Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Research papers 101: The do's and don'ts of scientific writing

Daniel Torres-Valladares *, Elvira Ballinas-García, Jessica Villarreal-Reyes, Valeria Morales-Álvarez, Carlos Ortiz-del-Ánge

Tecnológico de Monterrey, Escuela de Ingeniería y Ciencias, Ave. Eugenio Garza Sada 2501, Monterrey, NL 64849, Mexico

ARTICLE INFO

Article history:
Received 29 October 2020
Received in revised form 12 December 2020
Accepted 16 December 2020
Available online xxxx

Keywords:
Scientific Writing
Journal Publishing
Scientific Research

ABSTRACT

Writing and presenting results is important for any research work. Publishing is not enough; it should be written clearly and consistently for others to read and cite. We present a review of different techniques, for improving the writing in research papers intended for publishing. These include how to write the title, whether to use active or passive voice, how long sentences should be, how to present the uncertainty, understanding the International System of Units (SI) writing conventions, how to assign authorship, how to effectively use references, how to present the work, and how to publish it in a journal. This paper is a summary of the main and most recent recommendations in the scientific field on how to write a research paper to increase its impact.

Selection and Peer-review under responsibility of the scientific committee of the Tec.Nano 2019.

1. Introduction

Papers are a crucial part of research; if research does not produce published papers, it remains incomplete [1]. Proper writing is needed to get the results published. Furthermore, your work must be cited by others to be impactful [2]. Rejection rates are increasing for journal publishing [3], thus we have compiled a set of different writing techniques and recommendations for creating well written scripts. Writing better will facilitate getting past the copy editor, making it easier for your paper to arrive to the editor and the reviewers [2].

Many times, researchers refuse to invest time in reading about ways of improving their writing style [2]. However, as we will see, there is statistical evidence which supports that utilizing certain methods when writing an article can make it more understandable and successful in the journal rank. Two questions naturally arise: which is the strategy to follow to present and write about your research well? And, how does this strategy affect the scope of the article? We answer them by presenting a compilation of the most valuable information to make it readily accessible for authors looking to improve their writing.

The writing style can make a paper much more concise and easier to understand, but sometimes we use styles that might not be ideal. For instance, there is a strong preconception among engineers that the impersonal voice is more formal than the active voice [4]. However, the scientific journal Nature guidelines for How to write a First Class Paper explicitly state that "We should engage readers' emotions and avoid formal, impersonal language.". This can be achieved by use of the active voice (we performed the experiment...) [5].

In this review, based on the series of articles "How do you write and present research well?" by Gregory Patience, Daria Boffito and Paul Patience, we will cover the basics on how to do the presentation of your work. Then, we will dive into the specifics of strategic writing. We will discuss the use of active and passive voice when describing procedures and results, as well as the importance of the title and the words used in it. In terms of semantics and syntax, the length of sentences and the signs in writing, which include the use of words in hedging, will be discussed. We will also examine how to show the uncertainty in the results and how to present it, be it in graphs or as a ± sign after the result. The SI units writing conventions will also be explained, as well as authorship assignation and the proper use of references. At the end, we discuss how to submit your work to a journal.

2. How to present your work

Doing a presentation of your work is another crucial part of research. You should be able to communicate clearly and concisely what you did and what you found. In this section we will cover

* Corresponding author.
E-mail address: A01039732@exatec.tec.mx (D. Torres-Valladares).

https://doi.org/10.1016/j.matpr.2020.12.708
2214-7853/
Selection and Peer-review under responsibility of the scientific committee of the Tec.Nano 2019.

Please cite this article as: D. Torres-Valladares, E. Ballinas-García, J. Villarreal-Reyes et al., Research papers 101: The do's and don'ts of scientific writing, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2020.12.708

what to say and how to say it when you are giving an oral presentation.

2.1. What to say

Prepare to say less than you prepare is a great piece of advice [6]. People will need time to grasp all the information you give them, so you should try to ease them into the subject and not overwhelm them with technicalities. Know your audience, organize your content to fit their profile. Make your research accessible to them. You will always need to give a certain amount of context. Even if you are talking to specialists, "not everyone is familiar with your expertise" [6]. If you are giving a presentation for a multidisciplinary audience, save the technicalities for the poster session, and, if you are talking to the general public, use analogies instead of technical vocabulary. Remember, the goal is for people to understand your work. Try to adapt your presentation to the several learning styles to maintain the audience engaged.

As a general structure, present yourself as an expert, motivate your audience and support your claims with data. Try to use more images and graphs on your slides and, contrary to the recommended format for your research paper, repeat your message and explicitly state what you are going to talk about.

Tavares recommends not wasting your time by introducing your presentation with the generic structure: context, objectives, literature review, etc., but rather using an executive summary like a graphical abstract on the second slide [6].

Start by giving both the problem and the solution to interest your audience. Elaborate on the context, but be selective on what to include in the presentation, is it necessary for your message? Summarize and confirm your ideas on the last slide. You may repeat the executive summary. Do not introduce new data when you give your conclusion and remember to acknowledge all the participants and supporters of your work.

2.2. How to say it

Be careful with the pace of your words. Several researchers support the idea that the attention span of people is between 10 and 15 min [7]. When giving a presentation, you will encounter the difficulty of the audience's inability to focus. If you talk too fast or too slow, you risk losing your audience's attention. The right cadence for your words is between 120 and 160 words per minute [6].

Use well organized slides to complement your presentation. It is easier to follow numbered slides and tags. Replace bullet point lists with images and animate their entrances to avoid overwhelming the audience. Keep your slides simple: animate for emphasis and avoid long sentences. Favor graphs over tables, but, if the table is essential, limit it to a maximum dimension of 6×6 and try to use colors, bold fonts and frames.

The way you talk will make a huge difference in your presentation. Emphasize important parts by repetition and 1 or 2 s pauses. Avoid looking at your notes and slides as much as possible. Instead, focus your attention on the individuals in the room. Move around the stage calmly instead of remaining on a fixed spot. Keep a good posture and gesture to look confident and enthusiastic [6].

3. Main body manuscript's sections

3.1. Title

The title is the first thing they will see of your research paper and perhaps the only part they will read. A good title should describe the content of the paper with the less possible words [8]. The objectives, subject and the result should be included. It needs to show the main idea with the least number of words to optimize the search engine results [9]. The optimal length for a title is recommended to be between 8 and 15 words. The 500 top-cited articles have less than 19 words in their titles. Unnecessary words will make it more complicated for people to understand what your is work about; we list a few of them:

- Investigation of
- Observations of
- Assessment of
- An opening: A, An or The

These words just work as "fillers" to make the title longer. For example, in the article A detailed analysis of the KDD CUP 99 data set [10] the expression "A detailed analysis" does not give any crucial information about the research. The authors proposed a new data set (NSL-KDD) that solves some issues of the KDD CUP 99 in Intrusion Detection Systems (IDSs). This is not mentioned in the title, and neither are their results nor their objectives, so it becomes ambiguous. A more appropriate title could be "NSL-KDD data set for network-based anomaly detection systems", which results to be longer but also gives a more complete description of the work done. Titles should be short without omitting the main idea of the paper.

The title is a label not a sentence as is usually thought. The order of the words is utterly important [8]. People might not be able to find your article if the words are not chosen correctly.

3.2. Types of abstracts

Abstracts in Nature and Science have the following structure: research subject, description of the problem, contributions. A common mistake in writing abstracts is repeating the title in the first sentence [11]. An abstract should be viewed as a resumed version of the paper [8]. Usually, researchers do not fail to choose the type of abstract to be used but to include the right information. Another common mistake is mentioning just the expected results instead of also disclaiming the actual results of the experiment.

Abstracts can be classified as indicative, informative and graphical.

3.2.1. Indicative abstracts

This kind of abstract is designed similarly to a table of contents, indicating all subjects in the article, making it easier for potential readers to decide whether to read the paper or not. However, it can hardly be seen as a miniature version of the paper. Thus, indicative abstracts should not be used in research papers. They can be used in other types of context such as review papers, conference reports, and government reports [8].

3.2.2. Informative abstracts

This type of abstract condenses the paper. It should briefly state the problem, the method used for the study and the conclusions. It conveys an entirely accurate picture of the research and is the most common type that you will find in journals [8].

Graphical abstracts (GA) are like mini posters published in some journals to get the attention of readers. They serve to highlight the most notable features of a paper using mostly images. They are most common in Chemistry journals [12]. Since a GA is like a poster, it should be self-explanatory and eye-catching, so use colors, arrows, and images to interest the reader. Repeating the title, writing large blocks of text, and using tables would not serve the purpose [13]. Either an image or a very specific graph that depicts your results can serve as a graphical abstract [14].

You should make sure that all your figures and text follow the specifics of each journal's "Author guidelines", fitting in the specific window dimension and respecting the typography. A couple of common dimensions are squares and 5×2 rectangles [13]. Keep in mind that designers will modify your figure size to fit the window and will oftentimes convert it to a lower quality image. To keep your text legible after resizing, setting capital letters to be 5% of the vertical height of the image is recommended [13].

3.3. References

The quantity and quality of references in an article has an important impact in the reach of your publication. Nature journal authors reference recent research, and on average Nature articles are cited 41 times after two years of publication. This journal has an average impact factor (IF) of 41, which makes it one of the most important journals in scientific literature [15]. Articles that predominantly cite papers older than 2 years contribute little to a journal's IF.

3.3.1. References in the introduction

The introduction is usually where the incentives for writing the article are exposed. These could be economic, historical, scientific, etc. Reviewers looking at the article appreciate well written introductions that reference important and updated literature. Researchers should credit the research done previously over which they are working by citing it directly and replace sentences like "it is suggested" with a direct reference [16]. We recommend to overuse citations rather than under-use them for the following reasons [15],

1. Reviewers will be severe if their work is not properly cited.
2. Lack of proper literature revision could be plagiarism.
3. More people are likely to read your work.
4. More people are likely to cite your paper.

3.3.2. How many references to use?

It depends on the nature of the article since some disciplinary areas normally utilize more citations than others. For example, Mathematics and Philosophy papers usually have less citations than medicinal or cell tissue engineering papers [15]. In his paper, S. Patience researched the relation between the number of references and the frequency with which papers are cited. Research showed that a higher fraction of articles with fewer than 20 references are cited less times than the ones with + 40 references [15].

3.3.3. How old should papers be?

Citing more improves the chances of people seeing your work, and citing recent papers makes it more likely that people who are active in the field will see the work. The references in the Nature journal as of Jan 2015 frequently cite articles from 2011 to 2013, giving it an 8.2 scale parameter of age of references, each considered article was cited 100 times. On the other hand, the top 500 chemical engineering articles average 140 citations and have an age scale parameter of 6.2. These articles, besides citing more recent work, also include more references [15].

4. Elements of information

4.1. The use of error bars

To this date, there is still not a standard way to express error bars in graphs or the uncertainty in a measurand (the value following the ± symbol in a measurement), and not everyone includes them [17]. Moreover, even if some authors do, they often forget to specify either the sample size, *n*, the *p* value, or what the error bars or the uncertainty represent altogether. Do they portray the standard deviation, the standard error of the mean, or the confidence interval? If this information is not specified, it becomes impossible for the readers to interpret the error bars and what it means when they overlap.

To standardize the meaning and presentation of error bars in graphs and the uncertainty in a measurand, the Bureau International des Poids et Mesures has proposed some rules [18] that deal with the expression of measurement uncertainty. One of their recommendations is to add an indicator that tells the measurement quality or uncertainty.

4.1.1. Measurement uncertainty

The most widely used measures of uncertainty are the standard uncertainty (u, s.d., SD, s, σ), the standard error of the mean (s.e.m, SE, s_n), the expanded uncertainty (U, Cl, Δ), and the combined uncertainty (u_c) [18]. The standard deviation (SD) represents the inherent variability present in data. On the other hand, when few repeated measurements are made, the standard error of mean (SEM), becomes more relevant. Finally, confidence interval (CI) is a general expression the third equation shown in table 1, where one chooses a certain confidence level (95% and 99% are typical values). The equations and most common symbols of the discussed quantities are shown in table 1.

A confidence interval indicates a range of possible values for a given population parameter being estimated [19]. It is done with a t test, where the p value indicates the probability that two sets of measurements belong to the same parent population.

"A p value is the probability, under a specified statistical model, that a statistical summary of the data (for example, the sample mean difference between two compared groups), would be equal to or more extreme than its observed value" [20]. When the p value is less than or equal to the significance level, α, chosen by the researcher, the null hypothesis is rejected, meaning that the measurements are different. Nonetheless, the scientific findings cannot be rejected just because we rejected the null hypothesis (error bars do not overlap) since rejecting it does not mean that the means are different [21].

4.1.2. Error bars

An error bar represents a range of values where we expect to find a measurement, at a chosen level of confidence $(1-\alpha)$. Its main use is to show whether data is statistically different or not. To correctly interpret the statistical significance of error bars, we must state what they represent (SD, SEM, or CI). When using CI and SD error bars, if the bars touch, given that $n > 3$, the null hypothesis can be rejected. On the other hand, when SEM error bars touch, the p value is greater than 0.05 for all n, so the null hypothesis cannot be rejected [17]. Error bars that represent confidence intervals are recommended.

4.2. Stating the uncertainty

As mentioned in the previous section, an indicator of measurement quality is the uncertainty [17]. The most used metrics that researchers assign to the uncertainty are the standard deviation (σ, or s the sample standard deviation), the standard error of the

Table 1
Relevant quantities in measurement uncertainty, their common symbols, and their equations.

Quantity Name	Common Symbols	Equation
Standard deviation	u, s.d., SD, s,σ	$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2}$
Standard error of the mean	s.e.m, SE,s_n	$s_n=\frac{s}{\sqrt{n}}$,
Confidence Interval	U, Cl, Δ	$\Delta=t(\alpha,n-1)\frac{s}{\sqrt{n}}$,

mean (s.e.m), or the confidence interval (CI). However, if the uncertainty, x, is not added, it is assumed that it is equal to half the value of the last significant digit. On the other hand, if you present the uncertainty with the data, you need to explicitly state what it represents (SD, SEM, or CI). As a rule, the uncertainty should not be greater than the last significant figure of the measurement [22].

It is recommended that, to represent the uncertainty, the confidence interval is used. As presented in the previous section, the confidence interval is the product of the SEM equation (see table 1) and a coverage factor, which, for $n < 30$, is equal to the t statistic, and for larger samples the z statistic is used.

One way to reduce the uncertainty is to increase the sample size [22]. The most common method used to determine the number of experiments is to assume a value of n, then read the t statistic from a table or program and iterate to minimize the following equality:

$$\sqrt{n} - t(\alpha, n-1)\frac{s}{\Delta} = 0 \quad 1$$

4.2.1. Data types

The data can be categorized into 5 distinct classes [22]:

1. Pure physical quantities (distance, time, temperature),
2. Data recorded by an analytical instrument (concentration, surface area),
3. Derived (activation energy, selectivity),
4. Model parameters (coefficients in equations),
5. Statistical descriptors (standard deviation, mean, F value).

In most of the published papers, the correct number of significant figures is used for physical quantities. However, it is rarely stated if this uncertainty represents the SD, SEM, or CI. On the other hand, for the other data types, too many significant figures are included. In the following section, the appropriate number of significant figures for each data type is going to be presented.

4.2.2. Direct physical measurements

Temperature, T: Although it is derived from instruments that convert signals (ex. electrical) into a reading calibrated against a standard, temperature readings can vary as 10–20 °C during experiments. Therefore, it is considered excessive reporting 5 significant figures for temperature [22].

Pressure, P: Since it changes with meteorological conditions, 3 significant figures for pressure are often enough [22].

Porosity, ϕ: Because materials with porosity are anisotropic, ϕ will vary with each sample's position and size. Therefore, 3 significant figures are recommended [22].

4.2.3. Analytical data

Analytic data is obtained from analytic instruments that measure, for example, surface area, viscosity, particle size, and concentration, and report a single value. There are many factors that influence this value, such as instrument characteristics, calibration, and environment.

Surface Area, SA: Like porosity, surface area is anisotropic. Repeating the measurement decreases the uncertainty of the measurement. Even though instruments report surface area with a precision of 0.01 $m^2 \cdot g^{-1}$, measurements can vary as much as 1 $m^2 \cdot g^{-1}$. Therefore, 2 significant figures are enough [22].

Viscosity, μ: Since it varies with temperature, the number of significant figures must reflect the precision of the end use. It is therefore recommended to use 2 significant figures.

4.2.4. Derived data

Just like the analytic data, derived data depends on several factors. Their difference resides in that the instruments integrate all the factors and then just report one value (composite value) [22]. However, all these factors have their own uncertainty. Therefore, for a composite measure, f, that depends on factors $x_1, x_2, \cdots, x_n$, the combined uncertainty, f, depends on the individual uncertainties of each factor, x:

$$\Delta_f^2 = \left(\frac{\partial f}{\partial x_1}\Delta_1\right)^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\Delta_n\right)^2, \quad 2$$

4.2.5. Model parameters

Sometimes, researchers apply statistical, empirical, and engineering models to characterize experimental data. An example of a statistical model is heat capacity, where the model relies on a fourth-order polynomial in temperature. On the other hand, empirical models integrate physical observations into mathematical expressions, and engineering models start from first principles to derive simplified expressions. Whatever the model, they rely on experimental data that are uncertain. Therefore, the fitted parameters of the models must reflect that uncertainty [22].

Fitted parameters must have the same number of significant figures so that readers know that they contribute equally to the uncertainty in the final value. To this end, a single digit is enough to report the uncertainty in the fitted parameters. As well, the model coefficients should be limited to 3 significant figures, as they should reflect the uncertainty in the experimental data [22].

4.2.6. Statistical data

% deviation: A single digit, equal in magnitude to the most precise (smallest) digit of the associated number, is enough to express the uncertainty (Δ) [22].

R^2: If R^2 <80%, then it is recommended to use no more than two significant figures. If R^2 >98%, then it should be expressed with more than 3 significant figures. Finally, if R^2 >99%, include as many as 4 significant figures [22].

$\bar{x}, \mu$: The number of significant figures of the mean is calculated based on the magnitude of s. That is, whether the data varies within ± 50% or over several orders of magnitude [22].

s: Since s should not be reported with a greater precision than $\bar{x}$, it should rarely have more than 2 significant figures [22].

4.2.7. Rounding and truncating

NIST [23] recommends rounding converted values to the same number of digits of the original value and to choose prefixes such that the data lie between 0.1 and 1000.

4.3. SI writing conventions

There is an undeniable need for a standard that facilitates the understanding of units and symbols. In 1948, the 6th resolution of the CGPM instructed the CIPM "to make recommendations for a single practical system of units of measurement, suitable for adoption by all countries adhering to the Mètre Convention" [24]. That was the birth of the consensus that today outlines the format of modern unit writing.

Just as there are several ways to print a character [25], there are several rules for printing symbols and names of units [23]. They can be summarized as follows [26].

4.3.1. Units

Unit symbols are always printed upright (roman style), preceded by a separator, and, unless they come from a proper name, they must be lower-case letters. However, to avoid confusions between the lower-case letter *l* and the number 1, the 16th CGPM accepted the use of the capital-letter L for liter [26]. Since symbols are mathematical entities, they should not be followed by a period unless it is the end of the sentence, nor should they be turned to the plural form. For instance, centimeters are still represented by cm instead of cms. Algebraic rules apply in the operations of symbols. Multiplications will be expressed via a space or a half-centered dot and divisions, by either a horizontal line (-) or a solidus (/). If more than one solidus is used, it is necessary to use brackets.

Unit names are treated as nouns. Thus, their position in the sentence and their classification (proper or improper) will define whether they should be capitalized or not. A name that results from the juxtaposition of various names needs a separation, either by space or hyphen.

The use of abbreviations for both names and symbols is unacceptable; they should always be fully spelled. Furthermore, when combined with multiple or sub-multiple prefixes, neither the unit symbol nor the unit name needs a separator.

4.3.2. Quantities

Although the symbols for quantities are just recommendations, their format is still regulated by the consensus. Normally, quantity symbols are printed in *italic* and further identified by their subscripts or superscripts; for instance, the external pressure, p_{ext}. All extra information should be indicated on the quantity symbol and not on the unit symbol.

Moreover, their numerical values should precede the unit symbol with a separator in between. The sole exceptions to this rule are the symbols of degree, minutes, and seconds for plane angles. These are also exempt from the rule of having a single unit symbol per expression [26].

4.3.3. General format

Generally, anything that is well defined (a unit symbol, numbers, descriptive subscripts and superscripts, mathematical operators, or well-defined functions like the trigonometric functions) should be printed in roman. On the other hand, anything for which we can choose the value according to the situation, is written in italic, for instance, variables, vectors, running subscripts, undefined constants.

The most recent issue of NIST, includes the latest changes on the BIPM SI brochure and shows the standard practices for writing in the United states [26]. The main specifics are the English spelling follows Webster's Third New International Dictionary, and the unit t is now called a "metric ton".

4.4. Designing graphs

Understanding typography, like font types, typefaces, character height, etc., can help create an aesthetic graph [27]. A well-made, efficient graph or illustration communicates ideas, data, statistics, and relationships with minimal effort and far more effectively than tables and words do [28]. According to Tufte [29], graphical excellence maximizes data-ink ratio and is a matter of:

1. substance, statistics, and design;
2. clarity, precision, and efficiency;
3. communicating the maximum number of ideas with the least ink;
4. displaying the information honestly.

The following guidelines for line weights in plots respect the Can. J. Chem. Eng. typographical style. Kamat et al. [30] present complementary strategies to plot data, present results accurately, and define axes correctly.

If the journal's style is considered while designing graphs and illustrations, the result will be more aesthetically pleasing. Articles published in Can J. Chem. Eng. Have a 10pt Times New Roman typeface and the lines in tables are 0.75pt. Therefore, the text included inside graphs should be close to but not bigger than 10pt, sans serif. It is also recommended that axis lines be 1pt [28]. Overall, the whole process of designing a graph consists of two steps: establishing the graph's frame width to fit in the column, which is 91 mm wide; and setting the line weights and font size to match the journal. A guideline for recommended line and symbol dimensions can be found in [27]. Fig. 1 presents an example on what the graphs should look like based on the previous recommendations and on the guidelines presented below.

4.4.1. Frame dimensions

Although the column width is 91 mm, graphs should extend less than 85 mm from the y-axis label on the left to the end of the x-axis on the right (or the right y-axis label if there is one) [28]. To maximize the length of the x-axis, orient the y-axis labels vertically. In addition, choose prefixes that minimize the number of characters in the axis labels.

4.4.2. Axis length and line weight

If the graph has a single y-axis label, the x-axis length must be 70 mm. To maximize the data-ink ratio, one can remove the axis lines on the top and right sides. It is also recommended that the height of the y-axis lies between 50 and 70 mm [28]. For the line weight, all axis lines should be 1pt so they are slightly heavier than table lines.

4.4.3. Ticks

For the tick to be visible, it needs to be twice as it is wide. Therefore, a 1pt tick is recommended together with a gap in the axes to represent a tick [28].

4.4.4. Trend line

For experimental measurements and discrete data, use symbols. On the other hand, for mathematical models and continuous data, use lines. If these lines connect the dots (trend lines), then it should be stated in the caption. It is recommended that the line weight of the trend lines be 80–90% of the axes' weight [30].

4.4.5. Symbols

Colored symbols are recommended since they are clearer, even when printed in black and white. Light colors are adequate if

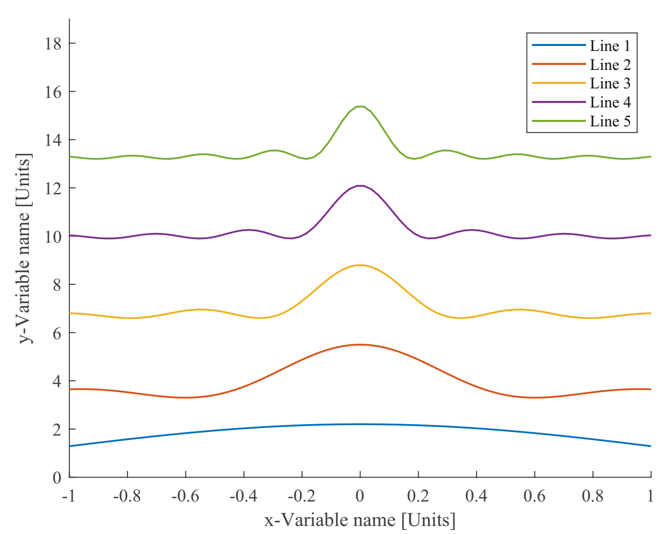

Fig. 1. Presents an example of a correct format for a graph based on the guidelines presented above.

accompanied with contour lines. It should be explicitly stated in the caption if the dimensions of the symbols represent the uncertainty. Apart from that, the size of the symbols must be at most the size of the text. Smaller symbols are recommended when there are hundreds of data points [30].

4.4.6. Confidence intervals

As previously mentioned, when error bars are presented in graphs, it should be explicitly clarified if they represent confidence intervals [22]. For this matter, standard line weights should be 0.5–0.75pt.

4.4.7. Bar charts

Bar charts should be replaced with tables when they require very long texts for the legends.

4.4.8. Grids

The main objective of graphs is to communicate concepts, not to derive precise numerical values. If grid lines are added to the graphs, it makes it more confusing to identify the essential information. Tufte [30] refers to grids as chart junk that carry no information. However, if grid lines must be added, then they should be no more than half the axis line weight and grey rather than black.

4.4.9. Excel

With some work, graphs that respect the previously presented guidelines can be generated in Excel. Excel assigns minimum and maximum axes reasonably well.

5. Elements of disclosure

5.1. Use of active voice

Currently many books/guides on scientific writing advise to use the active voice whenever possible. Nature and Science journals both prefer the active voice which normally uses the authors as agents of action [31]. In the abstracts, one in four sentences contains the word "we" or "our". Day and Sakaduski's writing guide for scientific English [32], states directly — "the passive voice should be avoided". The main argument is that it helps to make the writing clear and concise, contrary to a major part of the 20th century when the passive voice was characteristic of scientific writing [33].

The journals that encourage the active voice have less incidents of certain phrases. The past progressive form e.g., "it was observed that", "it was found that", "it was shown that" appears 70 times in the 500 most recent articles in Can. J. Chem. Eng. This form is almost absent in Nature. The present perfect progressive form is also uncommon in Nature: "it has been shown that" (3 times), "it has been found that" (twice), "and it has been observed that" (once) [31].

The active voice "we" takes responsibility for the research performed, the passive voice does not. Contrary to the belief of some, the active voice is as equally biased/unbiased as the passive voice, and personal pronouns are as objective as impersonal ones. The difference is that sentences become less vague; it is clear who did the work and who made the assessment [3]. Sainani recommends using the active voice as default and the passive voice only when there is a reason. In the top 500 abstracts in Nature in 2013, just 20% of the sentences were written in passive voice [34].

5.2. Replacing implicit agent with an explicit agent

Well-formulated active sentences are shorter than passive ones. Moreover, although "we" is a better agent than the implicit, sometimes it is not the best agent for an active sentence and it is better to identify who performed the action and make them the agent [35]. The best agents are the subjects who convey the action or state being expressed by the verb.

Here is an example.

1. We measured the pressure periodically.
2. The pressure was measured periodically.
3. It was shown that the pressure varied periodically.
4. Pressure varied periodically.

Out of these sentences, the best one is number 4 [35]. It contains the most amount of information in the shortest sentence. In sentences 2 and 3, the agent is implicit, and in the case of 2, we do not know if the pressure varied periodically or not. Sentence 3 does state it, but it is almost three times as long as sentence 4.

By the same logic, another example could be:

1. The contamination on the sample was measured through different testing methods.
2. We used Method A and B to measure contamination.

Between these two sentences, the second sentence is better because it states more information in a fewer number of words. Here "we" is an explicit agent that provides information as to who performed the method to measure contamination. It is common for non-native English speakers to struggle with using active verbs for inanimate objects and to feel the need to introduce themselves by writing in passive form, adding unnecessary verbs and expressions [36]. Nonetheless, replacing implicit agents for explicit ones, can make the sentence clearer.

5.3. Occam's razor

Classical philosophers argued that "entities must not be multiplied beyond necessity"— Occam's Razor (circa 1300 CE). Applying Occam's razor to writing is a way to make the paper better. Do not add unnecessary words; simpler sentences are usually better than complex ones [37]. Needless verbs can either support the main action or not: we performed, we measured, and we used are examples of actions that are not important for the reader but might support the main action [35].

5.4. Ambiguous and types of authorships

It is common to wonder who deserves authorship, especially when you do not have as much experience in collaborative work or research in general. An author must fulfil the following criteria [38]:

1. Contribute substantially to the conception or design of the work.
2. Draft the work or revise it critically for important intellectual content.
3. Approve the final version to be published.
4. Agree to be responsible for any aspects of the work.

Routine tasks require no acknowledgment. You need to go beyond the routine and give helpful observations to be acknowledged, and any person who contributes with intellectual content requires co-authorship status.

The position of the researchers' names in the author list matters, and the meaning of the order can vary depending on the journal. In the traditional research model, the first author corresponds to the person who leads the research and who contributed the most intellectual content. The successive authors are a list of researchers in descending order of contributions. The last author is usually the person with the most prestigious position and who coordinated the research.

5.5. Journals

5.5.1. Choosing the right journal

It is extremely important to choose the right journal to increase your article's impact. Not only the previously discussed topics on writing style, graphs and presentations are important, but also addressing the right audience. Submitting your paper to the right journal increases the likeliness of the paper to get accepted, shortens the submission process, and increases its possible impact.

In Patience's paper [39], his findings demonstrate that 80% to 90% of the scientific articles in the *Web of Science Core Collection* have been cited at least once. This is extremely important since an uncited paper is perceived as having little value [39].

Some interesting remarks to these numbers can be made. First, the citation rate increases with time and follows a sigmoidal distribution, with only around 10% of the papers of the current year cited at least once [27].

Another interesting remark is that researchers in Physics and Chemistry publish the most papers and have less than 10% uncited works, while Arts and Humanities publish the least articles and are the least cited with around 30% left uncited. Conversely, over 97% of the articles on the field of Cell Biology have been cited at least once; all these according to *Web of Science Core Collection*.

5.5.2. Submit to a journal you cite the most

The pertinence of the number of citations as a criterion to judge the impact of research is largely disputed. However, it does reflect research intensity and the number of people working in a certain field. The most cited paper in the *Web of Science* is one that treats density functional theory (DFT); it has been cited over 58 000 times. The citation drops quickly with rank, with the 20th ranked article of the *Web of Science* receiving 3.5 times fewer citations [40]. You should submit your manuscript to the journal you cite the most. This means that the topic of your choosing is closest to this journal's. This will likely increase the chances that the editor will send it out for review.

5.5.3. How to rank research

Granting institutions, universities and departments need to correctly assess the citation impact of the works they are funding. For this reason, it is necessary to measure not only the quantity of citations but the quality as well. There exist bibliometric indicators such as the impact factor [41] and the h-index [42]. For example, Google Scholar classifies journals by the h-5 index which equals the number of articles in the previous years with at least h citations. It is important to keep a balance between publishing in high impact factor journals which will probably bring more funding from research institutions and publishing in general science journals that maintain the vitality and the pureness of research.

6. Conclusion

Knowing how to communicate ideas is a crucial part of the scientific career; our work is incomplete if we fail to transmit our discoveries. In general, we have given the case for writing better in research papers, while also presenting a myriad of tools for this purpose. We challenge some outdated ideas about communication, like avoiding the active voice, that hinder the communication process. Overall, it is on the researcher's end to decide to apply these techniques and to learn how to make their discourse better. Nonetheless, it is a well-directed effort as it will be rewarded by the improvement on the publication's readability and availability to other people, which will in turn raise the paper's rank. Following these guidelines will help improve your chances of publishing, and thus, increase the reach of your work.

Disclosures

The authors of this review article certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

CRediT authorship contribution statement

Daniel Torres-Valladares: Investigation, Writing- Original. **Elvira Ballinas-García:** Writing - Original/Review & Editing. **Jessica Villareal-Reyes:** Writing - Original. **Valeria Morales-Álvarez:** Writing - Original, Investigation. **Carlos Ortiz-del-Ángel:** Writing - Original.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Dr. Flavio Contreras for his guidance and encouragement during this project.

References

[1] G.M. Whitesides, Whitesides' group: writing a paper, Adv. Mater. 16 (15) (2004) 1375–1377.
[2] G.S. Patience, D.C. Boffito, P.A. Patience, How do you write and present research well?, Canad J. Chem. Eng. 93 (10) (2015) 1693–1696.
[3] J. Kirkman, What is good style for engineering writing?, The Institute of Chemical Engineers, Grande-Bretagne, 1971
[4] Science, "STYLE IN SCIENTIFIC WRITING," Science, 02 May 1884. [Online]. Available: https://science.sciencemag.org/content/ns-3/65/534. [Accessed 28 August 2020].
[5] Nature, "How to write a first-class paper," Nature, 28 February 2018. [Online]. Available: https://www.nature.com/articles/d41586-018-02404-4. [Accessed 28 August 2020].
[6] J.R. Tavares, D.C. Boffito, P.A. Patience, G.S. Patience, How do you write and present research well? 15—Prepare to say less than you prepare, Canad. J. Chem. Eng. 94 (10) (2016) 1834–1837.
[7] Neil A. Bradbury, Attention span during lectures: 8 seconds, 10 minutes, or more?, Adv Physiol. Educ. 40 (4) (2016) 509–513.
[8] B. Gastel, R.A. Day, How to write and publish a scientific paper, Greenwood, Santa Barbara, 2016.
[9] Gregory S. Patience, Daria C. Boffito, Paul A. Patience, How do you write and present research well? 6—Tell it in the title, Canad. J. Chem. Eng. 94 (4) (2016) 713–715.
[10] M. Tavallaee, E. Bagheri, W. Lu and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, Ottawa, 2009.
[11] G. Patience, D. C. Boffito and P. A. Patience, "How do you write and present research well? Q4–Do not metastasize with metadiscourse," Canad. J. Chem. Eng., vol. 94, no. 2, pp. 205–208, 2016.
[12] A.M. Davis, The art of requirements triage, Computer 36 (3) (2003) 42–49.
[13] P.A. Patience, D.C. Boffito, G.S. Patience, How do you write and present research well? 14—Favour images over text in graphical abstracts, Canad. J. Chem. Eng. 94 (9) (2016) 1636–1639.
[14] Elsevier, "Graphical abstracts," Elsevier, [Online]. Available: https://www.elsevier.com/authors/journal-authors/graphical-abstract. [Accessed 28 August 2020].
[15] G. S. Patience, C. A. Patience, D. C. Boffito, P. A. Patience and F. Bertrand, "How do you write and present research well? 7—Cite to get cited," Canad. J. Chem. Eng., vol. 94, no. 5, pp. 805-808, 2016.
[16] G. S. Patience, G. Pounis, P. A. Patience and D. C. Boffito, Analysis in Nutrition Research; Principles of Statistical Methodology and Interpretation of the Results, San Diego: Elsevier, 2019.

[17] G.S. Patience, B. Srinivisan, M. Perrier, D.C. Boffito, P.A. Patience, How do you write and present research well? 9—show and state what error bars represent, Canad. J. Chem. Eng. 94 (7) (2016) 1221–1224.
[18] JCGM 100:2008 , ''Evaluation of measurement data — guide to the expression of uncertainty in measurement," 2010. [Online]. Available: https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf. [Accessed 28 August 2020].
[19] G. Cumming, Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis, Routledge, New York, 2013.
[20] R. L. Wassersteina and N. A. Lazara, ''The ASA Statement on p-Values: Context, Process, and Purpose," Am. Statist., vol. 70, no. 2, pp. 129-133, 2016.
[21] Martin Krzywinski, Naomi Altman, Points of significance: error bars, Nat. Methods 10 (10) (2013) 921–922.
[22] D.C. Boffito, P.A. Patience, B. Srinivasan, M. Perrier, G.S. Patience, How do you write and present research well? 10—State the uncertainty, but not too precisely, Canad. J. Chem. Eng. 94 (8) (2016) 1425–1430.
[23] A. Thompson and B. N. Taylor, ''Guide for the Use of the International System of Units (SI)," Special Publication (NIST SP), vol. 811, 2008.
[24] Mesures, Bureau International des Poids et, ''Comptes rendus des séances de la neuvième conférence générale des poids et mèsures," 1948. [Online]. Available: https://www.bipm.org/utils/common/pdf/CGPM/CGPM9.pdf. [Accessed 28 August 2020].
[25] P.A. Patience, D.C. Boffito, G.S. Patience, How do you write and present research well? 11—Respect SI writing conventions, Canad. J. Chem. Eng. 94 (8) (2016) 1431–1434.
[26] D.B. Newell, E. Tiesinga, The international system of units (SI), NIST Spec. Publ. 330 (2019) 1–138.
[27] G. Patience, D.C. Boffito, P.A. Patience, Communicate science papers, presentations, and posters effectively, Elsevier, New York, 2015.
[28] G.S. Patience, D.C. Boffito, P.A. Patience, How do you write and present research well? 12—Design graphs to fit within the journal column width, Canad. J. Chem. Eng. 94 (9) (2016) 1629–1632.
[29] E. Tufte, The Visual Display of Quantitative Information, Cheshire: Graphics Press, 11973.
[30] P. Kamat, G.V. Hartland, G.C. Schatz, Graphical Excellence, J. Chem. Phys. Lett. 5 (12) (2014) 2118–2120.
[31] G.S. Patience, D.C. Boffito, P.A. Patience, How do you write and present research well? 1—admit that you did it, Canad. J. Chem. Eng. 93 (11) (2015) 1889–1890.
[32] R. A. Day, N. Day and N. Sakaduski, Scientific English: A guide for scientists and other professionals, Santa Barbara: Greenwood, 2011.
[33] L. P. Alvin, ''The passive voice in scientific writing. The current norm in science journals," J. Sci. Commun., vol. 13, no. 1, p. A03, 2014.
[34] K. Sainani, S. Elliot and D. Harwell, ''Active vs Passive Voice in Scientific Writing," Retrieved March, vol. 14, p. 2016, 2015.
[35] P.A. Patience, G.S. Patience, How do you write and present research well? 2–Replace 007 with an explicit agent, Canad. J. Chem. Eng. 93 (12) (2015) 2095–2097.
[36] Peter Master, Active verbs with inanimate subjects in scientific prose, Engl. Specif. Purp. 10 (1) (1991) 15–33.
[37] L. P. Alvin, ''The passive voice in scientific writing. The current norm in science journals," J. Sci. Commun., vol. 13, no. 1, p. A03, 2014.
[38] D.C. Boffito, C.A. Patience, P.A. Patience, F. Bertrand, G.S. Patience, How do you write and present research well? 8—Assign authorship according to intellectual involvement, Canad. J. Chem. Eng. 94 (6) (2016) 1127–1134.
[39] G.S. Patience, D.C. Boffito, P.A. Patience, How do you write and present research well? 16—Target an audience and promote, Canad. J. Chem. Eng. 94 (10) (2016) 1838–1840.
[40] G.S. Patience, D.C. Boffito, P.A. Patience, How do you write and present research well? 17—Submit your manuscript to the journal you cite most, Canad. J. Chem. Eng. 94 (11) (2016) 2174–2178.
[41] E. Garfield, The History and Meaning of the Journal Impact Factor, J. Am. Med. Assoc. 295 (1) (2006) 90–93.
[42] J.E. Hirsch, An index to quantify an individual's scientific research output, Proc. Natl. Acad. Sci. 46 (2005) 16569–16572.

Further Reading

[1] WebofScience, ''Web of science core collection," WebofScience, 2020. [Online]. Available: http://apps.webofknowledge.com. [Accessed 28 August 2020].