≜UCL

How to write a good scientific paper

Dr Simon J.D. Prince
Department of Computer Science
University College London

17/01/07

≜UCL

PLAN OF TALK

•! Why should I publish papers?
•! When should I publish papers?
•! Where should I publish papers?
•! Parts of a published paper
•! Style
•! Tips for writing the paper
•! The review process

≜UCL

WHY WRITE A RESEARCH PAPER?

–! To contribute to the literature
–! For the sake of your career
–! To get feedback from the community on your work
–! To make your supervisor happy(!)
–! To stake a public claim to your ideas
–! To help bulletproof your PhD

≜UCL

PLAN OF TALK

•! Why should I publish papers?
•! When should I publish papers?
•! Where should I publish papers?
•! Parts of a published paper
•! Style
•! Tips for writing the paper
•! The review process

≜UCL

CONTRIBUTION

•! Depends on the standard of the particular journal/ conference and the norms of the field.
•! If it is only an abstract/poster then realistically the cost of you attending the conference is very little.
•! You have to make a decision about the standard of your own publications – if you don't go to any conferences you will probably not get better
•! Try not to get involved in disproving things that everyone knows are wrong already

---

≜UCL

PUBLISHING AND IP

-can't get a patent after you have made the first public declaration (1 year after in US)

-5 minute phone call to UCL business might be worth making

DUPLICATION

-can't publish same ideas twice

-can't dual submit to two conferences

-perfectly acceptable to submit to conference and journal simultaneously

-somewhat acceptable to publish variants of same paper to two different fields

PLAGIARISM

•! DO NOT claim other people's ideas as your own

•! DO NOT try to publish things that have already been done

•! DO NOT take ideas from inappropriate sources (reviewing etc.)

PLAN OF TALK

•! Why should I publish papers?
•! When should I publish papers?
•! Where should I publish papers?
•! Parts of a published paper
•! Style
•! Tips for writing the paper
•! The review process

WHERE SHOULD I PUBLISH?

•! journal vs. conference
•! length of format
•! impact factor
•! ease of publication
•! national vs. international
•! visa issues
•! special issues of Journals
•! smaller meetings w/ associated paper

POSTER vs. PAPER

•! Paper usually gets more exposure
•! Poster gets more interaction with members of the community – good for networking
•! Poster probably better if you don't have something world-beating or if you are not a great speaker(yet).

PLAN OF TALK

•! Why should I publish papers?
•! When should I publish papers?
•! Where should I publish papers?
•! Parts of a published paper
•! Style
•! Tips for writing the paper
•! The review process

COMPOSITION OF RESEARCH PAPER

MAIN COMPONENTS

- •! Title
- •! Abstract
- •! Introduction
- •! Methods
- •! Results
- •! Discussion
- •! References

OTHER COMPONENTS

- •! Headings
- •! Figures
- •! Tables
- •! Videos
- •! Appendices
- •! Acknowledgements

COMPOSITION OF RESEARCH PAPER

Q. Why this structure?

A.! Organises information to clearly show
  –! The problem being investigated
  –! What you did
  –! What you found
  –! What you think it means

INTRODUCTION

- •! Provides relevant background information to the process / feature / problem being investigated
- •! Usually includes an outline of the study area or relevant previously established aspects of the subject being studies
- •! Should end with a clear statement describing the aims of the work you are presenting
- •! DO NOT preview your results or conclusions

INTRODUCTION

What is the pitch? Should have the general form:

"Something is badly lacking from the literature, I am providing it"

INTRODUCTION

- •! The aims you define must be addressed in the Discussion/ Conclusion
- •! You also need to make clear WHY you are doing the work and why it is important
- •! This is where you SELL the paper – and one of the hardest parts to write well

---

Generic Introduction

1. What is the global application/scope/benefit of the work?
2. What is the specific sub-problem that you are trying to solve?
3. What have other people done?
4. In what ways are these deficient so that more research is required?
5. (optional) What recent developments have prompted the research in this paper?
6. The PITCH: In this paper you are going to move towards the goal in 1. by solving the problem in 2. in a way that avoids the problems in 4. using the developments in 5.
7. This is the structure of the rest of the paper…

METHODS

- •! A full record of exactly what you did
- •! Someone else should be able to use your methods to recreate your work exactly
- •! Details are really important here – exactly which parameters did you use?
- •! Will NOT include any results

RESULTS

REPORT JUST THE FACTS:

- •! Clear report of the data you obtained and processed summaries thereof
- •! Most results usually presented in Figures and Tables
- •! Also describe in words the key features, patterns and differences that you want to highlight

RESULTS

Results of statistical analysis should be presented in detail:
- •! Confidence limits
- •! Tests applied
- •! Numerical value of test statistic
- •! Degrees of freedom
- •! Significance value

Remember the point of this is to present your findings honestly…

RESULTS

- •! DO NOT make any interpretative statements (these belong in the discussion)

e.g. "We found pattern XYZ, which suggests that ABC might be occurring"

Your data should be kept separate from the interpretation – the interpretation may depend on current theories, but the data will always be true…

DISCUSSION

What does it all mean?

- •! Interpret results in the context of previous work
- •! May open by summarizing the key results & addressing aims defined in the introduction.

DISCUSSION

- •! Attempts to explain the pattern of results
- •! DO NOT introduce new evidence here
- •! DO cite results from the literature that are relevant to your results and argument
- •! State all the assumptions that you are making in the interpretation of your results: even if a reviewer does not accept your assumptions, they cannot reject your logic

DISCUSSION

- •! Should outline the implications of your findings
- •! May include a paragraph or section describing any limitations of the methods used
- •! Can briefly describe avenues for future investigation based on present work (claiming territory!)

CONCLUSION

- •! Some journals require a separate conclusion
- •! Conclusions should be proven from yhour data
- •! Summarize key findings and implications
- •! DO NOT introduce any new evidence from your work or new points from other published work.

ABSTRACT

- •! Typically 100-250 words
- •! Must be concise BUT sufficiently detailed to adequately describe your own work on its own
- •! Mirrors the structure and key points of the paper
- •! WRITE IT LAST?

ABSTRACT COMPONENTS

Abstracts have a standard structure that you need to recognize and use

(1)!Often open with a general statement defining the subject area
(2)!Summarize methods in next couple of sentences
(3)!Summarize key features of your results
(4)!Summarize your interpretation of these results
(5)!Finish with a statement of your main conclusions

ABSTRACT

- •! So the Abstract mirrors the structure and content of the paper
- •! You should be able to recognize the Introduction, Methods, Results and Discussion within it
- •! Sometimes Methods and Results may be interleaved for brevity or greater clarity
- •! In general, do not include references unless it is really key to understanding motivation

TITLE

- •! Write title last of all…
- •! Keep it short!
- •! Needs to be informative and precise

In general (at least for computer science) short is good: "Neural Networks", "Video Epitomes", "Style Machines", "Eigen-Faces"

FIGURES

The first thing everyone sees. If they look professional then your referees start with the preconception that you know what you are doing.

Someone should be able to understand the whole paper by just looking at the Figures and reading the legends.

FIGURES

- •! Journals require each figure to be presented on a separate sheet as a full-page figure
- •! Figure captions are usually listed at the end of the main text, followed by the pages with figures
- •! Insert only place-holder in main text ("Figure 1 about here")
- •! Make sure that text in Figures is of a reasonable size

FIGURES

- •! Composite figures bring data together to allow the reader to make clear comparisons
- •! Usually preferable to several smaller figures
- •! More efficient way of presenting results
- •! Each individual figure forms a panel (1A,B,C,D) etc.

FIGURES

- •! Axes should be labelled with correct units
- •! Avoid >5 lines or sets of data per graph or panel
- •! Only join data points with solid lines if there is a continuous process of change between them
- •! Different symbols for different datasets – should be defined in the caption
- •! If colour is vital to understanding then preface legend with "(Best viewed in colour)".

FIGURE CAPTIONS

- •! Figure captions should allow readers to interpret the figure fully without referring to the main text
- •! Insufficient: "Figure 7. Results of Experiment 1"
- •! Better:

"Figure 7. Plot of corner accuracy as a function of RMS contrast of 15 × 15 pixel region around corner. Accuracy estimated from ideal square corner contaminated with independent Gaussian noise (s.d. = 5) at each pixel. The variation in corner estimation asymptotes approximately at 0.16 pixels. We use these data as estimates of our measurement accuracy."

AUTHORS AND ACKNOWLEDGEMENTS

- •! Different conventions for author order – learn what is appropriate in your field or you will offend someone.

- •! General advice, adding authors is cheap. Not worth falling out with a colleague over.

- •! People who helped but made no intellectual contributions go in acknowledgements.

- •! Acknowledgements to grant funding body

APPENDICES AND VIDEOS

- •! Appendices generally not encouraged – the referees generally don't read them, so don't consider it free space.
- •! Should only put in proofs and mathematical details that would disrupt the flow of the main text
- •! Videos strongly recommended at some conferences, and are what the referees consider first.
- •! Use MPG format as this can be viewed by everyone.

REFERENCING

- •! Any factual statements that are not based on your own results should be backed up by cited references
- •! Makes it clear (1) what you are saying yourself and (2) what you think other people have said.
- •! Cited references should be primary peer-reviewed sources where possible.
- •! Be careful if you criticize other peoples work. Do not start an academic feud.

REFERENCING

- •! Different journals have different rules for citing references
- •! IEEE style is to have references in alphabetical order and cite with number in brackets e.g. [1].
  - –!1 author: Prince [2]
  - –!2 authors: Prince and Stroud [2]
  - –!More than 2 authors: Prince et al. [2]

REFERENCING

- •! Cited references listed in reference list
  - –!Alphabetically
  - –!Then single author papers before multiple authors
  - –!Then chronologically
- •! Different journals will have strict rules for whether official journal abbreviations should be used, where comma's and brackets should go, whether full-stops should be used after author initials etc.

HOW MANY REFERENCES?

- •! Regardless of what you think, it had better be roughly the same amount as everyone else in a conference paper if you don't want to stand out to the referees
- •! Make sure you reference the canonical papers in your area as the referees will look for these
- •! If possible find out who reviewed for the conference last year as they will probably review again. Reference them.
- •! Exhibit common sense: if the conference is national, it will be reviewed at a national level.

PLAN OF TALK

- •! Why should I publish papers?
- •! When should I publish papers?
- •! Where should I publish papers?
- •! Parts of a published paper
- •! Style
- •! Tips for writing the paper
- •! The review process

STYLE TIPS

"Vigorous writing is concise. A sentence should contain no unnecessary words, a paragraph no unnecessary sentences, for the same reason that a drawing should have no unnecessary lines and a machine no unnecessary parts. This requires not that the writer make all his sentences short, or that he avoid all detail and treat his subject only in outline, but that every word tell."

(From Strunk and White)

STYLE – OVERVIEW

- •! Simple is better than complex
- •! Precise is better than vague
- •! Short sentences are better than long ones
- •! Passive vs. active
- •! Cross-reference Figures/ Equations/ Sections
- •! Label all equations

STYLE #1 – Simple, but not too formal or casual

•! Don't artificially formalize writing style.

Bad: "Problem X is clearly a critical area that impacts our research agenda and hypothesis. Our ideas about problem X are embryonic and still evolving, and doubtless our ongoing work in this area will quickly yield fruitful results."

Better: "We recognize that problem X is central to our agenda, but we have only begun to investigate it."

•! Likewise, do not make the style too casual – if it sound a bit casual to you, it will sound REALLY casual to trained scientists. "Our results are somewhat pleasing…"

REMEMBER: The aim is to communicate your ideas, not to make yourself look clever!

- •! This is not a compliment:
  "Its so complicated, I can't understand the ideas"
- •! It's easier to be complicated
- •! If it were not unsimple then how could distinguished colleagues in departments around the world be positively appreciative of both your extraordinary intellectual grasp of the nuances of issues as well as the depth of your contribution?
- •! Best results are obvious in retrospect
  "Anyone could have thought of that"

Style #3 – From Simple to Complex…

From simple to complex
- –! Should be able to understand even if not in field at coarse level
- –! gets more complicated and loses readers as you go through
- –! appendices complex mathematics/ proofs / details which are not essential to the plot
- –! this structure repeated internally
  - •! explain idea in words
  - •! explain more precisely in words
  - •! write precise mathematical formulation

The EM algorithm alternately finds the expected values for the unknown pose-invariant vectors $\mathbf{c}$ (the Expectation- or E-Step) and then maximizes the overall likelihood of data as a function of the parameters θ (the Maximization- or M-Step). More precisely, the E-Step calculates the expected values of the invariant vector $\mathbf{c}_i$ for each individual i, using the data for that individual across all poses, $\mathbf{x}_{i1...iK}$. The M-Step optimizes the the values of the transformation parameters $\{\mathbf{F}_k, \mathbf{m}_k, \Sigma_k\}$ for each pose, k, using data for that pose across all individuals, $\mathbf{x}_{1k...Ik}$. These steps are repeated until convergence.

E-Step: For each individual, we estimate the distribution of $\mathbf{c}_i$ given the current parameter estimates θ_{t-1}. We assume that the probability distributions for $\mathbf{c}_i$ given each data point $\mathbf{x}_{i1...iK}$ are independent so that:

$$p(\mathbf{x}_{i1...iK}|\mathbf{c}_i, \theta) = \prod_{k=1}^{K} p(\mathbf{x}_{ik}|\mathbf{c}_i, \phi = \phi_k, \theta) \qquad (7)$$

where the terms on the right hand side are calculated from the forward model (Equation 3). Since all terms are normally distributed, the left hand side is also normally distributed and can be represented with a mean vector and covariance matrix. We use Bayes rule to combine this new likelihood estimate with the prior over the invariant space as in Equation 2. This yields a posterior distribution similar to that in Equation 5. The first two moments of this distribution can be shown to equal:

$$\mathrm{E}[\mathbf{c}_i|\mathbf{x}, \theta] = \left(\mathbf{I} + \sum_{k=1}^{K} \mathbf{F}_k^T \Sigma_k^{-1} \mathbf{F}_k \right)^{-1} \sum_{k=1}^{K} \mathbf{F}_k^T \Sigma_k^{-1} (\mathbf{x}_k - \mathbf{m}_k)$$

Style #3 – Tell a story

- ! The paper should flow properly and tell a story.
- ! You need to put in formal structure to make it a coherent whole.
- ! This might involve
  - ! Foreshadowing of results in introduction
  - ! Explicit declarations of structure (In Section 2.1…)
  - ! Explanation in each section for why we are here and how it relates to the whole
  - ! Relating the discussion back to the introduction etc.

Style #4 – Use the formal structure of the paper

- ! Use subheadings to organize your material and highlight key points / areas
- ! Use paragraphs to organize your ideas
  - ! In general one idea per paragraph
  - ! Expressed through several sentences
- ! Lack of paragraphing makes text difficult to follow and suggests that your ideas are not organized

Style #5 - Tersify

Omit needless words. Don't be surprised if this turns out to be 30-40% of the words you originally wrote. Your first effort rarely captures the most vigorous or concise way to say something. Spend time tersifying.

A good example is using "utilization' when "use" will do. Do not use jargon where regular language will do. Another example is the use of "in order to." Any time you write that phrase, delete it and replace it with "to."

Style #6 – Terms

You should (of course) define all variables in equations and abbreviations.

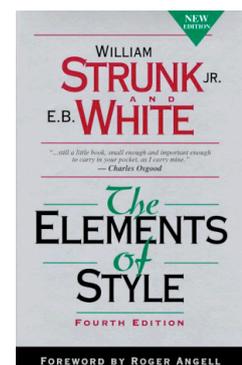You should (also) consider defining non-general technical terms as you use them

e.g. where by "feature-based approach" we refer to …

Style #7 – Know your audience

Scientific and technical writing can almost never be 'general purpose'; it must be written for a specific audience. In all cases, you must adopt the style and level of writing that is appropriate for your audience. Stylistic conventions and acceptable jargon can vary tremendously from one field to another, and to some extent, from one journal to another. If you are unfamiliar with the conventions of a field, study them as they are manifested in a selection of highly regarded papers and in the "Instructions for Authors" for key journals.

STYLE # 8 - Grammar

- ! #1 Read this book:

- ! #2 Find someone who is anal retentive about grammar to check your manuscript.

WARNING!

Your supervisor/professor is not here to teach you basic grammar and spelling

The more time and emotional energy she or he spends on correcting basic English usage, the less remains for issues of content or fine-tuning. You are responsible for mastering the basics of the language; save your supervisor's time for more substantive issues. A few glitches and non-parallel tenses will sip through your own careful editing, but there is no excuse for frequent ungrammatical sentences.

Style # 9 - PAPER vs THESIS

•! A thesis is not a string of papers, and neither is a paper part of a thesis
•! Paper is more positive
•! Paper much more concise

PLAN OF TALK

•! Why should I publish papers?
•! When should I publish papers?
•! Where should I publish papers?
•! Parts of a published paper
•! Style
•! Tips for writing the paper
•! The review process

TEMPLATES

•! Most conferences provide templates for Microsoft word and LaTeX.
•! Use this template!
•! Learn LaTeX! If you are going to be a professional scientist, then you cannot avoid this and there's no point delaying it.
•! Good online LaTeX help at Cambridge University: http://www.eng.cam.ac.uk/help/tpl/textprocessing/
•! There are also courses organized by the Graduate school

BEFORE YOU WRITE…

•! If possible, present your work in a short 5-10 minute talk to your colleagues before starting to **write**.
  –!Will help identify weaknesses
  –!May also notice strength that you did not see

WRITING A PAPER: Overview

1.! 1-page paper outline, with tentative page budget/section – 1 sentence per section to describe

2. Paragraph map1 topic phrase/sentence per paragraph, handdrawn figures w. captions

3. (Re)Write draft
Long captions/figure can contain details ~ Scientific American Uses Tables to contain facts that make dreary prose

4. Read aloud, spell check & grammar check
(MS Word; Under Tools, select Grammar, select Options, select "technical" for writing style vs. "standard"; select Settings and select)

5. Get feedback from friends and critics on draft; especially people unfamiliar with the research problem; go to 3.

WRITING A PAPER: Order

(1)! Write your introduction – if you can't make a case to justify the paper then there's not much point carrying on.

(2)! Decide on your figures. How many? What will each one show? Build the paper around these.

(3)! Write the methods section – usually the easiest section to write.

(4)! Write the discussion – should address aims in introduction.

(5)! Finally, write your abstract.

WRITING A PAPER: Time

•! Don't cram. You can't throw all your time into a **paper** at the last minute and expect a **good** result

•! Don't hand in a first draft! Ever! Most people's first drafts are terrible. I would not make anyone else suffer through mine. Don't make others suffer through yours. "**Good writing is rewriting**," and you should make a serious effort at editing, rewriting, and fine-tuning before you give the manuscript to anyone else to read. There are few things more frustrating to read than a paper in which you know there are pearls of wisdom, but where sloppy and ambiguous writing hides those pearls. If you need to put a piece of writing away for a few days before you can approach it dispassionately enough to rework it, do so.

•! Know when to say when. Even if you have written the **paper** with plenty of time and had a lot of outside review, after a certain point you will not be able to add much value without taking a break for a while (maybe a week or two).

FINAL CHECKLIST (from Armando Fox, Stanford)

1. Does the paper make clear precisely what your new contributions are, and how they are different/better than existing approaches to this or similar problems?

2. Does the outline of the paper (sections, subsections, etc.) cohere regardless of the granularity at which you view it?

3. Have you observed the following invariant: Before telling me what you did, tell me why I should care.

4. Have you made every important point three times--once in the introduction/ abstract, once in the body of the paper, and once in the conclusions?

FINAL CHECKLIST (from Armando Fox, Stanford)

5. Have you had it read by at least one person familiar with each of the areas the paper impinges on? (Think of them as consultants in that area. There is a risk that you will get some of the details wrong in talking about an area that is tangential to the paper but that you're not very familiar with, and if a reviewer happens to be versed in that area, it decreases your credibility. Such references are easy to get right, so there is no excuse.)

6. Have you searched carefully for any related work, and properly acknowledged it? The availability of papers and search indices on the Web makes it worse than ever to overlook significant related work.

7. Are you able to capture the non-experts in the audience with the opening of your paper, and impress the experts in the body of the paper?

8. Can you read only the abstract and conclusions and be able to give someone else a 30-second digest of what the paper claims it says?

ANONYMIZING

Many conferences are "double-blind"

-remove your name(!)

-do not cite yourself 20 times

-do not write "In previous work (Prince 2005) we have"…

-do not write "Prince 2000 produced an exceptional work of great brilliance in which he…"

LENGTH

If you find yourself with space to spare, then worry!

Usually papers are 150% of the maximum possible size, and are then edited down to fit the format.

It is possible to do this without really losing any information by an intensive process of tersifying.

DEADLINES

- ! Many conferences suffer from "deadline drift"
- ! Do not e-mail the committee begging after the deadline has passed
- ! Ideally, you should have a full first draft at least 1 month before the deadline.
- ! Sort out all the things that you can do early (references etc.) as soon as possible

PLAN OF TALK

- ! Why should I publish papers?
- ! When should I publish papers?
- ! Where should I publish papers?
- ! Parts of a published paper
- ! Style
- ! Tips for writing the paper
- ! The review process

PEER REVIEW

- ! Peer review ensures a degree of quality control in the publication process
- ! In general, referees will consider:
  - ! Is the work seriously flawed in approach?
  - ! Are the conclusions/ arguments valid?
  - ! Is the content sufficient, sufficiently important and suitable for publication in this journal?

Who will read your paper? (from Armando Fox, Stanford)

Remember that this will be read by people who

(a)!have never heard of you and the review is anonymous anyway,
(b)!have never heard of your project
(c)!are reading about 15-20 papers apiece, all in different subject areas.

They will spend the first 5 minutes deciding if your paper is actually good enough to be worth a fully detailed read; they will then spend an hour or so reading it in detail, trying to figure out

(a)!what your contribution is,
(b)!if the contribution is substantial enough to be worth publishing,
(c)!if the contribution is "feasible" (ie it is implementable and therefore would be useful to someone).

PEER REVIEW (PAPERS)

- ! Journal editors send submitted abstracts out to relevant researchers for assessment (usually 3)
- ! Referees accept or decline to review this paper
- ! The editor decides whether to publish the manuscript on the basis of the referees' reports
- ! Editor has the final decision

PEER REVIEW (PAPERS)

- ! Referees are usually kept anonymous
- ! Authors are sent copies of the referees' reports
- ! Even when referees recommend acceptance they usually insist on some changes/corrections
- ! Peer review is not infallible: papers can still be published with flaws outside referees' expertise.

THE REVIEW PROCESS (Conferences)

- •! Program committee –> Area Chair -> Reviewer
- •! 2-3 reviewers per paper
- •! usually more senior academics – trusted etc.
- •! May review up to 100 papers per year + grants
- •! Probably will not know exactly about the topic of your paper.
- •! Standard likely to be higher for journals than conferences

REVIEWING

Reviewing
  -Find a copy of the review form
  -Criteria:
     -technical correctness
     -novelty
     -comparison to other methods
     -clarity
     -relevance to field
     -contribution
(EXAMPLE FORM)

Responding to Reviewers

- •! Be polite!
- •! If they didn't understand something then take the attitude that is your fault for explaining it badly.
- •! If they are really wrong, then point out WITH EVIDENCE why this is the case.

WHAT IF YOU ARE REJECTED?

- •! Are reviews fair? Not always, but they are mostly reliable. If you are rejected, it is probably your fault.
- •! Use the advice the reviewers have given you – particularly if more than one referee makes the same point.
- •! Resubmit to the same conference next year, or possibly a conference more suited to the perceived contribution.

WHAT IF YOU ARE ACCEPTED? PROOFS

- •! For journals, you will be sent proofs. You cannot make technical changes now, only correct small details and typographical errors.
- •! Do not assume that errors that were not in the drafts have not been introduced by the printers.

PUBLISHING ONLINE

- •! Done by almost all academics now
- •! Use PDF format
- •! Be consistent with how you write your name
- •! Add page numbers and reference information so people can easily find your work
- •! Submit to online databases such as CiteSeer.

UCL

How to improve your chances of success:

5 Practical things that you can do after this session

1.! Get real review forms from your supervisor
2.! Set up a mock review panel
3.! Look at your favourite papers and examine the structure, paragraph by paragraph
4.! Practice tersifying
5.! Buy Strunk and White