Methodology and Tools for Research: Writing scientific material

Yannick Prié Polytech Nantes, University of Nantes Master DMKM, 2015-2016

CC BY-SA 4.0

Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)



- This course "Methodology and Tools for Research: Writing Scientific Material" by <u>Yannick Prié</u> is licensed CC BY-SA 4.0
- This license covers the general organization of the material, the textual content, the figures, etc. *except where indicated*.
- This license means that you can share and adapt this course, provided you give appropriate credit to the author and distribute your contributions under the same license as the original
 - for more information about this license, see <u>http://creativecommons.org/licenses/by-sa/4.0/</u>

Objectives of this course

- Writing scientific papers
- Writing other documents: research reports, posters, presentations

Ressources for the course

http://www.scoop.it/t/toolsandmethodologyforresearch

Preliminary remarks

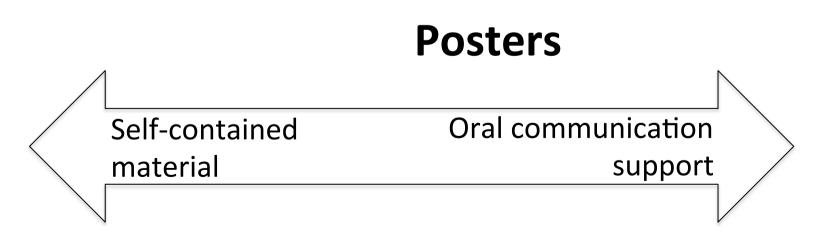
- Lots of advice can be found on the Web
- Often not exactly the same
 - depend on communities
- Often "common sense" advice: it is all about coherence and about your reader
 - common sense sometimes important to be remembered

What is written scientific material for?

To convey one's ideas to a dedicated public

Articles

Presentations



Outline

- Where should I publish?
- Papers
- Posters
- Presentations

Identify the idea you want to communicate

- New ways of looking at things (model)
- New way of manipulating objects (technique)
- New facts concerning objects (results)
- etc.

Honestly assess the quality of your work

- How good and important are your results?
- Why so?
- Differences between
 - Preliminary ideas on a new topic
 - First experimental results
 - e.g. from a master's thesis
 - Summary of a 3-year research project

Identify the relevant scientific (sub-)community

- What do they already know on the topic?
- Why would they read the paper?
- How will they read it?
- Focus both

(O. Goldreich)

- on experts scientists
- on their future and current graduate students
 - \rightarrow write for the good student

Choose the appropriate medium

- Workshop
- Average or top conference
- Average or top journal
- (Poster)
- Think long term
 - Defend the ideas that deserve it by making them progress and be better publisher

Write according to the publication target

- Identify the format of the conference / journal
 - One or two columns?
 - Number of pages?
 - Word/Latex model?
- Identify the "style of writing" of the target
 - Experimental papers?
 - Place of related works?
 - Auto-references?
 - Average number of references?
- \rightarrow Read published papers to get into the mood

Outline

- Where should I publish?
- Papers
- Posters
- Presentations

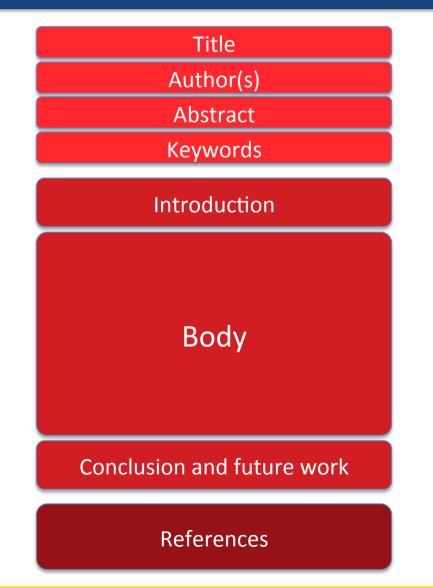
So, what's in an article?

- Here is a problem
- It is an interesting problem
- It has not yet be solved
 - Or not as good as I do
- This is my idea
- This is a working idea
- This is how it compares to others approaches

Conveying the idea to the target

- Presentation need to be clear
 - It is your duty to help readers extract relevant information from your paper
- Intuition is essential
 - A reader who catches the intuition will be willing to read the details
 - NOT the other way around
 - A reader can benefit from the article even if she does not read the details

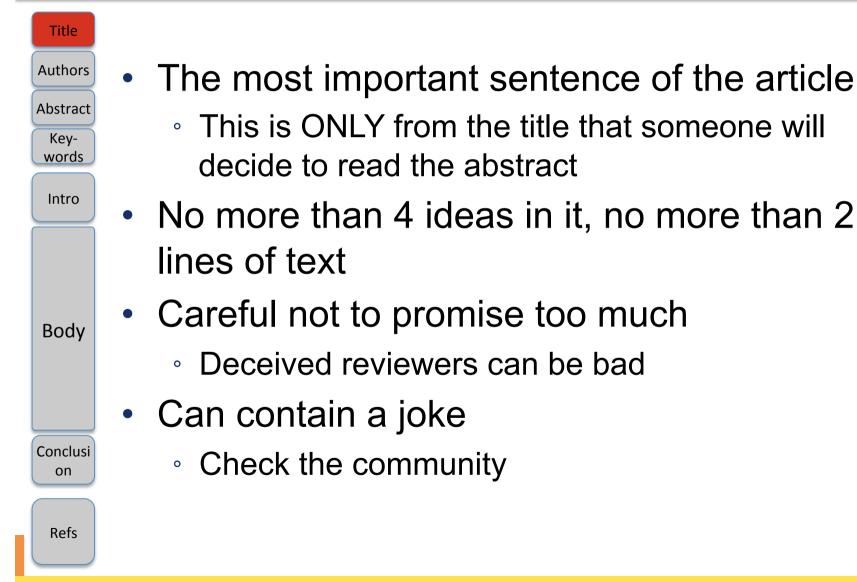
Reminder: What's in a paper?



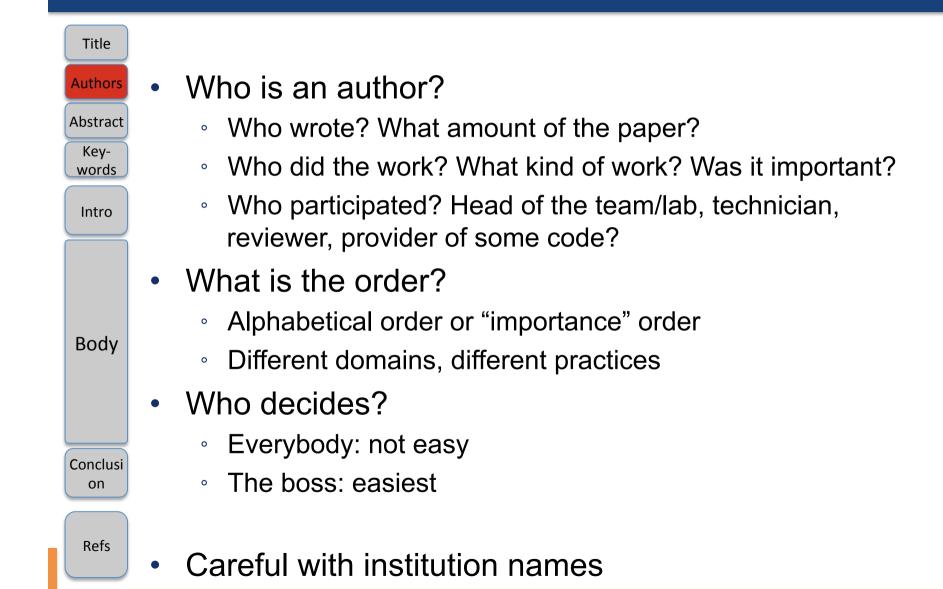
Companion reading

- Choose and open one article
 - From the PC assignment
 - From the case study assignment
- Keep an eye on it/them while the course unfolds
 - look for the various elements

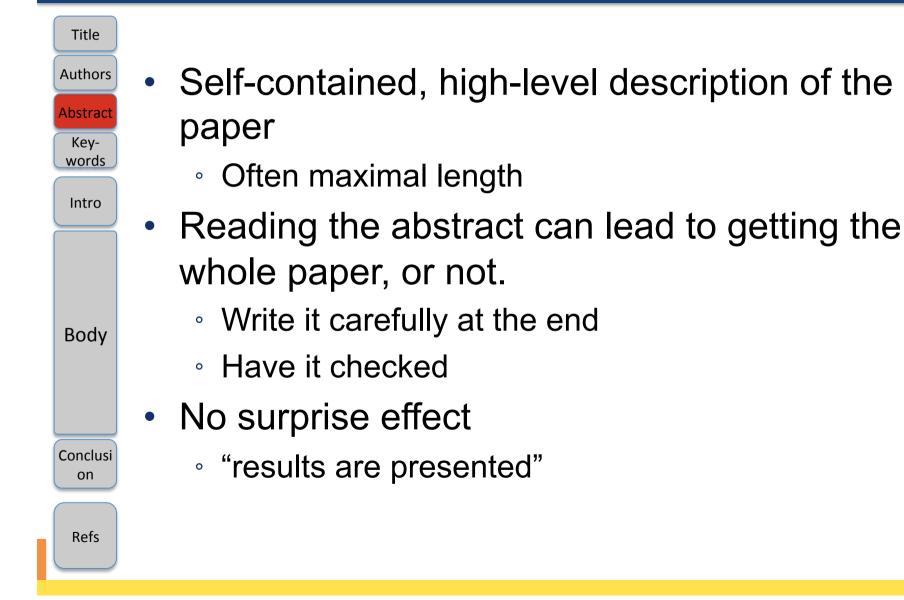
Title



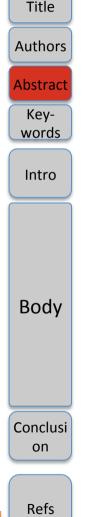
Author(s)



Abstract



Abstract: Four sentences proposal by Kent Beck



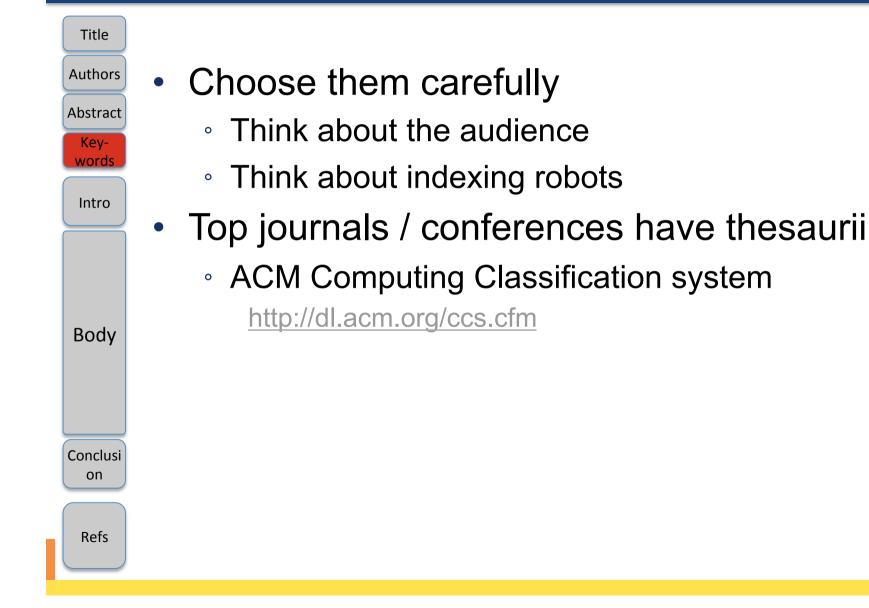
- The first states the problem.
- The second states why the problem is a problem.
- The third is my startling sentence (positive affirmation = main result)
- The fourth states the implication of my startling sentence

The rejection rate for OOPSLA papers in near 90%. Most papers are rejected not because of a lack of good ideas, but because they are poorly structured. Following four simple steps in writing a paper will dramatically increase your chances of acceptance. If everyone followed these steps, the amount of communication in the object community would increase, improving the rate of progress.

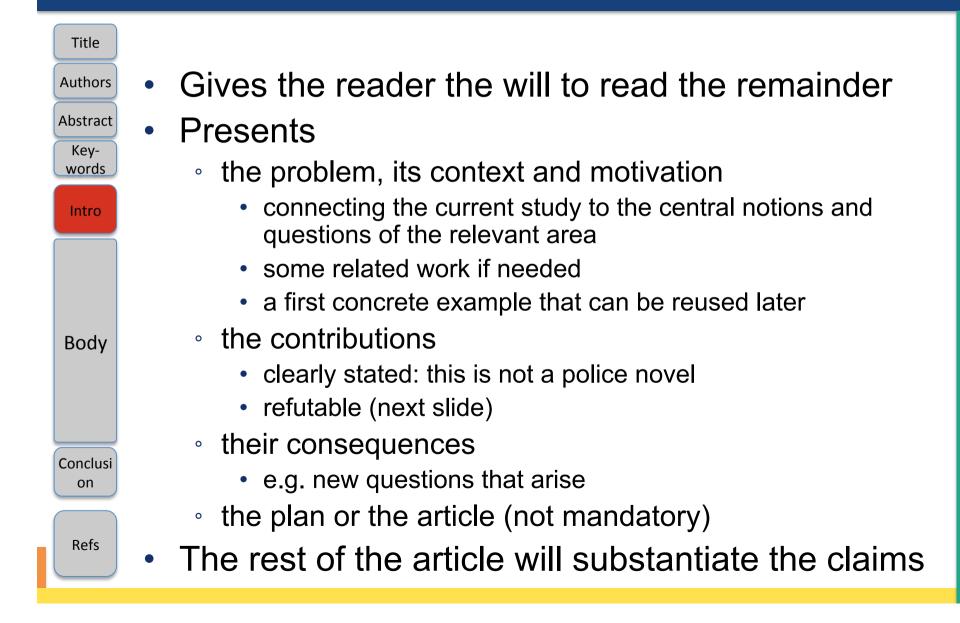
Nature's Authors Guide

Title	How to construct a Nat	ure summary paragraph
Authors		are summary paragraph
Abstract	Annotated example taken from <i>Nature</i> 435 , 114–118 (5 May 2005).	
Key- words	One or two sentences providing a basic introduction to the field, comprehensible to a scientist in any discipline.	During cell division, mitotic spindles are assembled by microtubule- based motor proteins ^{1,2} . The bipolar organization of spindles is
Intro	Two to three sentences of more detailed background , comprehensible to scientists in related disciplines.	essential for proper segregation of chromosomes, and requires plus- end-directed homotetrameric motor proteins of the widely conserved kinesin-5 (BimC) family ³ . Hypotheses for bipolar spindle formation include the 'push-pull mitotic muscle' model, in which kinesin-5 and opposing motor proteins act between overlapping microtubules ^{2,4,5} . However, the precise roles of kinesin-5 during this process are unknown. Here we show that the vertebrate kinesin-5 Eg5 drives the sliding of microtubules depending on their relative orientation. We found in controlled <i>in vitro</i> assays that Eg5 has the remarkable capability of simultaneously moving at ~20 nm s ⁻¹ towards the plus- ends of each of the two microtubules it crosslinks. For anti-parallel microtubules, this results in relative sliding at ~40 nm s ⁻¹ , comparable to spindle pole separation rates <i>in vivo</i> ⁶ . Furthermore, we found
Body	One sentence clearly stating the general problem being addressed by this particular study.	
	One sentence summarizing the main result (with the words "here we show" or their equivalent).	
	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.	
		that Eg5 can tether microtubule plus-ends, suggesting an additional microtubule-binding mode for Eg5. Our results demonstrate how members of the kinesin-5 family are likely to function in
	One or two sentences to put the results into a more general context.	mitosis, pushing apart interpolar microtubules as well as recruiting
Conclusi		microtubules into bundles that are subsequently polarized by relative sliding. We anticipate our assay to be a starting point for more
on	Two or three sentences to provide a broader perspective , readily comprehensible to a scientist in any discipline, may be included in the	sophisticated <i>in vitro</i> models of mitotic spindles. For example, the individual and combined action of multiple mitotic motors could be
Refs	first paragraph if the editor considers that the accessibility of the paper is significantly enhanced by their inclusion. Under these circumstances, the length of the paragraph can be up to 300 words. (This example is 190 words without the final section, and 250 words with it).	tested, including minus-end-directed motors opposing Eg5 motility. Furthermore, Eg5 inhibition is a major target of anti-cancer drug development, and a well-defined and quantitative assay for motor function will be relevant for such developments.

Keywords



Introduction

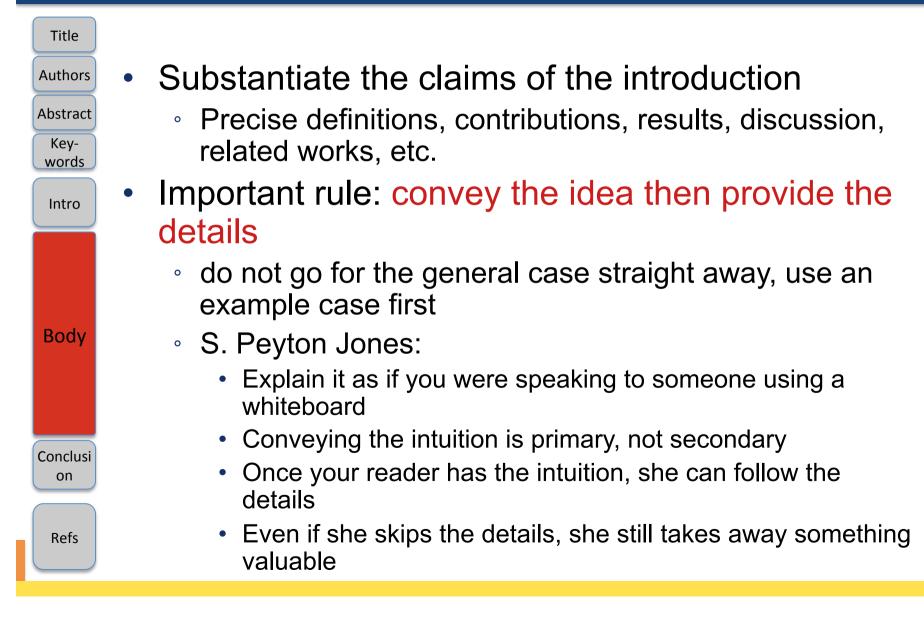


Introduction:

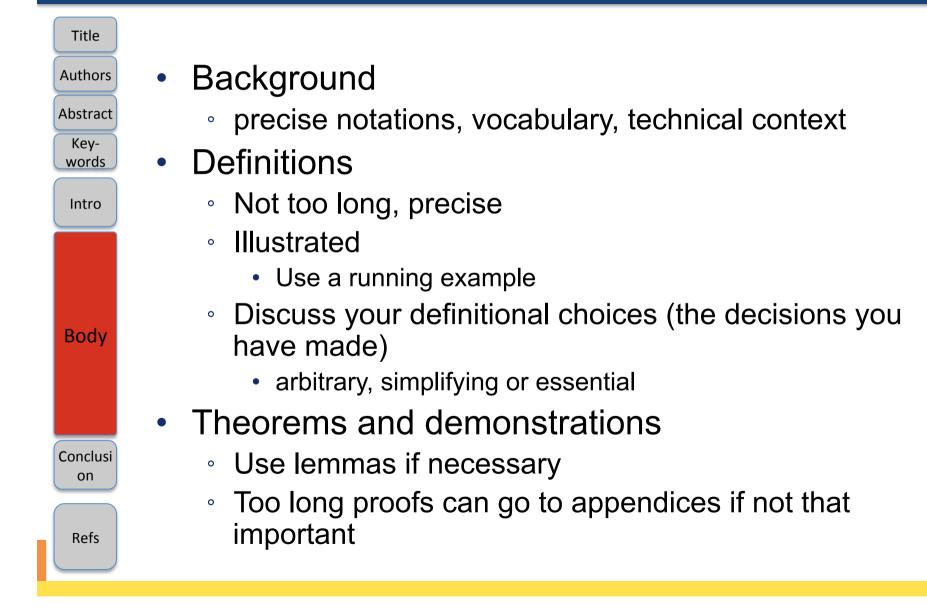
Refutable and clear contributions (S. Peyton Jones)

Title		
Authors	NO	YES
Abstract Key- words Intro	We describe the WizWoz system, it is cool.	We give the syntax and semantics of a language that supports concurrent processes (Section 3). Its innovative features are
Body	We study its properties	We prove that the type system is sound, and that type checking is decidable (Section 4)
Conclusi on Refs	We have used WizWoz in practice	We have built a GUI toolkit in WizWoz, and used it to implement a text editor (Section 5). The result is half the length of the Java version.

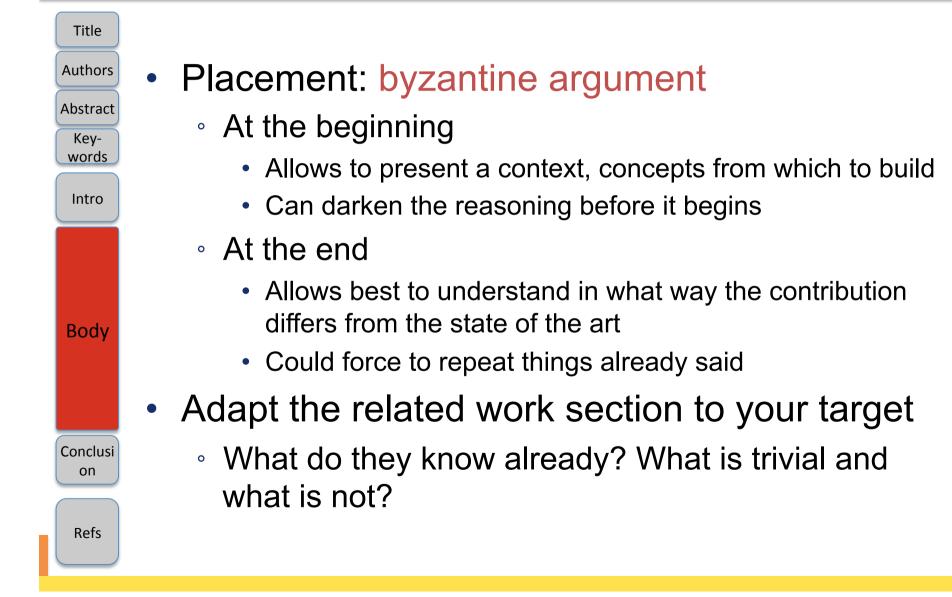
Body: "flesh" of the paper



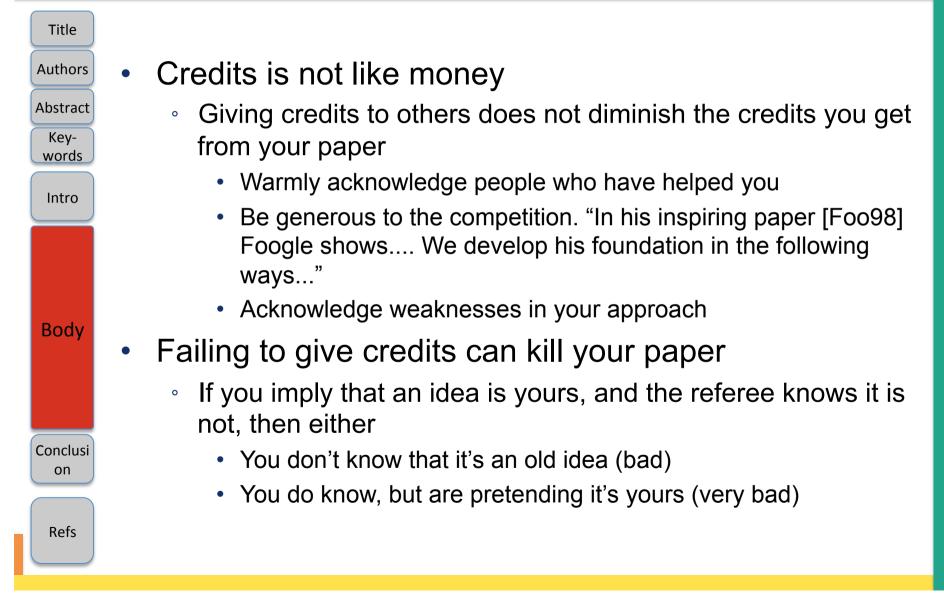
Body: Background, definitions, theorems and demonstrations



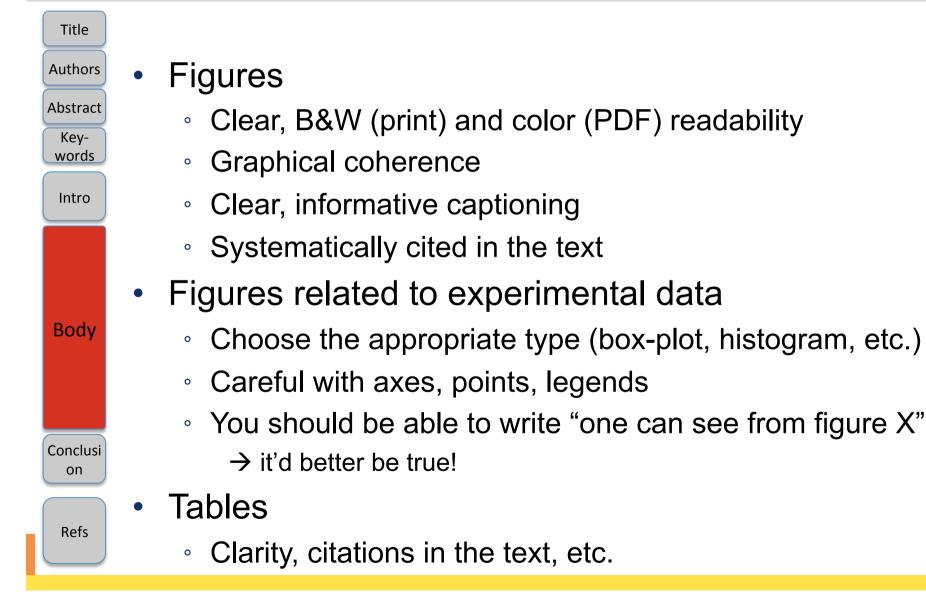
Body: Related work



Body: Related work (cont): S. Peyton Jones

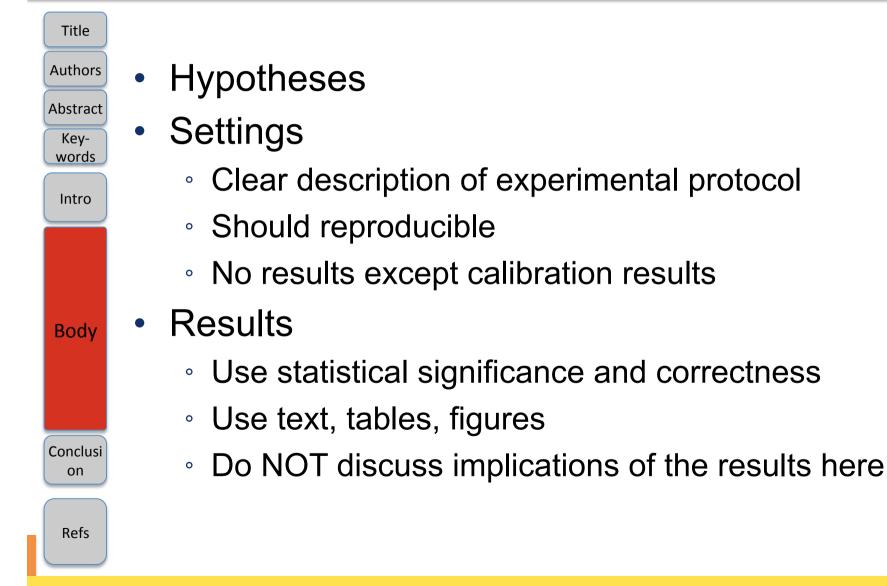


Body: Figures and tables

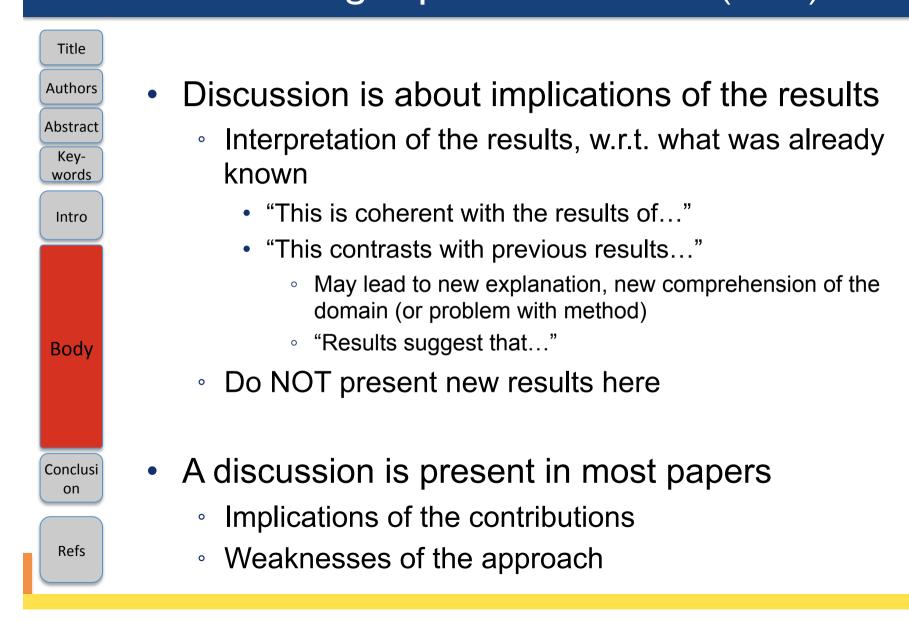


29

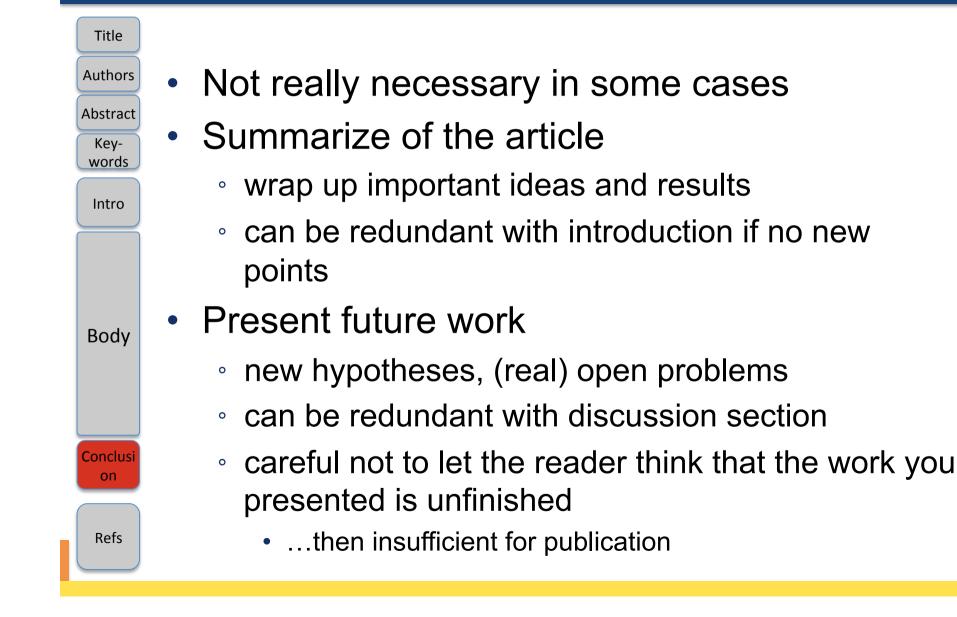
Body: Presenting experimental work



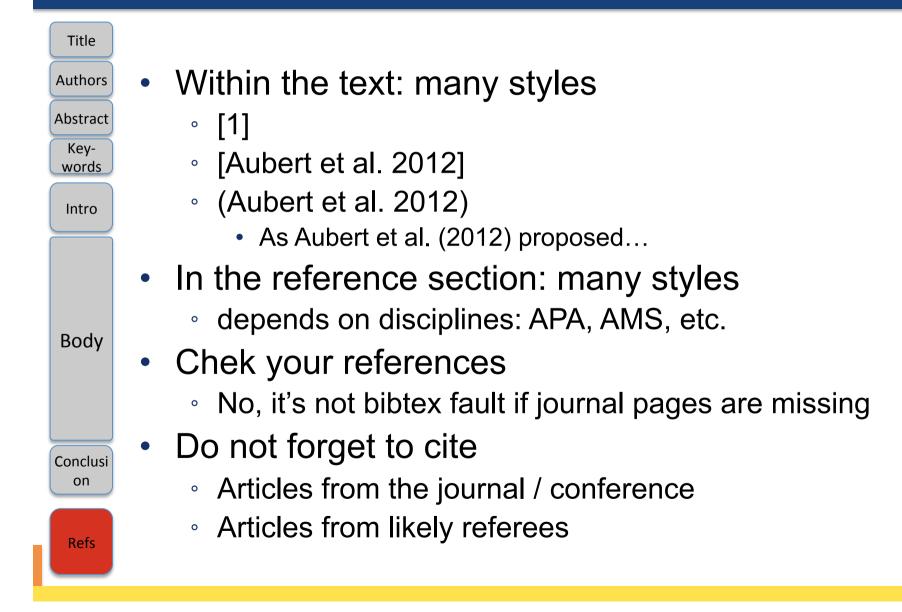
Body: Presenting experimental work (cont)



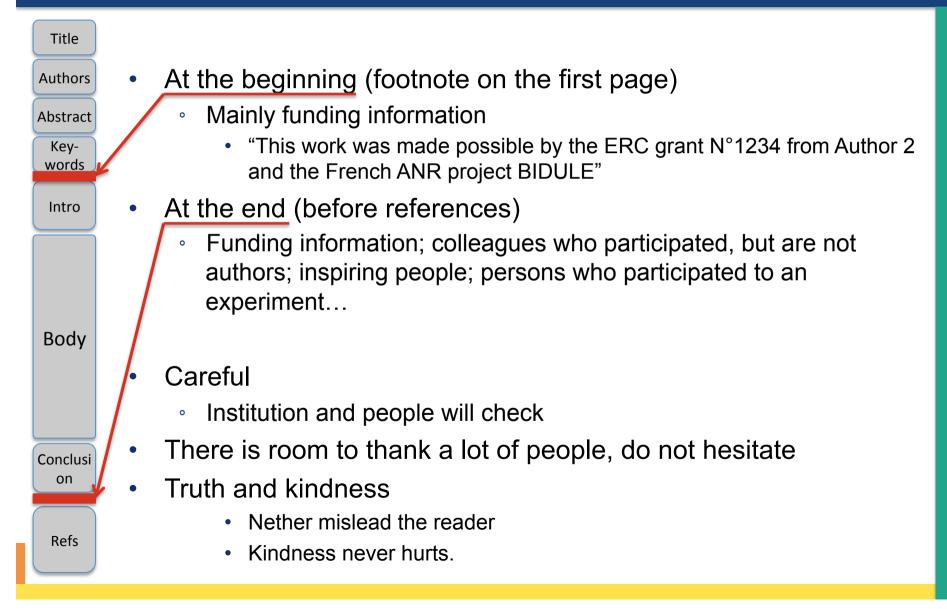
Conclusion



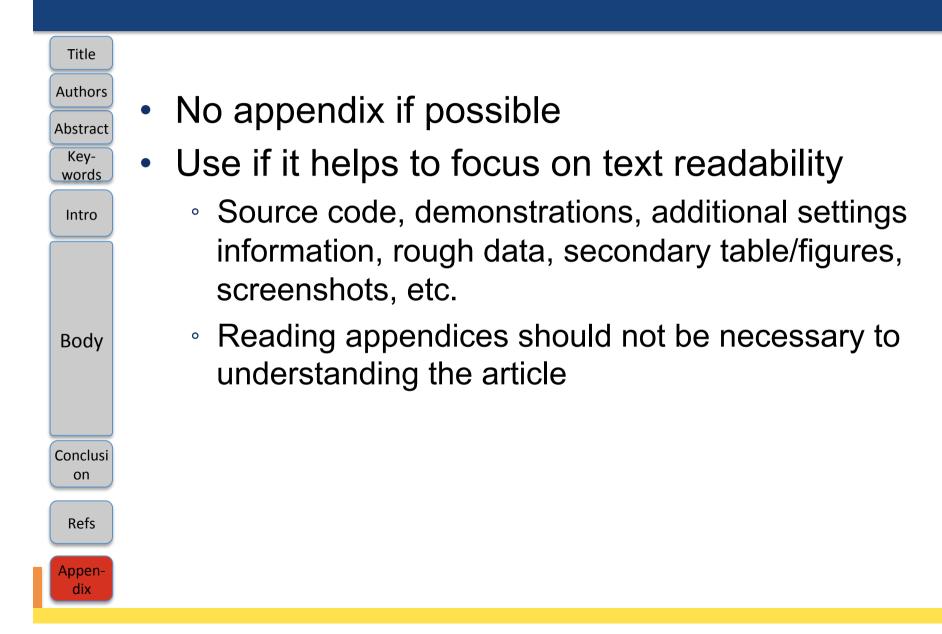
References



Acknowledgments



Appendix



Advice I: Remember pass I for reading a paper

- Title + authors
 - What is it about + Where does it come from?
- Abstract
 - What was done, what is the contribution?
- Medium
 - What is the audience?
- Introduction / conclusion
 - What is the context + what are the results?
- (Sub-)sections headings, figures, formulas
 - What is the paper general structure, contribution?
- References
 - Is it a serious paper?

Category Context Correctness Contributions Clarity

→ write so that any reader can answer these questions within 5 minutes

→ write the most important parts at the end (abstract, introduction)

Advice 2: Outline important ideas / message

- Write with honesty
 - ... but remember you have to be convincing
- Repeat important information
 - Title, abstract, introduction, discussion
- Place it where it will be recognised as such
 - Section/subsection titles
 - End/beginning of sections
 - Short paragraphs
- Be careful with the section/subsections titles
 - Informative enough to reveal the article structure

Advice 3: Be careful with language

- Systematically use a spell-checker
- Get inspiration from sentences found in good articles
- Read books
 - The elements of style (Strunk 1918)
 - <u>http://www.crockford.com/wrrrld/style.html</u>
- Take lessons of scientific english
- Wisely use punctuation
- He? She? She-he?
 - Copy on accepted articles

Advice 3: Be careful with language

- Keep a clear, concise, simple and direct language
 - something equivocal is deemed false
- No sentences with complex logical structure
 - particularly if you are no sure
- Banish labyrinths of indirections with implicit pointers (it, this)
 - prefer repetitions
- No acronyms
 - unless well known
- No cumbersome notations (a, b, c, d, e, f, g, h, i, j)-system
- Careful with mixtures of mathematical symbols and text

$$M^{O_b^c}_{i_{j,k_t}}$$

Advice 3: Be careful with language (cont)

- Focus on the subject and the public target
- Be careful with expressions that defy the reader
 - "Never, always"
 - "Clearly demonstrate"
 - "Unambiguous"
 - "It is obvious"
 - "Very"
- Use dynamic verbs
 - "We performed the measurement of" \rightarrow "we measured"
 - Use active voice (next slide)

Advice 3: Be careful with language: use active voice

NO	YES
It can be seen that	We can see that
It might be thought that this would be a type error	You might think this would be a type error
These properties were thought desira ble	We wanted to retain these properties
34 tests were run	We ran 34 tests
	(S. Peyton Jones)

Advice 3: Be careful with language: use simple, direct langage

NO	YES
The object under study was displaced horizontally	The ball moved sideways
On an annual basis	Yearly
Endeavour to ascertain	Find out
It could be considered that the speed of storage reclamation left something to be desired	The garbage collector was really slow

(S. Peyton Jones)

Advice 4: Reinforce the visual structure of the paper

- Use enumerations and lists
- Use figures, tables and take care of their position
- Wisely use sections, sub-sections, sub-subsections, paragraphs
- Emphasize, do not use bold fonts in the text
- Use dedicated style for source code and algorithms

Advice 5: Use the right tools for writing

- A writing tool is very important
 - Tool, text and ideas are not as separated as one believes
- Preparation / structure
 - Outliners, mind maps
- Writing
 - Word processor
 - WYSIWYG, Latex
 - General drawing tools
 - Visio, Inkscape, Omnigraffle
 - Dedicated drawing tools
 - Rstats
 - References management
 - Zotero, EndNote
- Versioning tools
- Collaborative features

Advice 6: Start early

- Papers and idea need time to mature
- Best papers have had a first version weeks before the deadline
 - papers should be reviewed: advisor, colleagues, etc.
 - too many conference papers are finished 10 minutes before deadline
 - difference being accepted and rejected paper can just be one or two cycles of reading / re-writing

Remark 7: Get help

- Find people to read the paper
 - Experts and non-experts
 - Only one first reading per reader!
- Get useful reviews
 - not just grammar/spelling \rightarrow understanding problems are better
- Really listen to the reviews and give attention to each point
 - If somebody had a remark, you may not necessarily follow her suggestion, but acknowledge that a problem has been spotted
- Thank the reviewers warmly

Advice 8: Take into account conference/journal reviewers comments

- There is always something to improve from reviewers' comments
 - Get over the form, even if very negative
 - Incomprehension may not mean that the reviewer is dumb
- For a journal
 - Send a letter that explains every modification to the reviewers
- For a conference with rebuttal
 - Explain how you will take into account the reviewers comments should the paper be accepted

Additional advice (Goldreich 2004)

- Focus on the reader's needs rather than on the writer's desires.
- Careful with
 - checklist phenomenon
 - obscure generality
 - idiosyncrasies
 - lack of hierarchy/structure
 - "Talmud-ism"

Additional advice (Goldreich 2004)

- Awareness of the knowledge level of the reader
 - Definition: the reader will not understand everything at first read
 - Proofs: focus on conceptual steps before technical ones
 - Ideas: do not begin with the general case, rather with special case
 - Difficulty should not be hidden, but discussed
 - New concepts: not too much

Outline

- Where should I publish?
- Papers
- Posters
- Presentations

What is a conference poster for?

- Helping communicate ideas to people who choose to spend 5 minutes with you
 - Small audiences (1-5 persons)
- Communicating these ideas on its own
 - A reader should be able to grasp the content by reading it from introduction to conclusion
- Both
 - After all, you will not be present all the time next to your poster
 - Most posters finish their lives hanging on labs' corridor walls

What the situation looks like (small venue)



The poster session at the 17th International Symposium on Graph Drawing, Chicago, 2009 by David Eppstein is licensed under CC BY SA 3.0

What the situation looks like (larger venue)

CAP

Part in

-

Computer science posters





A CONTRACTOR OF STREET, STREET	TU and Antonio frances frances for the formation of the second se
i le. Le.	



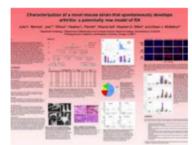
Las Alamos POSI Activities: Outreach, Characterization, Standar Gay Grite, Janes Karap, John Berl, Repter Gott: Los Karos National Laterat

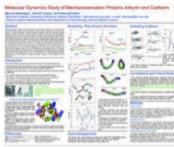
NAME OF TAXABLE PARTY.		2000
	States of the local division of the local division of the	
Briddler Bridden and Briddler	The second secon	
States of the local division of the state of the	states and states which is the state of the	
the state is not a state and a state		
A Date March State Street William	The other	
and all standard programs, And and	A COLORADO DE COLORADO DE COL	
100.001000000	And A CONTRACTOR OF A CONTRACTOR OFTA CONTRACT	
PARTICULARY	And the lot of the second design of the	
	The second secon	
	States and With a local date of	
Statements in a statement in the		
101100-		had and
And the second s	the lot into a building of	
	Address of the Addres	
station of a state of the state	characterization and statements	distant in the local distance of the local d
on these strength of the local division of t	and the second s	and a second second
Store on the second second	SOLUTION STATES	And in the second
Concession of the second second		The Rest of Concession of Concession, Name
	Contract of the local division of the local	and the second s
and the second second	100	and have been as a set of
	TRACTOR OF THE OWNER	100
Street & Western Starting	Statement in the second second	1. 1703/000
And Springer Springer		
THEFT	0	
ALC: NO. OF COMPANY		
Discharge and the second	APRIL PROPERTY AND ADDRESS	
		An other statements
		-

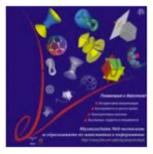
	and long' App and Ann And welling	
	0.0	
	-2825	

600	NUMBER OF STREET	(Allengana)

	THE RESEARCH I	tion on the Web
Reads 1975 menu (da pli) and a second and a second and a second and a second and a second and a second and a second a second and a second a second a second a secon	Name and a funge of the state of the state of the the state of the state of the	Ken pungke norst i satisfication - Andre se las noti. Loid aper un la gai lago - la constanti apertanti - Marcela de la constanti a de la constanti - Marcela de la constanti -
Nago copin-resolute data an A lanks - Marganak - Automation - Nata an Automation - Nata an Automation - Automatio - Automation - Automation - Automa	The factor of the state of the	and the second second second

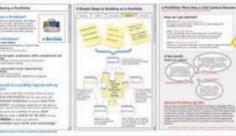




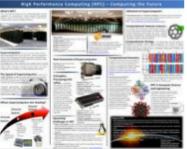




e-Portfolios: More Than a 21st Century Resume









Google image first result page, oct 2013, © Google | images © by their owners

54

Contents

- Short title
- Introduction
 - necessary concept / references
- Overview of the approach
- Results in graphical form

- Insightful discussion of results
- References

 not too much
- Brief acknowledgement
 - assistance and financial support

Bad Posters example

• see

http://colinpurrington.com/2012/example-of-badscientific-poster/

• or

https://www.google.fr/search?q=bad +poster&tbm=isch

Use the right tools

- Text design tools
 - Quark Xpress, In Design, Scribus (open source)
- Drawing tools
 - Illustrator, Omnigraffle (mac), Inkscape (SVG editor)
- (Latex)
- (powerpoint)

Some tips

- Length
 - 800 words max: under 5 minutes to read the whole content
- Illustrations
 - careful with the photo / image quality for printing (pixelisation)
- Fonts
 - non-serif font (e.g., Helvetica) for title and headings
 - serif font (e.g., Palatino) for body text.
- Text boxes
 - width: approximately 40 characters (av.11 words per line)
 - no longer than 10 sentences
- Logos
 - Avoid them
 - If not possible, hide them

Ask yourself one question

- Will I proudly stay next to my poster at the conference?
 - Does it contain every information I would need?
 - Is it attractive enough?
 - Is it clear enough?

ο

- Are there any typos?
 - (well, answer is always yes)

Outline

- Where should I publish?
- Papers
- Posters
- Presentations

Presenting a paper at a conference

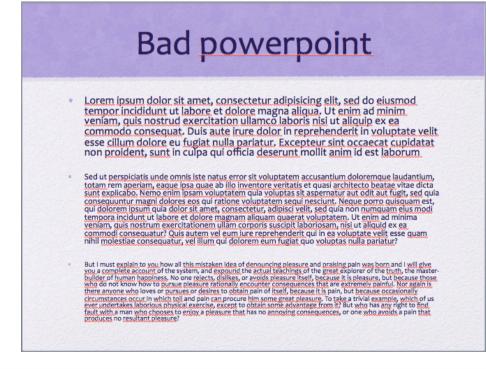
- Various situations
 - 10 to 1000 attendants
 - one or several sessions
 - 15 to 30 minutes (with questions)
- The occasion get interest from people in the room
 - Getting them to read the paper, to have students read the paper, to tweet about it
 - The presentation may be recorded and broadcasted

Visual aids for presentations

Туре	Pros	Cons
Projected slides	Images, underline key details	Can be boring Unefficient if too much details
White or black board	Mathematical demonstration	Not facing audience, slow, needs erasing
Videos	Stimulate imagination, movements, sounds	Audience focused the video Need short video, good quality prepared discourse
Demos	Get public interest	Can fail → rehearse/test Careful with idle times
Artefacts or props	Get public interest	Audience can get distracted
Paper handouts	Audience leaves with the written message	Audience can get distracted

Titles give the slide message (one sentence)

- Titles are substantived with
 - visual (image, figure) and
 - textual content (not too much)



What is bad for a poster is bad for slides

- TITLE SHOULD NOT BE
 IN CAPITAL LETTERS
- Use no more than 4 items per list
- Do not provide too much details
- Text should be readable from a distance
- The whole content of the paper should not be in the slides

- Use white space to visually arrange the slide and the reading order
- Use well designed figures
- Use images for outline slides
- ...
- Use animation if it supports the discourse

Some mistakes while presenting

- Forget Murphy's law
 - Rehearse, rehearse
 - Arrive early
- Miss the audience
 - Will the audience understand this point?
 - Will the audience get interest for this point?
- Be inattentive to the audience
 - Not speak loud enough
 - Move without a goal
 - Read the slides
 - Have no eye contact with all the audience
 - Have no idea of elapsed time
 - Do not listen to questions, do not reformulate

ANYTHING THAT CAN GO WRONG, WILL GO WRONG

Checklist for Scientific Presentations (not every item on this list applies to every presentation)

Speech	
Necessary information conveyed? Audience targeted? Terms defined?	Assertions supported? Tone controlled? Examples given?
Structure	
Organization of Beginning Scope defined? Topic justified? Proper background given? Talk memorably mapped?	Transitions Beginning/middle? Between main points of middle? Middle/ending?
Organization of Middle	Emphasis
Divisions of middle logical?	Repetition used effectively?
Arguments methodically made?	Placement used effectively?
Organization of Conclusion	
Main points summarized? Closure achieved?	

Checklist for Scientific Presentations (not every item on this list applies to every presentation)

Presentation Slides		
Slides orient the audience? Slides are clear to read? Slides have proper level of detail?	Slides show key images? Slides show key results? Slides show talk's organization?	
Delivery		
Speaker controls nervousness? Speaker shows energy? Speaker exudes confidence? Voice engages? Speed is appropriate? Filler phrases ("uh") are avoided?	Eye contact made? Movements contribute? Equipment handled smoothly? Questions handled convincingly? Questions handled succinctly?	

Ten commandments for (really bad) conference talk

1. Thou shalt not be neat

 Why waste research time preparing slides? Ignore spelling, grammar and legibility. Who cares what 50 people think?

2. Thou shalt not waste space

 Transparencies are expensive. If you can save five slides in each of four talks per year, you save \$7.00/year!

3. Thou shalt not covet brevity

 Do you want to continue the stereotype that engineers can't write? Always use complete sentences, never just key words. If possible, use whole paragraphs and read every word.

4. Thou shalt cover thy naked slides

• You need the suspense! Overlays are too flashy.

5. Thou shalt not write large

Be humble -- use a small font. Important people sit in front. Who cares about the riffraff?

Ten commandments for (really bad) conference talk

6. Thou shalt not use color

 Flagrant use of color indicates uncareful research. It's also unfair to emphasize some words over others.

7. Thou shalt not illustrate

 Confucius says ``A picture = 10K words," but Dijkstra says ``Pictures are for weak minds." Who are you going to believe? Wisdom from the ages or the person who first counted goto's?

8. Thou shalt not make eye contact

You should avert eyes to show respect. Blocking screen can also add mystery.

9. Thou shalt not skip slides in a long talk

You prepared the slides; people came for your whole talk; so just talk faster.
 Skip your summary and conclusions if necessary.

10. Thou shalt not practice

- Why waste research time practicing a talk? It could take several hours out of your two years of research. How can you appear spontaneous if you practice? If you do practice, argue with any suggestions you get and make sure your talk is longer than the time you have to present it.
- This commandment is the most important. Even if you break the other nine, this one can save you.

Outline

- Where should I publish?
- Papers
- Posters
- Presentations
- Conclusion

Writing rules: simple, (quite) obvious, numerous

- Mastering only comes with practice
 - reading and writing
- As for any design rules
 - Apply them, for they are accumulated wisdom
 - Understand them theoretically and practically
 - Develop your own style
 - Always remember not to stick to rules blindly
 - be flexible, apply good principles to the case at hand