How to write a scientific paper

Bedartha Goswami

7 July 2021

Section 1

Structure & components

Overview

- Title

Overview

- Title
- Abstract

Overview

- Title
- Abstract
- Main text (introduction, methods, results, . . .)

Overview

- Title
- Abstract
- Main text (introduction, methods, results, ...)
- Figures

Overview

- Title
- Abstract
- Main text (introduction, methods, results, . . .)
- Figures
- Data and code availability

Overview

- Title
- Abstract
- Main text (introduction, methods, results, . . .)
- Figures
- Data and code availability
- Acknowledgments

Overview

- Title
- Abstract
- Main text (introduction, methods, results, . . .)
- Figures
- Data and code availability
- Acknowledgments
- Conflicts of interest

Overview

- Title
- Abstract
- Main text (introduction, methods, results, ...)
- Figures
- Data and code availability
- Acknowledgments
- Conflicts of interest
- References

Overview

- Title
- Abstract
- Main text (introduction, methods, results, ...)
- Figures
- Data and code availability
- Acknowledgments
- Conflicts of interest
- References
- Supplementary

Title

- Short (max 15 words)

Title

- Short (max 15 words)
- Informative

Title

- Short (max 15 words)
- Informative
- Preferably no verbs, but only simple present tense if at all

Title

- Short (max 15 words)
- Informative
- Preferably no verbs, but only simple present tense if at all
- Avoid abbreviations

Title

- Short (max 15 words)
- Informative
- Preferably no verbs, but only simple present tense if at all
- Avoid abbreviations
- Avoid jokes (search engines might not get it)

Abstract

- Summarise the paper briefly

Abstract

- Summarise the paper briefly
- For a more general audience

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:
  - 1-2 sentences provide background and motivation / rationale

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:
  - 1-2 sentences provide background and motivation / rationale
  - 1-2 sentences describing what you do, e.g. “Here, we consider . . . ”

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:
  - 1-2 sentences provide background and motivation / rationale
  - 1-2 sentences describing what you do, e.g. "Here, we consider ..."
  - 1-2 sentences describing what you find, e.g., "We find that ...", "Our results indicate that ..."

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:
  - 1-2 sentences provide background and motivation / rationale
  - 1-2 sentences describing what you do, e.g. "Here, we consider . . . "
  - 1-2 sentences describing what you find, e.g., "We find that . . . ", "Our results indicate that . . . "
  - 1-2 sentences putting the results in context

Abstract

- Summarise the paper briefly
- For a more general audience
- Structure:
  - 1-2 sentences provide background and motivation / rationale
  - 1-2 sentences describing what you do, e.g. "Here, we consider . . . "
  - 1-2 sentences describing what you find, e.g., "We find that . . . ", "Our results indicate that . . . "
  - 1-2 sentences putting the results in context
- https://www.nature.com/documents/nature-summary-paragraph.pdf

Main text

- Introduction

Main text

- Introduction
  - Background

Main text

- Introduction
  - Background
  - Rationale

Main text

- Introduction
  - Background
  - Rationale
  - Related work

Main text

- Introduction
  - Background
  - Rationale
  - Related work
  - Contribution

Main text

- Methods

Main text

- Methods
  - Data used

Main text

- Methods
  - Data used
  - Methods used

Main text

- Methods
  - Data used
  - Methods used
  - Setup of experiments

Main text

- Results and discussion

Main text

- Results and discussion
  - Main figures of the paper

Main text

- Results and discussion
  - Main figures of the paper
  - Tables

Main text

- Results and discussion
  - Main figures of the paper
  - Tables
  - Brief (non-judgemental) interpretation of the figures

Main text

- Conclusion

Main text

- Conclusion
  - Summarise the main points

Main text

- Conclusion
  - Summarise the main points
  - Repeat what you did / set out to do

Main text

- Conclusion
  - Summarise the main points
  - Repeat what you did / set out to do
  - Repeat what you achieved

Main text

- Conclusion
  - Summarise the main points
  - Repeat what you did / set out to do
  - Repeat what you achieved
  - Provide a brief context of how your results will impact the field

Figures

- Use a plotting software that generates vector graphics (SVG, PDF)

Figures

- Use a plotting software that generates vector graphics (SVG, PDF)
- Avoid PNGs and JPEGs as far as possible

Figures

- Use a plotting software that generates vector graphics (SVG, PDF)
- Avoid PNGs and JPEGs as far as possible
- Ensure the smallest font size in the figure is no less than 11 pts

Figures

- Use a plotting software that generates vector graphics (SVG, PDF)
- Avoid PNGs and JPEGs as far as possible
- Ensure the smallest font size in the figure is no less than 11 pts
- Find out how big should the figure be (in/mm) and design accordingly

Figures

- Use colorblind friendly color-schemes

Figures

- Use colorblind friendly color-schemes
  - Use a color blindness simulator

Figures

- Use colorblind friendly color-schemes
  - Use a color blindness simulator
  - https://colororacle.org/

Figures

- Use color maps appropriately

Figures

- Use color maps appropriately
  - sequential color maps for data which are all non-negative or all non-positive (e.g., precipitation)

Figures

- Use color maps appropriately
  - sequential color maps for data which are all non-negative or all non-positive (e.g., precipitation)
  - divergent color maps for data which have both positive and negative values with the zero in between (e.g., precipitation anomalies)

Figures

- Use color maps appropriately
  - sequential color maps for data which are all non-negative or all non-positive (e.g., precipitation)
  - divergent color maps for data which have both positive and negative values with the zero in between (e.g., precipitation anomalies)
  - https://matplotlib.org/2.0.2/examples/color/colormaps_reference.html

Figures

- Use line plots and bar plots appropriately

Figures

- Use line plots and bar plots appropriately
  - line plots for continuous data, i.e., if it is meaningful to think of data in between two plotted data points, then a line plot is okay, e.g., hourly or daily temperature

Figures

- Use line plots and bar plots appropriately
  - line plots for continuous data, i.e., if it is meaningful to think of data in between two plotted data points, then a line plot is okay, e.g., hourly or daily temperature
  - bar plots for discontinuous data, e.g., monthly averages or sums, categorical data (e.g., country-wise averages or sums)

Figures

- Use line plots and bar plots appropriately
  - line plots for continuous data, i.e., if it is meaningful to think of data in between two plotted data points, then a line plot is okay, e.g., hourly or daily temperature
  - bar plots for discontinuous data, e.g., monthly averages or sums, categorical data (e.g., country-wise averages or sums)
- Always provide axis labels (even for the color bars)

Figures

- Use legends appropriately

Figures

- Use legends appropriately
- Figure legends (color keys, etc) should be such that the figure is understandable without reading the caption

Figures

- Use legends appropriately
- Figure legends (color keys, etc) should be such that the figure is understandable without reading the caption
- If you have space, then use the figure captions to quickly the results in the context of the article's narrative, e.g.

Figures

- Use legends appropriately
- Figure legends (color keys, etc) should be such that the figure is understandable without reading the caption
- If you have space, then use the figure captions to quickly the results in the context of the article's narrative, e.g.
  - "(A) shows the Niño 3.4 SST anomalies where the El Niños and La Niñas are clearly visible as large excursions from the normal SST climatology (the zero line)", as opposed to

Figures

- Use legends appropriately
- Figure legends (color keys, etc) should be such that the figure is understandable without reading the caption
- If you have space, then use the figure captions to quickly the results in the context of the article's narrative, e.g.
  - "(A) shows the Niño 3.4 SST anomalies where the El Niños and La Niñas are clearly visible as large excursions from the normal SST climatology (the zero line)", as opposed to
  - "(A) shows the SSTA anomalies in the Niño 3.4 region".

Data and code availability

- Important for reproducibility of research

Data and code availability

- Important for reproducibility of research
- Mention data sources

Data and code availability

- Important for reproducibility of research
- Mention data sources
- Provide code sources (Github repo, etc.)

Acknowledgments

- Mention funding sources

Acknowledgments

- Mention funding sources
- Sometimes people with whom you had informal but helpful discussions about the paper

Conflicts of interest

- E.g., if you are working for Shell and writing about climate change

References

- Use a reference manager

References

- Use a reference manager
- Make sure you cite classical papers as well

References

- Use a reference manager
- Make sure you cite classical papers as well
- Make sure you cite the most recent papers as well

References

- Use a reference manager
- Make sure you cite classical papers as well
- Make sure you cite the most recent papers as well
- Don't use "filler" references

References

- Use a reference manager
- Make sure you cite classical papers as well
- Make sure you cite the most recent papers as well
- Don't use "filler" references
- Keep to a minimum (length != intelligence)

Supplementary material and appendices

- Typically not essential to the main narrative but helpful

Supplementary material and appendices

- Typically not essential to the main narrative but helpful
- Results from additional experimental setups

Supplementary material and appendices

- Typically not essential to the main narrative but helpful
- Results from additional experimental setups
- Sensitivity / cross-validation / robustness checks

Supplementary material and appendices

- Typically not essential to the main narrative but helpful
- Results from additional experimental setups
- Sensitivity / cross-validation / robustness checks
- More background about the data and methods

Supplementary material and appendices

- Typically not essential to the main narrative but helpful
- Results from additional experimental setups
- Sensitivity / cross-validation / robustness checks
- More background about the data and methods
- Details about algorithmic and implementation aspects of your ideas

Section 2

Language and style

Overview

- Why is language important

Overview

- Why is language important
- Tenses

Overview

- Why is language important
- Tenses
- Grammar

Overview

- Why is language important
- Tenses
- Grammar
- Sentences

Overview

- Why is language important
- Tenses
- Grammar
- Sentences
- Paragraphs

Overview

- Why is language important
- Tenses
- Grammar
- Sentences
- Paragraphs
- Common errors

Overview

- Why is language important
- Tenses
- Grammar
- Sentences
- Paragraphs
- Common errors
- Source: https://www.gfdl.noaa.gov/wp-content/uploads/2018/08/Elements_of_Style.pdf

Why is language important

- Poor language can delay or even block publication

Why is language important

- Poor language can delay or even block publication
  - Reviewers need to understand your message with minimal effort

Why is language important

- Poor language can delay or even block publication
  - Reviewers need to understand your message with minimal effort
  - Reviewers need to understand your idea to be able to judge them

Why is language important

- Poor language can delay or even block publication
  - Reviewers need to understand your message with minimal effort
  - Reviewers need to understand your idea to be able to judge them
- Reviewing process does not involve language corrections

Why is language important

- Poor language can delay or even block publication
  - Reviewers need to understand your message with minimal effort
  - Reviewers need to understand your idea to be able to judge them
- Reviewing process does not involve language corrections
  - Authors' responsibility to convey ideas

Why is language important

- Poor language can delay or even block publication
  - Reviewers need to understand your message with minimal effort
  - Reviewers need to understand your idea to be able to judge them
- Reviewing process does not involve language corrections
  - Authors' responsibility to convey ideas
  - Not the reviewer's responsibility to read and re-read your draft

Tenses

- Preferable using present tense:

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"
  - **Do**: "We take SST data from ERA5"

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"
  - **Do**: "We take SST data from ERA5"
- If you have to use past tense, restrict it to describing experimental conditions:

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"
  - **Do**: "We take SST data from ERA5"
- If you have to use past tense, restrict it to describing experimental conditions:
  - "The simulations were run on a computing cluster with 120 nodes"

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"
  - **Do**: "We take SST data from ERA5"
- If you have to use past tense, restrict it to describing experimental conditions:
  - "The simulations were run on a computing cluster with 120 nodes"
  - "The proposed method was applied to all three data sets"

Tenses

- Preferable using present tense:
  - **Don't**: "We had taken SST data from ERA5"
  - **Do**: "We take SST data from ERA5"
- If you have to use past tense, restrict it to describing experimental conditions:
  - "The simulations were run on a computing cluster with 120 nodes"
  - "The proposed method was applied to all three data sets"
- Avoid shifting tenses in a basic unit of text: paragraph / section

Grammar

- Prefer active voice

Grammar

- Prefer active voice
  - **Don't**: "It was found that . . . "

Grammar

- Prefer active voice
  - **Don't**: "It was found that . . . "
  - **Do**: "We find that . . . "

Grammar

- Avoid the following:

Grammar

- Avoid the following:
  - **Contractions**: Use “do not” instead of “don’t”, “is not” instead of “isn’t”, . . .

Grammar

- Avoid the following:
  - **Contractions**: Use “do not” instead of “don’t”, “is not” instead of “isn’t”, . . .
  - **Abbreviations**: Except common abbreviations, any new abbreviations (if absolutely necessary) should be defined on first usage

Grammar

- Eliminate redundant phrases

Grammar

- Eliminate redundant phrases
  - due to the fact that -> because or since

Grammar

- Eliminate redundant phrases
  - due to the fact that -> because or since
  - immediately apparent -> apparent

Grammar

- Eliminate redundant phrases
  - due to the fact that -> because or since
  - immediately apparent -> apparent
  - in the case that -> in case

Grammar

- Eliminate redundant phrases
  - due to the fact that -> because or since
  - immediately apparent -> apparent
  - in the case that -> in case
  - and also -> and

Grammar

- Eliminate redundant phrases
  - due to the fact that -> because or since
  - immediately apparent -> apparent
  - in the case that -> in case
  - and also -> and
  - in order to determine -> to determine

Sentences

- Short, sharp sentences

Sentences

- Short, sharp sentences
  - Avoid long, compound statements

Sentences

- Short, sharp sentences
  - Avoid long, compound statements
- One idea per sentence

Sentences

- Short, sharp sentences
  - Avoid long, compound statements
- One idea per sentence
  - avoid multiple statements in one sentence

Sentences

- Short, sharp sentences
  - Avoid long, compound statements
- One idea per sentence
  - avoid multiple statements in one sentence
- Keep subject and verb close together

Sentences

- Short, sharp sentences
  - Avoid long, compound statements
- One idea per sentence
  - avoid multiple statements in one sentence
- Keep subject and verb close together
  - "An El Niño event, which is defined as an anomalous positive excursion of a three month running mean SST anomaly over the equatorial Pacific, typically causes a failure of the Indian monsoon in the following year."

Paragraphs

- Basic unit of conveying a complex idea

Paragraphs

- Basic unit of conveying a complex idea
- Prefer to have one idea / topic per paragraph

Paragraphs

- Basic unit of conveying a complex idea
- Prefer to have one idea / topic per paragraph
- Maintain the same tense all throughout

Paragraphs

- Basic unit of conveying a complex idea
- Prefer to have one idea / topic per paragraph
- Maintain the same tense all throughout
- Arrange paragraphs logically to create the ‘flow’ of your text

Common errors

- Avoid using “This” without qualifying it first

Common errors

- Avoid using "This" without qualifying it first
  - **Don't**: "We found this to be the most important facet of the ocean's dynamical response."

Common errors

- Avoid using "This" without qualifying it first
  - **Don't**: "We found this to be the most important facet of the ocean's dynamical response."
  - **Do**: "We found this feature of the thermocline to be the most important facet of the ocean's dynamical response"

Common errors

- Avoid too many successive prepositional phrases

Common errors

- Avoid too many successive prepositional phrases
  - **Don't**: "We ran a model simulation of the ocean for research into the evolution of the thermocline"

Common errors

- Avoid too many successive prepositional phrases
  - **Don't**: "We ran a model simulation of the ocean for research into the evolution of the thermocline"
  - **Do**: "We ran an ocean model simulation to conduct research into thermocline evolution"

Common errors

- Avoid subjective words / phrases that will get outdated over time

Common errors

- Avoid subjective words / phrases that will get outdated over time
  - “high resolution”

Common errors

- Avoid subjective words / phrases that will get outdated over time
  - “high resolution”
  - “new result”

Common errors

- Avoid subjective words / phrases that will get outdated over time
  - “high resolution”
  - “new result”
  - “latest finding”

Common errors

- Avoid subjective or judgemental adjectives

Common errors

- Avoid subjective or judgemental adjectives
  - "*simple* model"

Common errors

- Avoid subjective or judgemental adjectives
  - "*simple* model"
  - "*crucial* result"

Common errors

- Avoid subjective or judgemental adjectives
  - "*simple* model"
  - "*crucial* result"
  - "*fundamental* approach"

Section 3

The process of writing

Overview

- What are you writing? Who are you writing for?

Overview

- What are you writing? Who are you writing for?
- Find a narrative

Overview

- What are you writing? Who are you writing for?
- Find a narrative
- Writing

Overview

- What are you writing? Who are you writing for?
- Find a narrative
- Writing
- Editing and review

What are you writing? Who are you writing for?

- Know your demographic

What are you writing? Who are you writing for?

- Know your demographic
  - general / technical audience

What are you writing? Who are you writing for?

- Know your demographic
  - general / technical audience
- Research article or review or tutorial

What are you writing? Who are you writing for?

- Know your demographic
  - general / technical audience
- Research article or review or tutorial
- Graphical abstract or schematic explaining workflow / core concepts

Find a narrative

- My workflow:

Find a narrative

- My workflow:
  - First prepare the main result figures

Find a narrative

- My workflow:
  - First prepare the main result figures
  - Core narrative is, at this stage, simply figure description

Find a narrative

- My workflow:
  - First prepare the main result figures
  - Core narrative is, at this stage, simply figure description
- In general, ask these questions:

Find a narrative

- My workflow:
  - First prepare the main result figures
  - Core narrative is, at this stage, simply figure description
- In general, ask these questions:
  - What is interesting about your work?

Find a narrative

- My workflow:
  - First prepare the main result figures
  - Core narrative is, at this stage, simply figure description
- In general, ask these questions:
  - What is interesting about your work?
  - Interesting for whom (e.g., for climate scientists or computer scientists)?

Writing

- Write first, edit later

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then
  - Results and discussion, then

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then
  - Results and discussion, then
  - Conclusion, then

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then
  - Results and discussion, then
  - Conclusion, then
  - Introduction, then

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then
  - Results and discussion, then
  - Conclusion, then
  - Introduction, then
  - Abstract

Writing

- Write first, edit later
- I like to start with the methods and end with the abstract:
  - Methods, then
  - Results and discussion, then
  - Conclusion, then
  - Introduction, then
  - Abstract
- Use a spell check simultaneously

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well
- What to look for:

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well
- What to look for:
  - Unfounded generalisations

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well
- What to look for:
  - Unfounded generalisations
  - Results don't match the claims

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well
- What to look for:
  - Unfounded generalisations
  - Results don't match the claims
  - Research questions / hypothesis are not clearly laid out

Editing and reviewing

- Write a first draft, take a step back (a day or two), then edit
- Ask someone from a different group / field to review as well
- What to look for:
  - Unfounded generalisations
  - Results don't match the claims
  - Research questions / hypothesis are not clearly laid out
  - Figures are not easily understandable

Section 4

Miscellaneous issues

Overview

- Authorship

Overview

- Authorship
- Selecting a journal

Authorship

- Different disciplines have different criteria

Authorship

- Different disciplines have different criteria
  - Math: Mostly single authors

Authorship

- Different disciplines have different criteria
  - Math: Mostly single authors
  - Physics: Around 3-6 people

Authorship

- Different disciplines have different criteria
  - Math: Mostly single authors
  - Physics: Around 3-6 people
  - Geoscience: Large teams of 10-15 people

Authorship

- Different disciplines have different criteria
  - Math: Mostly single authors
  - Physics: Around 3-6 people
  - Geoscience: Large teams of 10-15 people
  - High energy physics: International conglomerates (100s of people!)

Authorship

- Typically, an author has **done** one or more of the following

Authorship

- Typically, an author has **done** one or more of the following
  - Conceptualised the study

Authorship

- Typically, an author has **done** one or more of the following
  - Conceptualised the study
  - Written the code

Authorship

- Typically, an author has **done** one or more of the following
  - Conceptualised the study
  - Written the code
  - Prepared the figures

Authorship

- Typically, an author has **done** one or more of the following
  - Conceptualised the study
  - Written the code
  - Prepared the figures
  - Written or edited the manuscript

Authorship

- Typically, an author has **done** one or more of the following
  - Conceptualised the study
  - Written the code
  - Prepared the figures
  - Written or edited the manuscript
  - Carried out related fieldwork or experiments

Authorship

- Typically, the following **does not** make you an author

Authorship

- Typically, the following **does not** make you an author
  - Providing feedback after a talk or in an isolated discussion

Authorship

- Typically, the following **does not** make you an author
  - Providing feedback after a talk or in an isolated discussion
  - Providing previously published data set

Authorship

- Typically, the following **does not** make you an author
  - Providing feedback after a talk or in an isolated discussion
  - Providing previously published data set
  - Providing previously published code

Authorship

- Typically, the following **does not** make you an author
  - Providing feedback after a talk or in an isolated discussion
  - Providing previously published data set
  - Providing previously published code
  - Reviewing someone else's manuscript

Selecting a journal

- Which discipline would benefit from your work?

Selecting a journal

- Which discipline would benefit from your work?
- Which journals come up more often in your own list of references?

Selecting a journal

- Which discipline would benefit from your work?
- Which journals come up more often in your own list of references?
- Which journals do your peers publish in?

Q & A

- Any questions or comments?