PLEASE READ BEFORE WRITING A DRAFT OF A PAPER OR PROPOSAL

First, please watch this video on common grammar mistakes: http://digg.com/video/yet-another-new-weird-al-track-word-crimes

Make absolutely sure that you do not lift sentences or paragraphs from something written by someone else and/or from something that you published previously. From the Elsevier Ethics in Publishing, Instructions for Authors document:

- "All reporting, writing, and editing that make up the content of the submitted paper shall be the original work of the authors and shall not plagiarize the work of others.
  - *Plagiarism* can mean the literal copying of the entirety of another's article or paper or other text.
  - *Plagiarism* can also mean the literal copying of large portions of another's work or even the substantive "paraphrasing" of another's work.
  - In all of these cases of plagiarism, the authors whose work is being copied or reproduced may also have legal claims with respect to copyright infringement or violations of their moral rights.
- Short quotes from the work of others is typical in the preparation of scholarly or professional manuscripts, but all such quotes should be properly referenced with full bibliographic details of the quoted work, as it is important to place the reported research or conclusions in a scholarly context.
  - Note that to quote or copy text or illustrations beyond a "short quote" will require the author to obtain permission from the rights holder."

---

1) Papers should generally be written in the past tense when presenting results. The exception to this might be the description of a crystal structure, where I find it hard to describe the structure in the past tense. But for other results, in a paper, you should use past tense.
From one website about writing scientific papers:
 Use present tense to report well accepted facts - for example, 'the grass is green'
 Use past tense to describe specific results - for example, 'When weed killer was applied, the grass was brown'

2) The word “data” is plural.
The data were…. CORRECT
The data was… INCORRECT

3) Use one space, not two, after the end of a sentence. See
http://www.dailywritingtips.com/one-space-or-two-at-the-end-of-a-sentence/

4) i.e. versus e.g.
http://grammar.quickanddirtytips.com/ie-eg-oh-my.aspx
I.e. and *e.g.* are both abbreviations for Latin terms. *I.e.* stands for *id est* and means roughly "that is." *E.g.* stands for *exempli gratia*, which means “for example.”

5) Use a comma after i.e. or e.g. (http://grammar.quickanddirtytips.com/ie-eg-oh-my.aspx)

6) “It’s” means “It is” and “its” is the possessive form of the pronoun it.
CORRECT: It’s the first day of its life.
INCORRECT: Its the first day of it’s life.

7) The word “comprised” should not be followed by the word “of,” but the word “composed” must be followed by “of.”
e.g., “A Quidditch team comprises three chasers, two beaters, one keeper and one seeker” or “A Quidditch team is composed of three chasers, two beaters, one keeper and one seeker.” It is NOT correct to say, “A Quidditch team is comprised of three chasers, two beaters, one keeper and one seeker.”

8) Should you start a sentence with the word “however”?
See http://www.quickanddirtytips.com/education/grammar/starting-a-sentence-with-however-right-or-wrong

Here are some quotes from the above website:

It is fine to start a sentence with however. You just need to know when to use a comma and when to use a semicolon.
The comma is important because however is a conjunctive adverb that can be used in two different ways: it can join main clauses and it can modify a clause.
If you use however at the beginning of a sentence and don't insert a comma, however means “in whatever manner,” “to whatever extent,” or “no matter how.”

For instance, Winston Churchill said, “However beautiful the strategy, you should occasionally look at the results,” and for those of you who like more modern examples, on the TV show House, Dr. Foreman said, “However bad you think you’re going to be in that room, not being there is worse.”

In both those cases, however isn't playing a role as a conjunction. It's not joining anything to anything else. It means “no matter how.” “However bad you think you’re going to be” and “No matter how bad you think you’re going to be” mean the same thing. I don't think anyone has ever disputed starting a sentence with however when it is used that way.

When you put a comma after however at the beginning of a sentence, everyone knows it means “nevertheless.” There's no reason to outlaw a perfectly reasonable use of the word when you can solve the problem with a comma. Some writers have even gone so far as to say it is preferable to start sentences with however instead of burying the word in the middle of a sentence, because putting it at the beginning makes the connection between sentences more clear and therefore makes the text easier to scan.

9) Use of “which” versus “that”
This is confusing and British English has different rules than American English, but most people incorrectly use “which” in American English when they mean “that.”
Read http://grammar.quickanddirtytips.com/which-versus-that.aspx
before you start writing a paper or your thesis.

10) Be careful with adjectives – what might strike you as “very” different, “extremely” large, or xx (insert your favorite phrase with an adjective) may not seem that way to someone else, so just say that it is different or it is large. I usually remove unnecessary adjectives from paper drafts. I also usually avoid the words “unique” and “novel” (note that, by definition of these words, it is impossible for something to be “very unique” or “very novel”). Also, be careful about the word “significant.” Some journals will not allow the word “significant” to be used in a paper unless you include statistics establishing significance.

11) Quote from Russell Doolittle’s book: “Probably no word causes more confusion in this field than the word “homologous”. When two sequences are homologous, they share common ancestry. In this sense, there are no degrees of homology. Sequences are either homologous or they are not. Many investigators use the word when the mean "similar". Two sequences may be

similar by chance, for example. They may resemble each other to a high degree, but they ought not to be very homologous or slightly homologous. Also, two sequences may be 60% identical, but they are not 60% homologous."

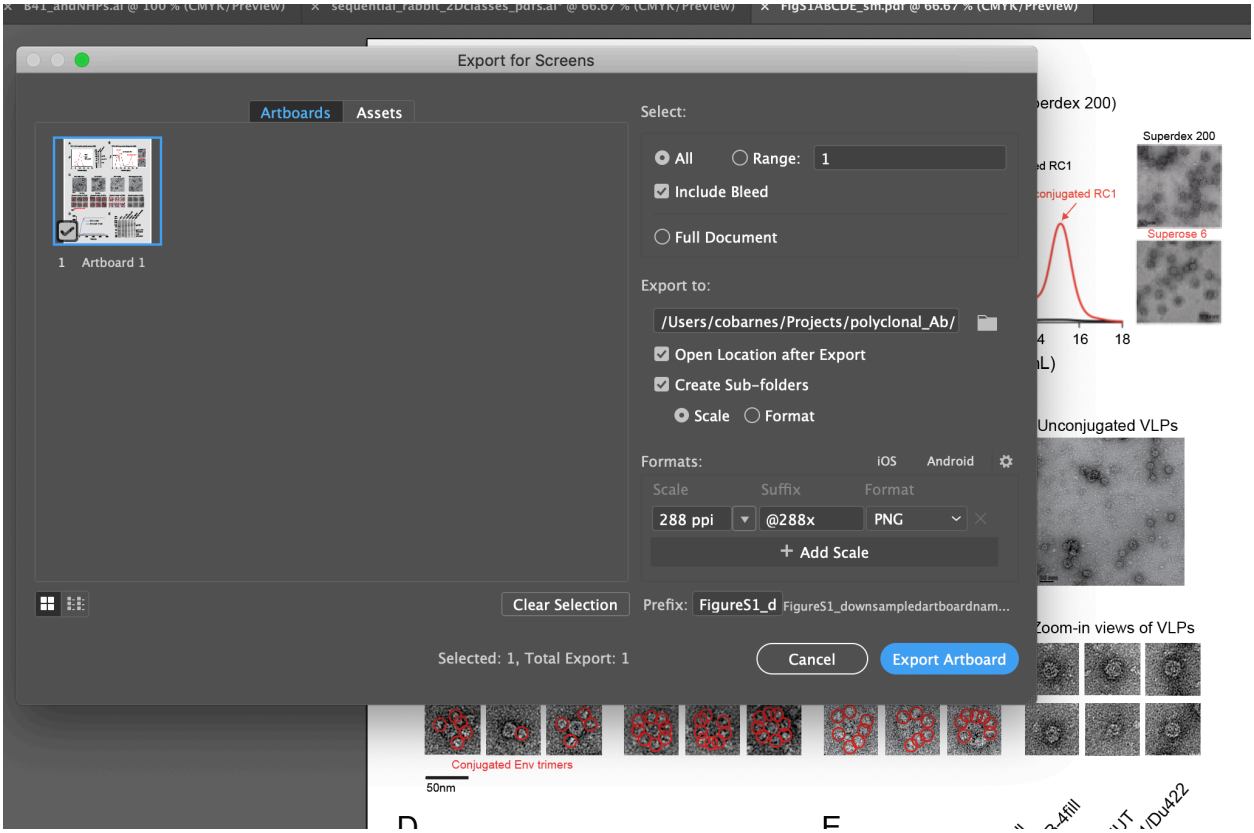
See http://www.theallium.com/science-life/evolutionist-loses-it-as-colleague-conflates-homology-and-similarity-yet-again/ for possible consequences when someone gets this wrong.

12) To report equilibrium dissociation and kinetic constants, use approved nomenclature: K_D (uppercase K in italics and uppercase non-italics D as subscript) for affinity; k_a and k_d for on- and off-rates (lowercase k in italics and lowercase non-italics a or d as subscript). Remember to report units (molarity for K_D; $M^{-1}s^{-1}$ for k_a; s^{-1} for k_d.

13) Don't create figures in Powerpoint. Many journals won't accept powerpoint figures even if they've been converted (e.g., from the J Exp Med website: Please note that files saved as TIFF or EPS from within the PowerPoint application are NOT at sufficiently high resolution to meet our formatting requirements.

14) **How to reduce the size of Illustrator files to put them into a Word document for submission to a journal:** First, save the full-size file as an Illustrator pdf in the folder where you're storing your files. When you're ready to insert a figure into a Word file, open the pdf using Illustrator, go to File -> Export -> Export for screens, and then save images as pngs and manually change resolution from 288 ppi to 150 or 96 ppi (changing resolution is done under Scale/resolution).

In Word, you can also use the menu option under File that says Reduce file size to make a Word file small enough to email.



15) If you are including a crystal structure in your paper, note that you MUST have refined that structure to have Rfree and Rwork values that are in the top 50% for the resolution that you are reporting. Check out other structures in the PDB at your resolution and make sure that your statistics are in the top 50% and make sure that you have refined to minimize Rfree, not Rwork (i.e., the difference between Rfree and Rwork should not be unacceptably large). The following papers discuss common problems in crystallographic refinement that result in high Rfree values (e.g., not using non-crystallographic symmetry restraints/constraints at low resolution, refining individual B factors at low resolution, etc.)

1) Bränden & Jones (1990) Between objectivity and subjectivity.
Nature 343: 687-689. (written before R_{free} commonly used)
2) Kleywegt & Jones (1995) Where freedom is given, liberties are taken.
Structure 15: 535-540.
3) Kleywegt & Jones (2002) Homo Crystallographicus -- Quo Vadis?
Structure 10: 465-472.
4) 2017 paper: Validation of structures in the Protein Data Bank: https://www.sciencedirect.com/science/article/pii/S0969212617303374

16) If you are reporting buried surface area (BSA) for an interaction, specify whether the number you are reporting refers to the entire interface or to the BSA on one of the interacting partners.

17) If you are calculating a root mean square deviation (rmsd) after superimposing structurally-related proteins or domains, this must be calculated on carbon-α atoms only (NOT ON ALL ATOMS) and you must say how many carbon-α atoms were used for the calculation.

18) Example of what should be included in a Methods section of a structure paper (but don't copy this exactly – that would be plagiarism since this is from a lab paper).

Buried surface areas were calculated using PDBePISA (Krissinel and Henrick, 2007) and a 1.4 Å probe. Potential hydrogen bonds were assigned using the geometry criteria with separation distance of <3.5 Å and A-D-H angle of >90°. The maximum distance allowed for a potential van der Waals interaction was 4.0 Å. Protein surface electrostatic potentials were calculated in PyMol (Schrödinger LLC). Briefly, hydrogens were added to proteins using PDB2PQR (Dolinsky et al., 2007), and an electrostatic potential map was calculated using APBS (Baker et al., 2001). Epitopes for antibodies in Figure 3 were identified as gp120 residues containing an atom within 4 Å of an antibody as calculated in PyMol (Schrödinger LLC).

19) Please read the instructions for authors for the journal to which you'd like to submit the paper so that you know whether they want you to say Figure 1A, Fig. 1A, Figure 1a, Fig. 1a or some other permutation, length restrictions for various sections of a paper, order of presentation, and other stylistic rules that differ from one journal to another.

20) I would prefer that people prepare drafts of papers using Arial 11 as the font because that font is allowed in NIH grants and I sometimes copy sections from a paper into a grant (Times and Cambrian fonts are not allowed for NIH grants).

last updated: 11/29/2021