

Plagiarism in Science and How to Avoid It

ChatGPT

ChatGPT: yes or no?

Check and refine results

Check regulations on use

References are (currently) the worst part about ChatGPT ...

External vs. internal ethics

Rules of good scientific practice = research integrity

Research ethics

Research ethics: external

Good scientific practice: internal

Importance for science

Compliance topic? Boring?

John A. Ioannidis, “Why Most Published Research Findings Are False” (2005)

“Replication Crisis” (2011 ff.)

Growth of (secure, well-demonstrated) knowledge (epistemology ... Plato ... Aristotle)? Validity and reliability?

Extremely important for integrity of science!

Good scientific practice

Rules of good scientific practice (“RGP”)

Codes of conduct

retractionwatch.com

Abuse of power

False incentives

“Publish or perish!”

Funding proposals

ffp prevention

Core of the rules of good scientific practice: ffp prevention

F – falsification (misconduct regarding data)

F – fabrication (= misconduct regarding data)

P – plagiarism (= misconduct regarding attribution)

F: borderline cases ... where does it begin, where does it end?

STEM-humanities divide

Who gets into REAL trouble for what?

STEM-humanities divide, with social sciences in the middle

In STEM, people get punished for falsification and fabrication

In the humanities, they get punished for plagiarism

Why is that? Do STEM people not plagiarize? Do humanities scholars not falsify or fabricate?

RGP and QRP

“Below” RGP: Questionable research practices (“QRP”)

Ioannidis 2005

“Replication Crisis”

Popperian ethos vs. confirmation bias (“cargo cult science”)
working “lege artis”

Researcher degrees of freedom

IPR and plagiarism

Now for the p in ffp – plagiarism!

What is the difference between IPR / Copyright violation and plagiarism in science?

Think about politicians' plagiarism scandals? What happened to the politicians/what did not happen to the politicians?

Vanity/career PhDs in Law or Business Studies

What do you need to reference?

Answer 1: text!

Involuntary plagiarism?

Check for yourself (not necessarily now) ...

“Ο Ελευθέριος Βενιζέλος είναι ένα από τα πιο εξέχοντα πρόσωπα που συνδέονται με τα Χανιά”

“Eleftherios Venizelos is one of the most prominent persons linked to Chania”

No hits for either; “Eleftherios Venizelos” produced 1,380,000 hits, “Ελευθέριος Βενιζέλος” 1,320,000 this morning from Google

Criminal energy

The Cyrillic (and the Greek) letters k, t, o as plagiarism tools

Quotation marks white on white

Paraphrases

What about paraphrases/rewrites?

Do you need to reference them?

Yes, you do:

A recognizable rewrite that is not references makes things *worse*, not *better!*

Once it is *truly unrecognizable*, no problem ... as far as *text* goes.

What do you need to reference?

Answer 2: content!

Science is about proving/arguing points

Axioms

Types of proof:

logical

empirical

references = literature

The common knowledge problem

“Common knowledge” need not be referenced, unless ...
... you are using someone else’s *text*, remember?

But what is common knowledge? tricky ...

The common knowledge problem

What is the number of molybdenum in the periodic table?

Although maybe (most of) you did not know it, the answer, 42, is *common knowledge* ...

What is the capital of Burkina Faso?

Although maybe (most of) you did not know it, the answer, Ouagadougou, is *common knowledge* ...

Who was the foreign minister of Greece as it entered the then European Community on January 1, 1981?

Although maybe (most of) you did not know it, the answer, Konstantinos Mitsotakis, is *common knowledge*

The common knowledge problem

How many people are living on the island of Antikythera?

That is not common knowledge. Why not?

When was the alliance of Athens with Egesta during the Peloponnesian war?

That is not common knowledge. Can you guess why?

Common knowledge is clearly a misnomer:

Most of “common knowledge” is clearly *not* common knowledge!

The common knowledge problem

What does it mean? What are the characteristics of common knowledge?

Lack of controversy

Ubiquitous checkability

Fixed, binary “true or false” character

No perpetual changes

Of course, you need to be sure; if you do not give a reference, implying common knowledge, then we cannot accept mistakes!

Borderline common knowledge

Once you criticize others' positions (= the more so, the more so) ...

... or once an fact is very central to your argumentation ...

... or, on the contrary, a fact is very clearly outside your area of expertise ...

... it is safer to include a reference.

But do not sprinkle references, or even quotations if you are in the social sciences, everywhere: as many as necessary, as little as possible

Mopping up important special cases

Translation plagiarism

Figures/graphics/illustrations

Self-plagiarism

How is that even possible? What is meant by it?

How do you solve it?

Misrepresentation of literature

Many plagiarizers plead innocence: “Why, there *is* a reference”!

The reference also needs to be correct regarding its content and as clearly attributed as possible.

*Yanis Varoufakis is the most clever politician in Greece (Tsipras 2023a, b; Mitsotakis 2022).

*Some people think *Nea Dimokratia* is the best party for Greece, some people think it is *Syriza* (Mitsotakis 2022, Tsipras 2023a, b).

Sometimes people attribute their opinions to others; that is of course unacceptable, too.

Misrepresentation of literature

The same is true for extent.

Some people take many ideas, statements, arguments (and perhaps even text) from just one publication, but only put one reference of this publication in one paragraph.

Plagiarism researchers call this a “pawn sacrifice”; it is unacceptable.

There are many ways of demonstrating the extent of reliance on others properly.

Indirect references

Indirect references (often quotations) are another problem

What is an indirect reference?

*“Weizmann 1948:35, as cited in Eshkol 1966:123”

Two *good* ways to go about it. Which are they?

Get hold of Weizmann 1948 yourself!

If you cannot for good reasons, indeed use:

“Weizmann 1948:35, as cited in Eshkol 1966:123”

Indirect references

So, back to the example:

*“Weizmann 1948:35, as cited in Eshkol 1966:123”

Two *bad* ways to go about it. Which are they?

Eshkol 1966:123 (cutting out Weizmann 1948)

worse:

Weizmann 1948:35 (cutting out Eshkol 1966, but suggesting to readers you have used Weizmann 1948 yourself)

This means any mistakes Eshkol made in 1966 in describing Weizmann 1948 are now owned by you!

The chain of transmission

Does this mean we have to report the whole chain of transmission and growth of knowledge?

Like in Islam, you have the isnād of a ḥadīth
(Or in Judaism in the Mishna or the Talmud)

Of course not, that is generally totally undesirable and even impossible unless you are writing a piece on the history of science!

I agree with yesterday's takes:

Refer to the pioneers

Refer to the newest relevant research;
“complete, but not exhaustive”

Authoritative sources

How many people are living on the island of Antikythera?

That is not common knowledge.

What sources do you use?

When was the alliance of Athens with Egesta during the Peloponnesian war?

That is not common knowledge.

What sources do you use?

Disreputable sources

More dangerous than less-than-perfectly authoritative sources are disreputable ones.

What might they be (depending on subject)?

Wikipedia

Journalism, general websites, blogs, posts ...

Unpublished presentations, manuscripts

School textbooks etc.

Why are such sources a problem – and how do you solve it?

Use them and reference them

If you cannot reference them, do not use them

Exactitude

Many scholars obfuscate by not providing information as exact as it should be, and can be.

They e.g. might write:

*“The number of people living on Antikythera was 68 in 2011 (Hellenic Statistical Authority, *Census of Greece*, 35 volumes, Athens 2015).

Is this number “68” in all 35 volumes? We need a volume and indeed a page number!

Be as exact in ascription as the style permits.

Anti-plagiarism software

Anti-plagiarism software like Turnitin – is it good or bad?

It misses a lot of things (false negatives regarding plagiarism), but ...

... it also creates a lot of false positives.

Therefore it must be used responsibly.

Literature management software

Who is using a literature management software?

If you do not use one yet, my advice is: do.

Conclusion

The rules of research ethics and good scientific practice ...

... the rules on plagiarism avoidance ...

... the rules on attribution ...

... were not invented and refined to hinder or torture students.

They are necessary to ensure progress in striving for validity and reliability of knowledge.

Thank you very much
for your attention!