Principles of Scientific Writing

J.P. Kastelic

Purpose and importance

Scientific progress is usually incremental; a small amount of new information is added with occasionally a large leap of new knowledge. A scientific paper is the primary method of communicating research results to other members of the scientific community. In general, it is the first publication of original research results, described in enough detail that the experiment can be repeated by others, and published in a journal or other readily accessible source. Publication is also a means of demonstrating productivity and maintaining accountability.

Ethics and professionalism

The quality of scientific publications reflects on the reputations of the authors and their institution. Therefore, carefully consider any material prior to submitting it for publication. Authors have a responsibility to be honest in their publications; flagrant falsification of data, withholding conflicting data or purposefully misleading the reader is clearly inappropriate. For a report of original work, there is no substitute for research that was properly planned, conducted and analysed. Not all work is worthy of publication. In general, it is more difficult to publish negative results than positive results. If there is legitimate concern regarding the validity of the work, it is often advisable to withhold publication until the results can be confirmed or refuted. Avoid publishing work in small pieces, merely to increase the number of publications.

A scientific paper contains original work that has not been previously published. If the work (or portions of it) were reported as an abstract or in the proceedings of a meeting, it is valid to submit it, but clearly indicate the previous places that it appeared (often as a single sentence in the acknowledgements or in the introduction). A paper can only be considered by one journal at a time (you cannot legitimately send it simultaneously to two or more journals and wait to determine which has the most favourable review).

Journal selection

There are several factors to consider when choosing a journal. First, identify your audience and the journals they are likely to read. The scope of articles published in a particular journal is usually listed and can also be inferred by examining recent issues. Some journals are more prestigious than others; this can be assessed objectively (e.g., impact factors, citation indices) or more subjectively by asking experienced scientists. In general, submit your paper to the most prestigious (and appropriate) journal that you think will accept it (be realistic). If the chosen journal rejects your paper, you may subsequently choose to send it to a less prestigious journal. The availability of funds to support publication may affect your choice as some journals have page charges and others do not. Some aspect of journal format or quality (e.g. resolution of published

photographs) may affect your choice. Many journals list the dates that a manuscript was first received and when it was accepted, enabling you to assess the typical interval from submission to publication. In some journals, this interval can be reduced by publishing an article as a 'Rapid Communication,' 'Note,' or similar format. Avoid having all or most of your publications in a single journal; publishing in several different journals indicates that your work is accepted by different editors and reviewers.

Instructions to authors

These instructions are published at least once a year, typically in the first or last edition of the year. Carefully reading and following these instructions will greatly improve the submission and publication process. In contrast, ignoring these instructions will almost inevitably delay the process and may make an editor much more critical. Many questions (e.g., regarding format) can usually be answered by looking at recent issues.

Title and title page

The title is a critical method of attracting attention to your work. The title should be specific, informative, and relatively short; it should contain the fewest possible words that adequately describe the contents of the paper. Titles that are too long may be ignored, whereas titles that are too short are often too general. Since many computer searches use the title, careful selection of appropriate terms will increase exposure. Although a title is usually not a complete sentence, correct syntax (word order) is important. Abbreviations, jargon and proprietary names are generally not appropriate. The title can be a complete sentence, but does not have to be. Designating papers as a series is discouraged as they may not be accepted and published in the order designated (they may even be rejected). For some journals, a short title ('running title') may also be required.

Authors

Authorship of scientific papers remains an important method of assessing productivity. In general, only persons that made substantial or important contributions to the work should be included as authors. In some journals, the specific role of each author must be described. Although performing a critical laboratory assay or complex statistical analyses would probably merit authorship, limited technical assistance (or intellectual input) or routine statistical analyses generally would not. The order in which names are listed typically reflects decreasing responsibility for the work, with the person that had principal responsibility being the first (primary) author. In some cases, it will be noted that the first two authors were equally responsible. In many cases, the director of the laboratory is the last author. Indicate the author responsible for correspondence and the author responsible for reprints (or state that reprints will not be available). The addresses of all authors (including telephone and fax numbers and e-mail addresses) should be included. If an author has subsequently moved, it is common to use a superscript and then to give their present address as a footnote.

Abstract or summary

The abstract is typically the first (and sometimes the sole) access to the paper. The abstract contains all the major elements of the manuscript in a short form, including principle objectives, methods, results (including means and probabilities), discussion and conclusions; it is essential that the information in the abstract is consistent with that in the rest of the paper. An abstract usually does not contain references and abbreviations are discouraged (unless they are used repeatedly). Most journals impose a word limit; sometimes a computer database may impose an even lower limit and will truncate the abstract once this limit is reached. Therefore, the abstract must be concise (typically there is little opportunity for discussion). A summary is usually shorter and less descriptive than an abstract and typically contains mainly methods, results and conclusions. If requested by the journal, include a list of key words; be as specific as possible (consult other papers or an index) and put them in the order of decreasing importance. In most cases, the abstract or summary is the last section of the paper to be written.

The IMRAD format

This format refers to Introduction, Materials and Methods, Results and Discussion. This approach has been popular for only a half-century. Based on this model, the following questions are addressed: 1) What question or problem was studied? 2) How was the problem studied? 3) What were the findings? and 4) What do the findings mean? The advantage of this approach is that it gives structure and consistency and makes a paper easier to write and easier to read. Obviously, this format is not appropriate for all papers; review, papers, for example, may have a very different format.

Introduction

The main purpose of the introduction is to tell readers why you have undertaken the study. It is often convenient to divide the introduction into three parts: a description of the problem, a very brief review of existing knowledge, and a statement how the submitted work will challenge, expand or improve what is currently known. Typically only a few critical references should be included (avoid a comprehensive review of the literature); three citations from different groups are usually adequate to establish that a fact is well accepted. The most common error is an introduction that is too long and detailed. The introduction usually concludes with a statement of the objectives of the study or the hypotheses that are being tested. In some journals, it is common to put a very brief summary of the new findings of the paper at the end of the introduction.

An effective introduction includes knowing your audience, writing succinctly, providing a context for the study and justifying your work, and trying to capture the reader's imagination with the opening. To capture a reader's imagination, writers often use one of the following techniques: tell an interesting story; describe a scene vividly; use a strong quotation; give intriguing facts; or express a strong opinion or controversial

pronouncement. Although not all of these approaches are relevant for a scientific paper, they provide ideas to capture the imagination of the reader.

Materials and methods

It is essential that sufficient detail be given to allow the experiment to be repeated. Describe all components used, including animals (breed, age, weight, management, etc.), equipment, supplies and chemicals. Common procedures can be referenced (if the original source is readily available). However, if there were substantial modifications to the original procedures, these should be explained in detail. Procedures should be described in a logical sequence (usually in chronological order). Several techniques can make it much easier to present what was done, including defining a fixed time point (e.g. day of ovulation or day of treatment as Day 0), or making a flow chart, diagram or table. A complete description of all statistical analyses is important (but often absent). Define experimental units, independent variables, response variables and risk factors (as appropriate). The methods of assigning experimental units into treatment or study groups must be described. There should be complete agreement between the materials and methods section and the results section; if something was done, the results should be listed and if the results are listed, it should have been described. Carefully compare these two sections to confirm agreement.

Results

The results section has two key features: an overall description of the major findings and clear, concise presentation of the data. This section should be brief, accurate and generally should follow the same sequence as the materials and methods. The goal is to present the information in a clear, comprehensive and easy to interpret format. Keep the results section as brief and uncluttered as possible and organize the presentation of the information.

Differentiate between data and results; data are factual findings (e.g. measurements), whereas results include the meaning of the data. Data can rarely be listed without stating the results. However, avoid excessive interpretation of the results; that belongs in the discussion.

Many readers prefer to look at tables and figures in lieu of written descriptions of results; therefore tables and figures should have strong visual appeal, be easy to understand and effectively stand alone from the rest of the text. Try different methods and formats for presenting the data; the preferred method is often not the first one tried nor the most complicated. Include summary statistics, often means and a measure of variability (e.g. standard deviation or standard error) or a confidence interval. If data are presented in tables or figures, highlight some key results in the text while avoiding either inadequate or excessive reference to the tables and figures. Do not use table headings or figure legends as topic sentences; state the results directly and cite tables and figures in parentheses. Wherever possible, be consistent from table to table and figure to figure, to make it easier for the reader.

Tables should be used only when you cannot easily present data in the text. Data should be organized so that common elements are in a column (vertical). The number of decimal places should be consistent and reasonable. Several forms of graphs can also be used. Where appropriate, the vertical axis represents the dependent variable and the horizontal axis represents the independent variable. In some cases, two or more panels are combined (usually vertically). To facilitate comparisons between two or more graphs, use consistent axes (also try to maximize the utilization of the available area). Ensure that all headings, lines, symbols, etc. are clear, include probability values, and indicate significant differences (e.g. with superscripts or asterisks).

Photographs should be truly representative (avoid the rare 'perfect' image) and carefully cropped. Including a ruler or scale accurately indicates size (regardless of photographic reductions during publication). Tables and figures generally have a title and often have legends or footnotes (enabling them to be interpreted independent of the text).

When appropriate, indicate the top of the figure, and always ensure that the sequence number (e.g., Figure 1) and preferably reference to the author and the manuscript are included. If necessary, prepare labels (preferably typewritten) and put them where they will not interfere with the image (e.g. on the back of a photograph). If figures are prepared with a computer, print each figure on a separate page, and print the captions separately. For type-set journals, print captions consecutively and allow the journal to determine their location in the paper. For 'camera-ready' journals, indicate the location of the figure by printing the caption and leaving a blank space for the figure. In this situation, it is also advisable to submit at least one copy of the manuscript with the figures included (typically a photocopy of the original with the figures 'in situ').

Discussion

The discussion is the 'heart' of the paper. The results should not simply be restated; they should be discussed in the context of what is already known. A discussion should include the following: summarize major findings; discuss limitations and problems of current study; discuss scientific and clinical implications; suggest further work; and provide a concise conclusion. As a guideline, the discussion should not exceed one-third of the length of the paper (excluding references). The order of presentation is typically in decreasing priority, starting with hypotheses or objectives; answer the questions that were posed in the introduction and give the primary evidence supporting those answers. This should be followed by a critical assessment, including shortcomings in design, limitations of methods, potential flaws in the analysis, and validity of the assumptions. Once you have stated the results of the present study, make comparisons with other studies; start with the most important references, specifically those that generally agree with our findings. Thereafter, refer to contradictory sources, discuss similarities and differences, indicate potential shortcomings (do not be excessively critical) and perhaps suggest further work. Finally, discuss the conclusions and implications for your study (perhaps discuss the potential for further work). Avoid

discussing every paper ever published on the topic; this will often make the paper excessively long and boring.

Make sure that the objectives are consistent with what was done and furthermore, that the conclusions are supported by the data. Ensure that your arguments progress logically, with each paragraph leading the reader step by step. To verify that this has been done, carefully read the manuscript and highlight key sentences that advance your argument. The location of these sentences is critical; most should be at the beginning of paragraphs and some could be at the end, but very few should be in the middle. If there are large blocks of text that are not highlighted, they may be redundant and can be eliminated.

Regarding statistical differences, in general, $P<0.05$ is accepted as 'significant' and $P<0.1$ is accepted as a 'tendency.' Although it is common to cite a difference as significant and to include a P value, it is preferable to indicate that there was a difference and then include the probability in parentheses. Sometimes it is also appropriate to indicate the nature of the difference (e.g. something is bigger than something else) or to list the means. If a difference was significant, then clearly state that it was significant. In contrast, if results were not statistically significant, do not pretend otherwise. If certain results simply cannot be explained, a simple sentence (e.g. ' For unknown reasons.....') should be included. If appropriate, include brief, reasonable speculation. Additional information can also be presented, including a brief description of peculiarities of individual animals in the experiment. Direct comparisons are valid when animals are appropriately assigned to groups (as part of experimental design). However, unplanned comparisons, including comparisons between experiments (including experiments described in the paper or one experiment in the paper and one or more in the literature) should be made 'by inference.' In this case, words such as 'seemed' or 'appeared' are appropriate.

It is important that the discussion have a strong conclusion; many readers will read the abstract and if still interested, will read the final paragraph of the discussion. A conclusion that more work needs to be done is not a strong conclusion; this is wasting a critical final opportunity to summarize the key findings of the paper. In the concluding paragraph, avoid new arguments and references to the literature and make as strong a statement as warranted by the findings.

Acknowledgments

This section should be very specific. In some cases, an institutional reference number is reported first. Specific reference to funding sources (e.g. grants or scholarships, including reference number if applicable) should appear next, followed by reference to donated materials, etc. Finally, list people who were important to the work, but that did not merit being included as an author. Discretion is needed as an excessively long list will clearly diminish the intended honour. Furthermore, avoid the term 'wish' (e.g. We wish to thank…) as it is redundant. It is desirable (essential in some journals) to have written permission from persons before including their name in this section. If

some or all of the data were previously reported as an abstract or proceedings article, this can be noted in the Acknowledgements.

Citations and bibliography

In the text, reference to the cited work should be in a logical location in the sentence (which may not be the end of the sentence). Insure that all information listed in the bibliography is accurate and that the format is correct (errors are common). Minimize the use of errors. If you were unable to locate the primary source, indicate the primary source as well as the secondary source (the source that quoted or cited it). Minimize the use of secondary citations and abstracts. If the article was originally published in another language, cite the original title and include an English translation of the title (in square brackets). When the paper is complete, verify that all references cited in the text are listed in the bibliography and that all references in the bibliography are cited in the text.

Tables and figures

The primary purpose of a table or figure is to present data that are too complex to describe in the text of the results. However, if data are described in a table or figure, do not describe them again, in great detail, in the text. If numbers have an interesting trend, put them into a figure; otherwise, put them into a table. A column chart is appropriate for discrete data (no carryover effect); however, for continuous data, a line chart is recommended.

All tables have a similar structure: title, column headings, body, and footnotes. Keep the title brief and ensure it is consistent with the content of the table. The title of table or figure can be either descriptive (describes content in general terms, without listing every end point) or declarative (describes a clear trend or relationship). The table or figure should 'stand alone'; i.e. it should be understandable in the absence of the text. Concurrently, avoid describing the experimental design or treatment groups in great detail. Column headings should be brief, minimize repetition, use subheadings if necessary, and include units of measure (in lieu of repeating them excessively). For tables, put like elements in a column; typically the control (or the independent variable) will be listed first.

A table should convey one or two main ideas; if there are multiple ideas, it may be better to create additional tables. Avoid an excessively wide table; it may be necessary to put SD, SEM, range or confidence intervals below the mean; this is valid if the primary comparisons are horizontal (among columns). However, if the primary comparisons are vertical (within a row), measures of variation should be placed beside the mean. Minimize the number of lines within a table; typically there should three horizontal lines (above the column headings, below the column headings and at the bottom of the table) and no vertical lines.

Figures should be large enough that they are still legible after a 50% reduction in size; use the reducing feature on a photocopy machine to determine their suitability.

Since the primary space limitation in a journal is typically the width, avoid excessively wide figures; if you have two figures, arrange them vertically and not horizontally. Minimize blank spaces, avoid excessive numbers of lines, and where possible, directly label the line.

A figure legend typically has four parts: brief title, brief description of the experiment, definitions (e.g. symbols or abbreviations), and statistical information. Keep the title brief, use the same key abbreviations as used in rest of paper, avoid excessive details regarding design, and give key statistical outcomes (without describing the statistical tests used). If identical symbols or abbreviations are used more than once, define them in detail in the first figure legend and subsequently refer the reader to the that legend.

The best font for labels on figures is a sans serif typeface, e.g. Helvetica. Either use different lines or different symbols to depict different groups, but do not do both. Keep all lines of similar thickness, use standard symbols (e.g. open circles, triangles, squares, and then closed circles triangles, squares) and avoid excessively large symbols. The lines should be thicker than the data axes and in general, avoid grid lines on the figure. For a column chart, the columns should be wider than the spaces between columns. Columns that are shaded grey are generally more aesthetic than stripes or dots.

Tense

In general, the present tense is used for reference to the literature and figure captions whereas the past tense is used for everything else.

Numbers

Depending on the journal, numbers <10 are often written as a word, whereas numbers $\geq$10 are often written as an Arabic number. Do not start a sentence with a number; revise the sentence to have at least one word at the start. Avoid the use of two numbers consecutively (e.g. 10 500 mg tablets. Minimize the number of zeros, use scientific notation, but avoid the use of the term 'billion' (means 10^9 in the USA and 10^{12} in most European countries). It is best to avoid the use of a comma for numbers with more than 3 digits before the decimal place; some journals will simply leave a blank space (e.g. 20 310).

Percentages are commonly used, but often used incorrectly. The following is a suggested use of percentages, based on the number of units used to calculate the percentage: <25 (use proportions and avoid percentages); 25 to 100, give a percentage as a whole number (no decimal); 100 to 100 000, use one decimal place; >100 000, use 2 decimal places. When using a percentage, always give the original data, never use a slash mark (i.e. do not say 11/25), and always give the original number first, e.g. 209 (7.2%) of the 2901 oocytes reached the blastocyst stage.

Logistics of submission

Follow the instructions to authors to determine the format, number of copies, etc. If a paper was initially submitted to another journal and rejected, make sure that you revise the format in accordance with the subsequent journal; leaving it in a distinct but incorrect format is a clear sign that it was previously rejected.

Photographs and figures should be appropriately labelled and should be placed within a separate, labelled envelope (that is placed within the envelope or package sent to the journal). A cover letter to the editor, typically signed by the corresponding author, must be included. The cover letter should refer to the title and the authors, indicate that the work is original, and politely request that the manuscript be considered for publication.

Retain copies of all text, figures, photographs, etc. Insure that the envelope or package is secure, properly addressed and has sufficient postage. Most journals will send a letter or post card confirming receipt of the manuscript and indicating the reference number that it has been assigned. If confirmation has not been received after a reasonable interval (e.g. 4 to 6 weeks) contact the editorial office and politely inquire whether the paper was received.

On-line submission is becoming more common. Follow the instructions and retain copies of all documents. Using common file formats and minimizing the changes that need to be made before the paper is sent for review (e.g. typically include page and line numbers) will be appreciated and will facilitate your submission.

Reviews, reviewers, editors

You can typically expect a response within 4 to 8 weeks of submission; in the absence of a response within 8 weeks, a polite inquiry to the editor is appropriate. There are several common general assessments of a manuscript; it can be accepted without revision (uncommon), accepted with revisions (another review may be requested) or rejected. Regardless of the editorial decision, a rapid response is desirable (in many cases a specific time limit is imposed by the journal). Typically the reviewers and editors will request revisions or identify issues that need to be considered; many of these will be helpful and valid, although some may seem equivocal. Although the author generally has the right to refuse these changes, it is best to make most (or all) of the changes, and only refuse to change issues that you feel very strongly about. When returning the manuscript to the editor, include a cover letter. Graciously acknowledge the comments and suggestions that were made and specifically indicate how each was dealt with (avoid a general statement indicating that the changes were made as requested). For most comments, a single word (e.g., changed) or a short phrase or sentence will be adequate. Issues that are not changed will generally require a detailed rebuttal.

If the paper is rejected by the editor, you generally cannot send it back to this journal. Consider the reasons given for rejection and also the comments and criticisms. It may be possible to revise the paper, modify the format (as needed) and send it to another journal. In some cases, the paper is not rejected, but it will clearly be extremely

difficult or impossible to address the comments, criticisms and revisions. In this case, the most reasonable approach may be to politely withdraw the paper and perhaps consider submitting it to another journal.

Conclusion

Scientific publications remain a critical part of the scientific process. Good science, well designed, conducted, and reported, is a positive contribution and a pleasure to read. Reading well-written papers and much practice are useful methods to improve writing skills.

References Consulted

Day, Robert A. How to Write and Publish a Scientific Paper. Fifth Edition, 1998, Oryx Press, 275 pp.

Gustavii, Björn. How to Write and Illustrate a Scientific Paper. 2003, Cambridge University Press, 141 pp.

Hall, George M. How to Write a Paper. Third Edition, 2003, BMJ Publishing Group, 176 pp.

John P. Kastelic, DVM, PhD, Diplomate ACT
Co-Editor in Chief, Theriogenology

Senior Research Scientist
Agriculture and Agri-Food Canada
Lethbridge Research Centre
5403-1 Ave South, P.O. Box 3000
Lethbridge, Alberta, Canada T1J 4B1
403-317-2236 (office direct)
403-382-3156 (fax)
kastelicj@agr.gc.ca
therio@shaw.ca