

How to Write a Scientific Paper

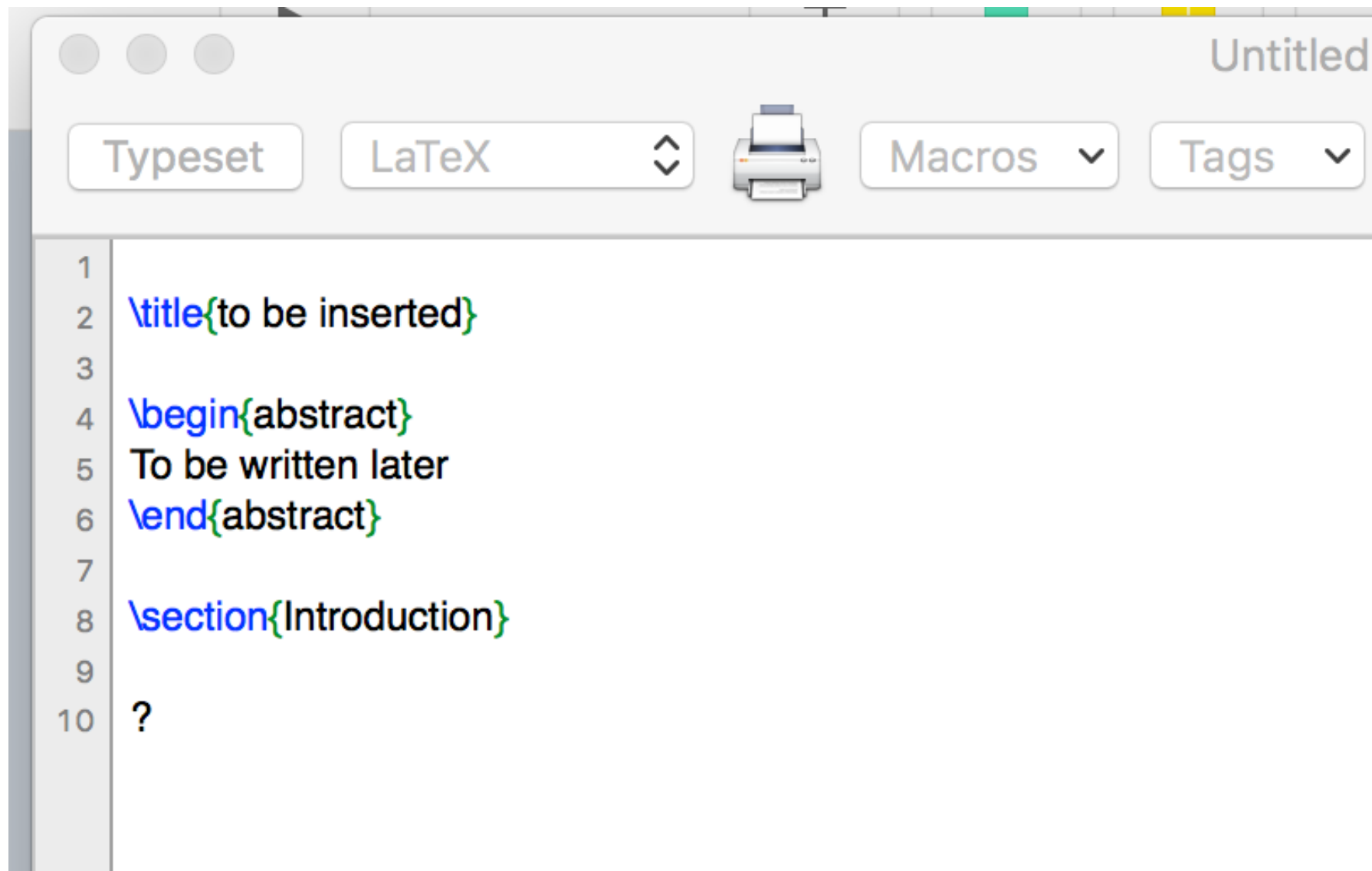
Jari Saramäki

CCS Warm-Up 2022

<https://jarisaramaki.fi/>

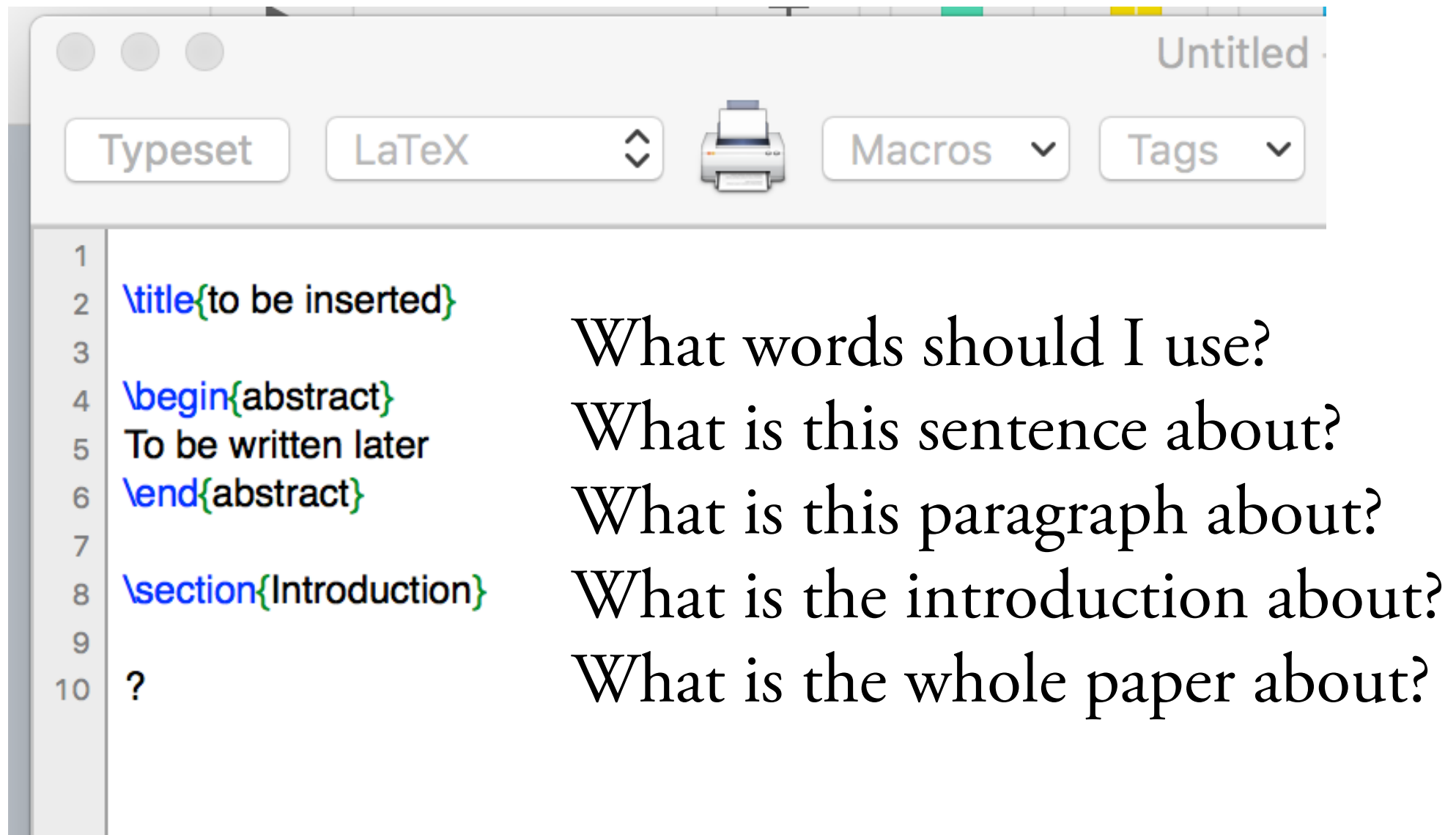


This Is Not The Way!



```
1
2 \title{to be inserted}
3
4 \begin{abstract}
5 To be written later
6 \end{abstract}
7
8 \section{Introduction}
9
10 ?
```

This Is Not The Way!



The screenshot shows a LaTeX editor window titled "Untitled". The window has a toolbar with buttons for "Typeset", "LaTeX", a printer icon, "Macros", and "Tags". Below the toolbar, there is a list of lines of code with corresponding questions:

Line	Code	Question
1		
2	<code>\title{to be inserted}</code>	What words should I use?
3		
4	<code>\begin{abstract}</code>	What is this sentence about?
5	To be written later	
6	<code>\end{abstract}</code>	What is this paragraph about?
7		
8	<code>\section{Introduction}</code>	What is the introduction about?
9		
10	?	What is the whole paper about?

One cannot solve these problems all at once!

Q: How to Write a Scientific Paper?

A: From **top to down**, start with **planning** and **thinking**.

This Is The Way!

I Developing the story

Choose your point

Choose which
results to include

Write the abstract

II Outlining & writing

Outline the sections

Quickly write
a full draft

Revise for
structure & content

Revise for
language & clarity

(Iterate from beginning
if necessary)

III Getting published

Write cover letter
& submit

Revise & resubmit

Rejoice & market!

I Developing the story

Step I: Choose Your Point

Before anything else, you should decide what your paper is about.

You should be able to express this key point in 1-2 sentences.

If you have many results at hand,
pick the most important one
and make it the focus of your paper.

Step I: Choose Your Point

This step is **really important!** Spend time on it!
Discuss it with your colleagues/coauthors/supervisor!

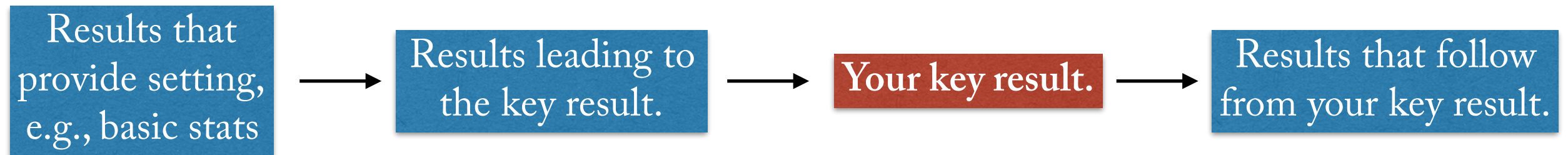
When the point of the paper has crystallized,

- i) the paper is easier to write,
- ii) the paper will be easier to read,
- iii) the paper will be easier to market.

Step 2: Choose what results to include.



Step 2: Choose what results to include.



Step 3: Write the abstract!

It is unconventional to **write the abstract first**,
but once you've tried it,
you know it is the right thing to do.

The abstract is a **miniature version of your story**.

Once you have written it, the rest of the paper
becomes much much easier to write.

Structure of the abstract

Broad

The broad setting: what domain are we talking about here?

Moving closer to your research question: what and why?

What exact problem did you set out to solve?

Narrow

How did you solve it, and what was the outcome?

Immediate consequences of having solved the problem

Broad

The broader implications of your result



Case in point: EVERY Nature abstract

1. One or two sentences providing a basic introduction to the field, comprehensible to a scientist in any discipline.

2. Two to three sentences of more detailed background, comprehensible to scientists in related disciplines

3. One sentence clearly stating the general problem being addressed by this particular study

Networks of coupled dynamical systems have been used to model biological oscillators, Josephson junction arrays, excitable media, neural networks, spatial games, genetic control networks and many other self-organizing systems. Ordinarily, the connection topology is assumed to be either completely regular or completely random. But many biological, technological and social networks lie somewhere between these two extremes. Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. We call them 'small-world' networks, by analogy with the small-world phenomenon (popularly known as six degrees of separation). The neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors are shown to be small-world networks. Models of dynamical systems with small-world coupling display enhanced signal-propagation speed, computational power, and synchronizability. In particular, infectious diseases spread more easily in small-world networks than in regular lattices.

4. One sentence summarising the main result (with the words "here we show" or their equivalent).

5. Two or three sentences explaining what the main result reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge

6. One or two sentences to put the results into a more general context. Two or three sentences to provide a broader perspective, readily comprehensible to a scientist in any discipline

Watts & Strogatz *Nature* **393** (1998)

The title of the paper

- Many conflicting requirements!
- Short better than long, but informative better than uninformative — if you can compress the result into a short title, perfect.
- Avoid the use of jargon or concepts that the average reader wouldn't know in the title!
- But do use the right keywords for searchability!
- Catchy is good, but ridiculous isn't... there's a fine line

Example titles from Nature, Jan 19th, 2022

Microwave background temperature at a redshift of 6.34 from H₂O absorption

Novel flight style and light wings boost flight performance of tiny beetles

Malaria protection due to sickle haemoglobin depends on parasite genotype

An autoimmune stem-like CD8 T cell population drives type 1 diabetes

Domain-wall dynamics in Bose–Einstein condensates with synthetic gauge fields

Comments? Which work well, why?

II Outlining & writing

Step 4: Outline the paper

Introduction

Methods

Results

Discussion

- Next, sketch & plan what goes into each section.
- First, roughly sketch the contents of the sections, then refine to the point where you have a paragraph-level plan

Step 4a: The Introduction

- What is the context?
- What leads to your question?
- What IS your question?
- How did you approach it?
- What did you find out?

Much like the abstract, but longer, and emphasizing context (the top of the hourglass).

Introduction: 4-paragraph template

- *The first paragraph* provides context and background for your research by introducing the problem area and the knowledge gap that has led you to ask your research question, citing key papers.
- *The second paragraph* zooms in on your particular research problem, forming a funnel from the broader motivation to the exact question. At the end of this paragraph, you should explicitly state the research question that your paper addresses.
- At the beginning of *the third paragraph*, the point of view moves from what others have done to what you have done. The third paragraph describes how you have approached the research question.
- *The fourth paragraph* moves from your approach to your findings. It reveals the outcome of your work and briefly summarises your results.

Step 4b: Methods

After reading Methods,
the reader should be
able to reproduce
your findings.

The scientific method: any result can be verified or falsified!

Methods

- Use schematics and figures to provide an overview before presenting the details!
- Motivate your choices of methods! State your reasoning. Clarity is key!
- If possible, talk about limitations already at this stage.
- Materials, data, etc are also discussed in this section.
- Code, data, etc: make available using repositories, give pointers to those.

Methods

How **not** to begin a Methods section

Method X entails
computing Z
and then...

Way better, isn't this?

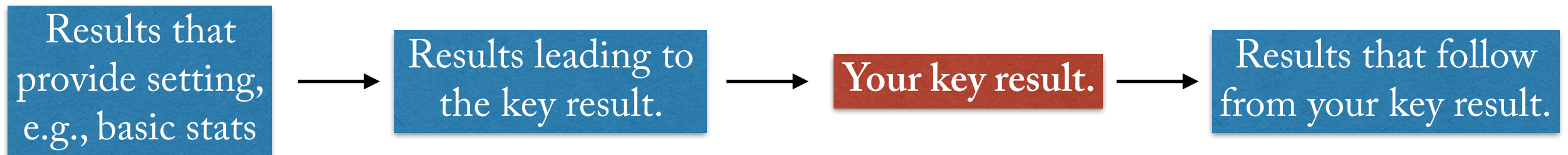
To do A, we need a
method.
We have chosen X
because...
Method X entails...

Reading the Methods section is always hard. Be gentle
to your reader and guide her hand.

Step 4c: Results

Think of the Results section as a court case in some TV drama (but build your case with honesty and integrity).

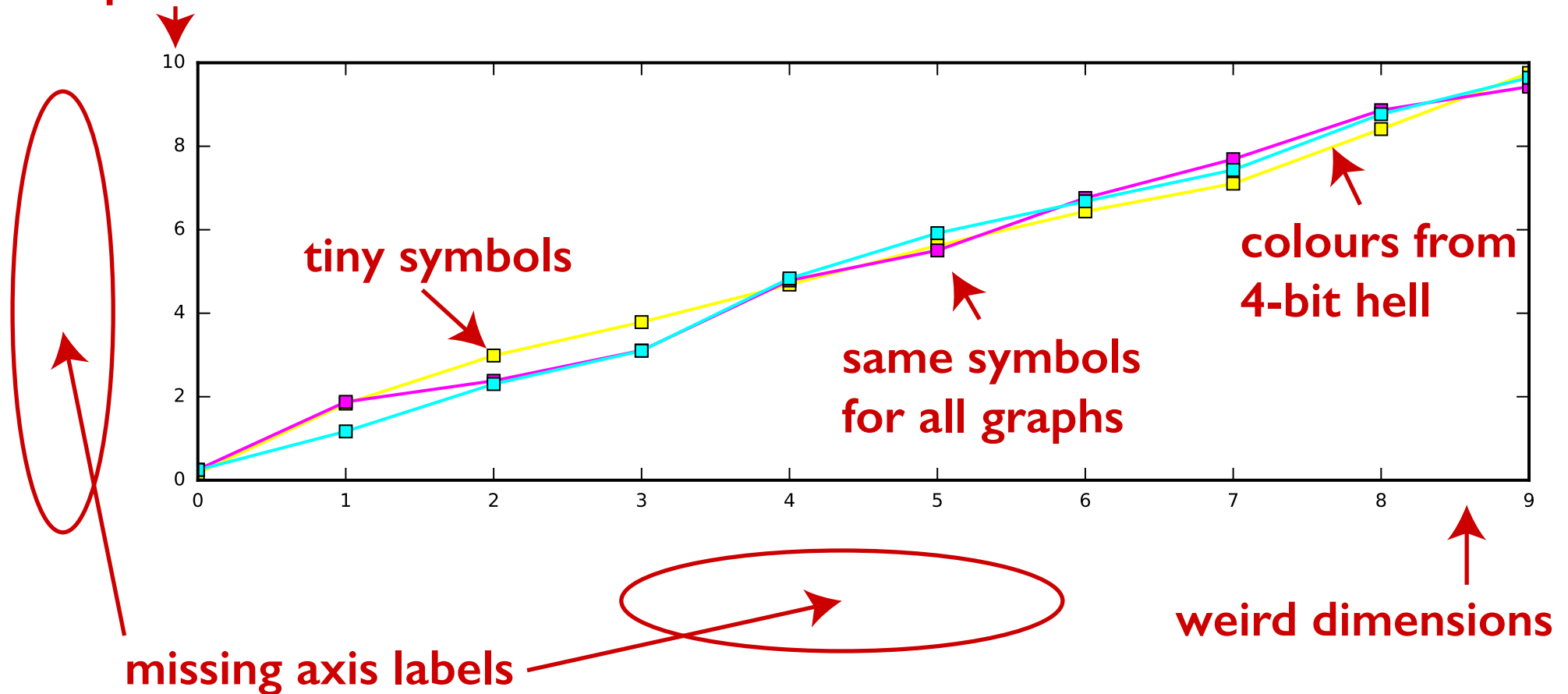
Results



- First, sketch your story with **figures** and **their captions**!
- **Pro tips for figures:**
 - **Caption:** what the figure means, not “Fig 1: X as a function of Y”
 - **Axes:** 1) always label them, even in the 1st draft for your supervisor, 2) do not use tiny fonts for labels
 - **Colors:** one color = one meaning, throughout the paper!
 - **Align** panels and other parts of composite figures
 - **Tools:** Learn to use vector graphics tools — **Adobe Illustrator, Inkscape, or similar** — for post-processing

Figures: do not do this!

nanopixel-sized fonts



Results: interpretation

- Three levels of knowledge:
 1. **Pure data:** e.g. the data points in my plot
 2. **Facts that everyone would agree on:** my plot shows that $a(x)$ is mostly above $b(x)$
 3. **Interpretations:** $a(x)$ being above $b(x)$ is in line with the hypothesis that...
- Always make it clear whether you talk about data, facts, or interpretation.

Tip for results: use subsection titles

2 Results

2.1 Screen-on events can be used for chronotype assessment

We use time-stamped data on “screen-on” events from the smartphone data-collection apps to assign a behavioral chronotype to each participant. Whenever the participant uses the smartphone, from making a call to checking the time, the phone’s screen is turned on, and the data-collection app records the time of this event. We use the frequency of these events as a statistical proxy for the daily activity rhythm of the participant, since frequent screen-on events tell that the participant is awake, and night-time event frequencies are

2.2 Owls have larger personal networks than larks

We first construct the personal networks of all participants based on both call and text data. For this we use each (hashed and anonymized) phone number that the participant communicates with (through calls and text messages) as a proxy of a social relationship. In this network, each individual is a node and communication events (calls and text messages) between people are the links. The degree of a node (the personal network size)

2.3 Owls are more central than larks in the social network of participants

In order to study the network centrality of each participant, we constructed the social network of participating students, so that two individuals i and j are connected with an

one result per
subsection;
the title *is*
the result

Step 4d: Discussion

- Wrap up and condense results and their meaning, and discuss future directions
- Discuss limitations but don't dwell on them, and don't finish with them — rather, the point of view should be that the limitations provide new, interesting and important research questions
- Discuss impact: what follows from your work now? What future doors does it open? Do you have future research suggestions?
- [Note: in some fields, results are interpreted in Discussion, not in Results. For those fields, I am here talking about General discussion].

Discussion: a template

- *First paragraph:* Recap of the knowledge gap and the research question
- *2nd-Nth paragraphs:* Summary your results, in the order leading to your main point, discussing limitations/alternative approaches/etc while you go, as well as new questions that have now opened
- *Final paragraph:* conclusion and take-home message — “To conclude, we have shown that ... Because of our result, it is now possible to ... Our result also highlights the need to ...”
- *Always end on a high note! Your last sentence should be strong!*

Step 5: Now Go and Write a Crappy First Draft!

- *“To write is human, to edit is divine”* -Stephen King
- Productive writers write crappy drafts quickly
- Only then do they edit them to perfection.
- Be a productive writer. Quick and dirty first, edit later.
Draft your entire paper before polishing anything.

Paragraphs

- **One paragraph is about one thing only!**
- 1-2 first sentences define the topic of the paragraph.
- Last 1-2 sentences conclude the paragraph and lead to the next paragraph.
- Do not stray off the path! No tangential sentences! Split long paragraphs ruthlessly!

Sentences

- The beginning of a sentence is its **setup** (“Topic position”). It tells what the sentence is **about**.
- The end of a sentence is the **resolution** (“Stress position”). This is what the reader focuses on, so **place important material at the end**.
- It is easier to read text where **sentences begin with familiar words/material and end with new things**.
- The beginning of the sentence also provides a **link to the previous sentence**.

Sentences

- Keep your **subject and verb close!** (Writing short sentences helps).
- Avoid the **passive voice!** (“The passive voice should be avoided!”)
- How to spot passive voice? See if you can insert “**by zombies**” after the verb without violating grammar.

(this ingenious tip comes from Rebecca Johnson, <https://twitter.com/johnsonr>)

Sentences

- Use **verbs** that **describe** actions!
 - Very clunky: There is a dependence between X and Y
 - Still clunky: X is dependent on Y
 - Better, isn't it: X depends on Y
 - Best by far: X grows linearly with Y

Step 6: Edit your draft

- Do at least two passes of edits
- On the first pass, focus on flow, clarity, and structure.
 - Add signposts wherever needed: words and sentences to guide the reader
 - “We will first present the general summary statistics for our data, and then move on to...”
- On the second-Nth pass, focus on details
 - One paragraph = one point. Split long paragraphs
 - Shorten sentences and cut out words that are not needed. Aim at cutting 10-30%!
 - Remove jargon, remove passive voice.
 - **Aim at simple, short, easy-to-understand sentences with active verbs.**

III Publish & market

Step 7: Write a cover letter (if needed) and submit

- Be succinct — get to the point quickly.
- Explain the key point of the paper and why it is important (if you have done your homework, this should now be easy).
- Explain why the paper is a good match for this journal/conference.

Step 8: Revise and resubmit

- When receiving critical comments, relax. Breathe. Don't be hasty.
- Don't let yourself get into a fight-or-flight response.
- Let the comments simmer for a while, then reread with a critical eye.

Step 8: Revise and resubmit

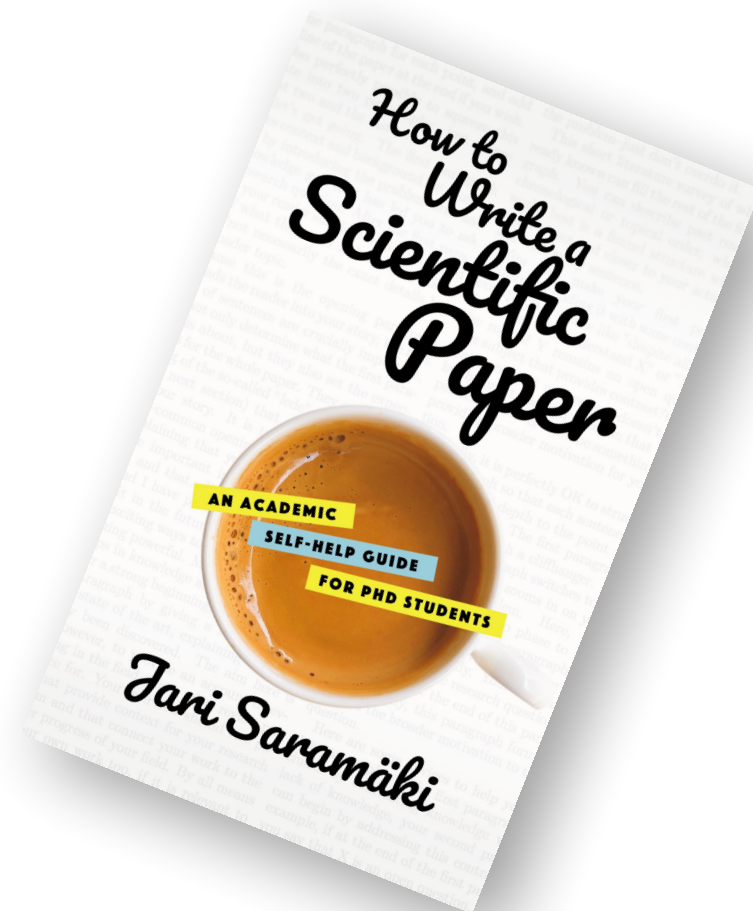
- Sort the comments into doable, maybe, and WTF categories.
- Doable: clearly improve the paper/correct mistakes
- Maybe: borderline, reviewer misunderstanding something, out-of-scope, etc.
- WTF: the rest. The toxic waste bin. Unfortunately, this category exists and is common.

- Doable comments: just do them.
- The maybe category:
 - The key is to make the referee feel heard. Change something in response to each comment—even a very small, cosmetic thing.
 - If the referee has misunderstood something, revise the text so that other readers won't. Think of the referee as a typical reader.
 - However, if the referee wants something that e.g. is beyond the scope of your work, politely say no (and explain why).
- The WTF category:
 - If you do not understand what the referee wants, say so, or just interpret their comments in your chosen way
 - Say no to inappropriate requests, like to cite N off-topic papers by someone (=the reviewer)
 - Fight, if need be (but politely, and writing the response so that it is also clearly intended for the eyes of the editor).

Step 9: ACCEPTED!!

- Congratulations!
- Still work to do: the proofs will arrive one day, usually with a 24-hour turnaround time :-)
- Market your work! Conferences, Twitter, etc
- Twitter: a 1/N style thread with actual results works well.
- Press release? See <https://jarisaramaki.fi/2019/09/19/how-to-write-a-press-release-that-journalists-want-to-publish/>

The End



For blog posts on writing, see <https://jarisaramaki.fi/>
For my writing book, see <https://books2read.com/howtowriteapaper>
(paper version available from Amazon)

Further recommended reading: Joshua Schimel — Writing Science