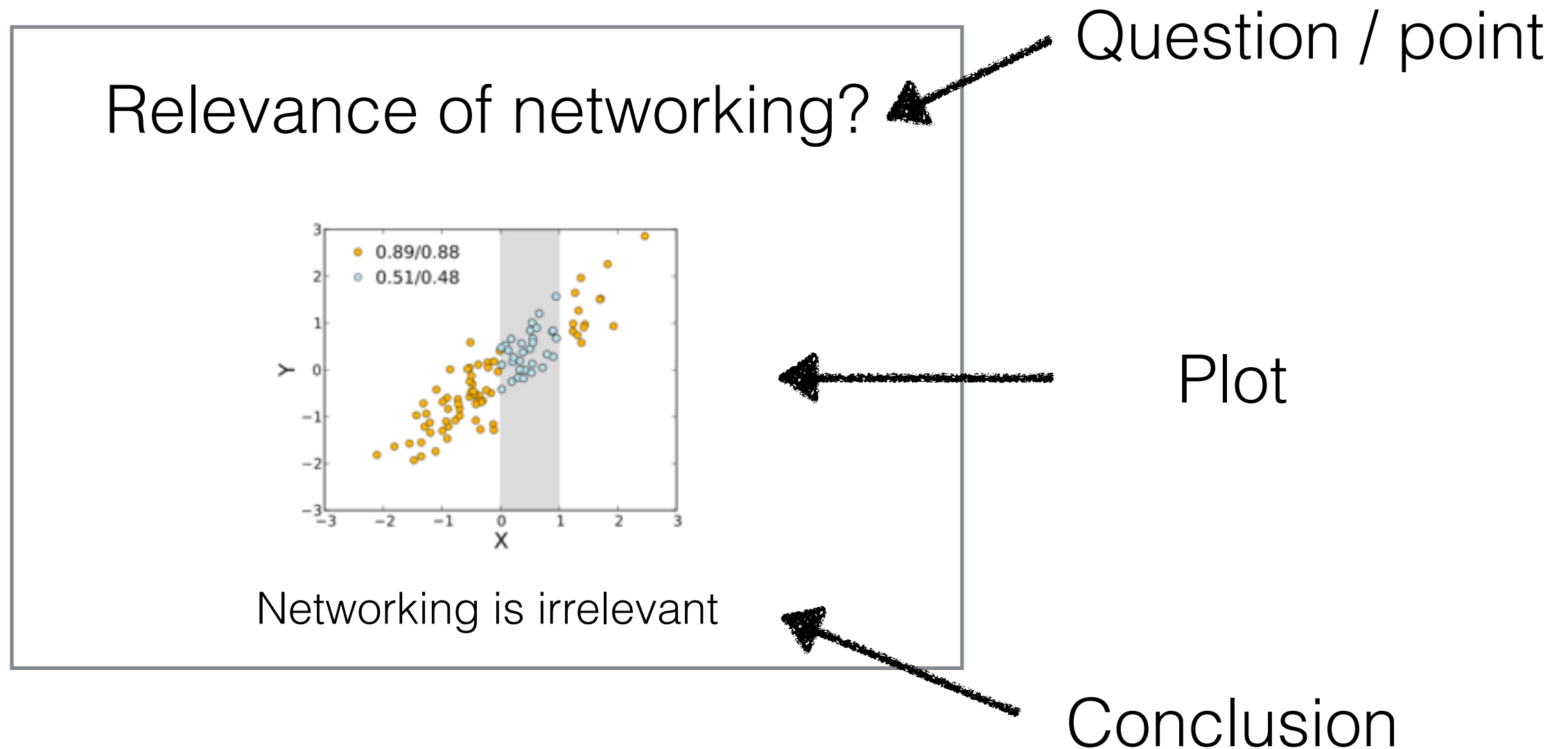
- Many people find it hard to revise / delete text
  - They are proud to have written something
- Get rid of it
  - Revise often
  - Delete unnecessary text
  - Text is only a communication tool



# Advice by Nick Feamster

- “An attempt to write down a concept, idea, or solution can actually help you realize that you don’t know what you are talking about!”
- “writing isn’t just something that happens at when the research is “done”; rather, writing should occur throughout the course of a research project.”
- “A well-written paper with a half-baked idea that people can understand and appreciate will often be accepted to a conference over a paper with a (possibly) stronger result that nobody can understand or appreciate.”

# Craft your story as slides



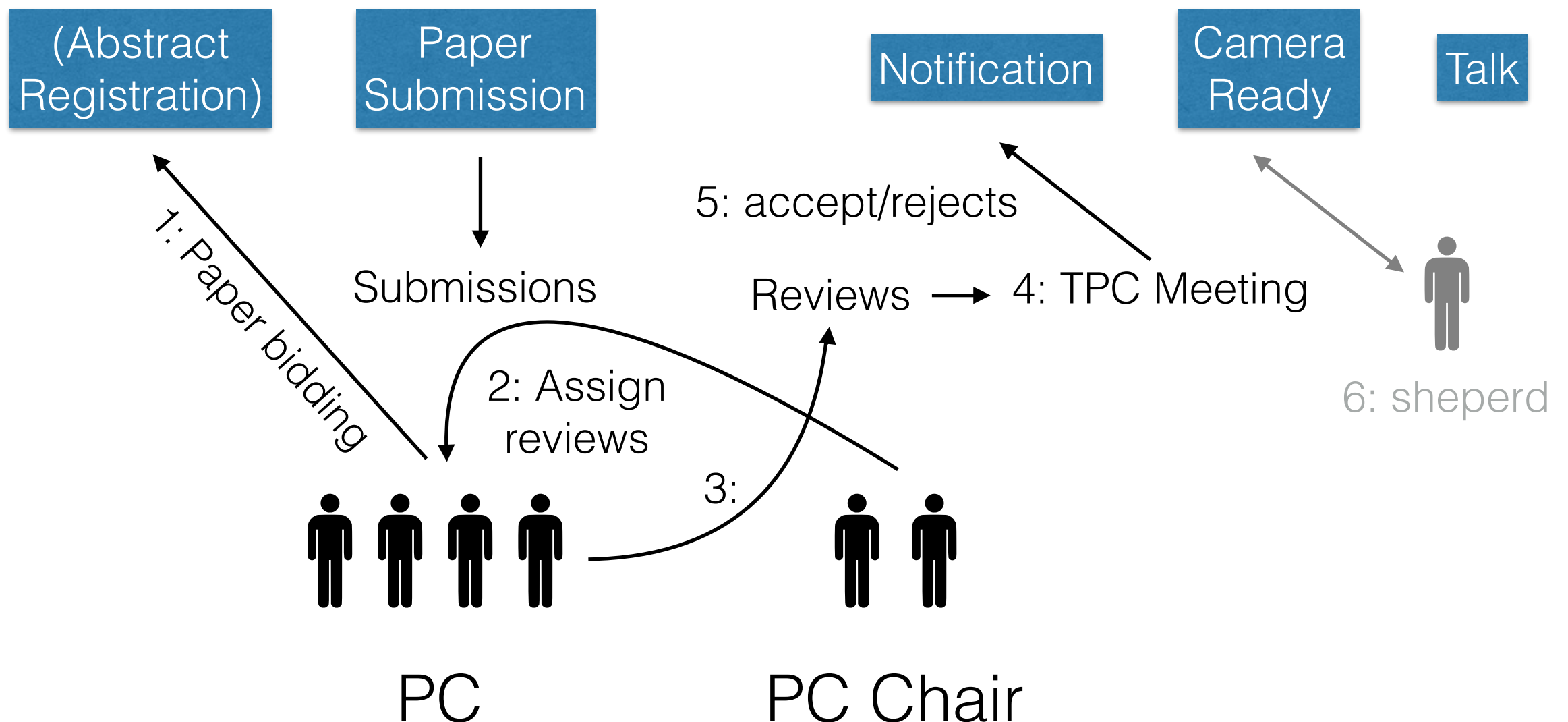
Easy to create / easy to share / simplifies text writing

Behind the Scenes:  
What happens to your  
paper after submission?

# Review Process

Research

Writing



# The Result

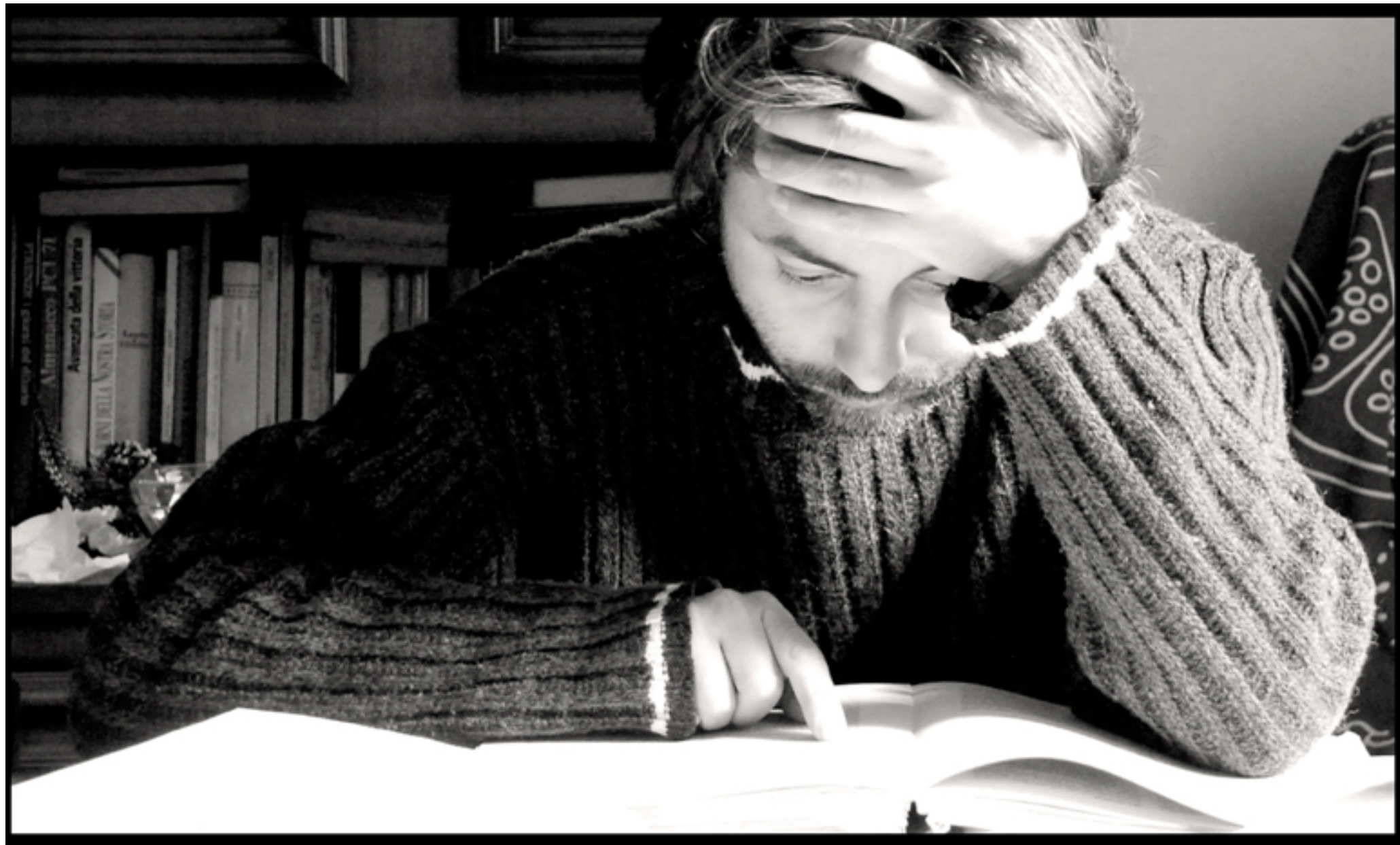
- **Accept** = be happy :)
- **Reject** = my work was bad?
  - Maybe
  - Maybe not

# Influence Factors

- Reviewer selection: who gets to review your paper?
- Mood / background etc. of the reviewer
- TPC Meeting
- ...



Noisy process



# Idealistic reviewer





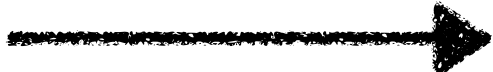
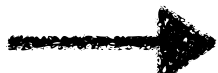
# More realistic reviewer



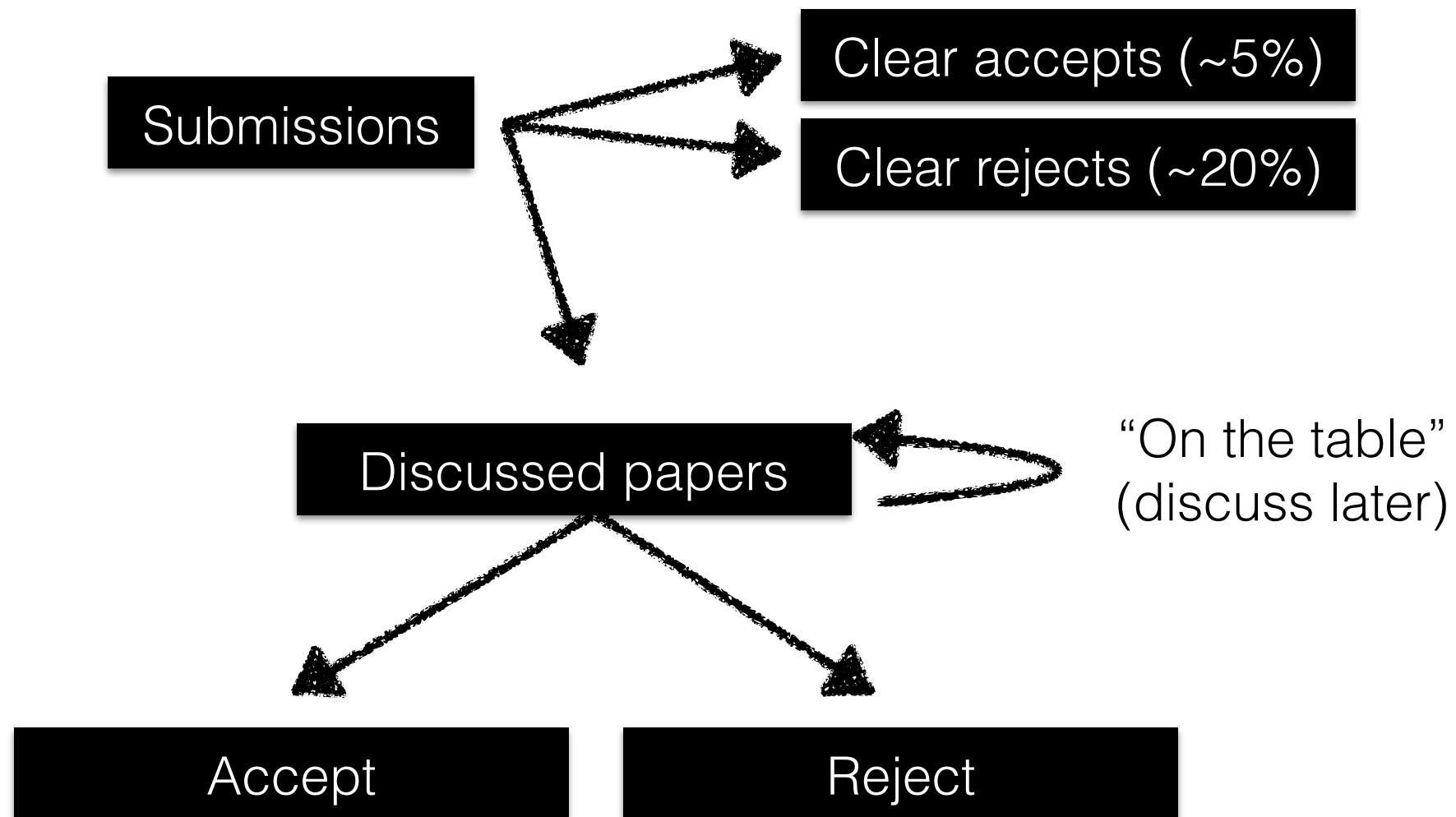
# How your paper is read

- Reviewers skim (e.g., <30 min / paper)
- Low acceptance rates -> bias towards reject
  - Find reasons to reject rather than accept
  - Acceptance / Rejection mood made during the first minute spend with the paper -> convince in the beginning
- They need to fill our review forms, help them


# Review Forms

- Summary
  - Strengths 
  - Weaknesses
  - Detailed comments
  - Comments for TPC 
- Help filling it out:
- Bullet point list of contribution in the intro
  - Some authors prefer it on the first page
- Prevent comments like
- Yet another paper on X
  - Boring read
  - Incorrect

# TPC Meeting



# TPC Meeting

- 
- 9:00 Welcome address by PC chairs: define goals, e.g., #accepted papers
  - 9:15 Everyone is connected to the Wifi
  - 9:30 Normalisation round: 10 papers (very good, normal, and very bad)
  - 10:00 Discuss clear accepts
  - 10:30 Last TPC member joins and excuses for delayed flight
  - 10:33 Start discussion for papers in “to be discussed” pile
  - 12:30 Lunch
  - 13:30 Continue discussion
  - 15:00 Realize that only accepting 5 papers is not enough. Start over
  - 15:30 First TPC members stressfully start thinking about catching their flight
  - 16:30 Still insufficient accepts. Accept papers on top of discussion pile to fill program
  - 17:00 Official end

# TPC Meeting

- Anything can happen if not in clear accept/reject pile
- It depends when your paper gets discussed
  - Beginning: everyone very critical
  - Towards end: accept it, just to get rid of it (and catch flight)
- Discussion time per paper: <4 min
  - <1 min per reviewer
  - Not a lot of time to convince for acceptance
  - You need a champion!