Day 1: Writing an English journal paper

Geoff Hart
Beijing Forestry University
October 2010

Introduction

A bit about me and the three lectures:

- My training: a scientist (plant physiology and ecology) at the University of Toronto.
- Career: For 25 years, I have helped scientists from every continent publish papers.
- Peer reviews: I have helped authors of more than 5000 journal articles respond to peer reviews.

Notes: My goal is not to brag or seem prideful about my accomplishments; I want to establish that I have been doing this for a long time and that your students can trust me.

Introduction

Overview of today's lecture:

- Goal: I will not teach you the "best" way; I will teach you the *Western* way.
- Writing for English journals is like learning English: the rules are different from the Chinese rules.
- I will teach you what I have learned from helping authors respond to more than 5000 journal reviewers.
- I will discuss experimental design, the sections of a manuscript, and the requirements of literature reviews.

Notes: Western culture and Chinese culture have different rules. When you are working in Western culture, you must follow Western rules. English journals follow a specific set of Western rules, and not all of those rules have been written down. Some are rules that you only learn from experience. I will discuss some aspects of experimental design that often cause problems because many graduate schools do not teach this course as a separate subject, and some papers are rejected because the design is not adequate. I will discuss the sections of a journal manuscript to help you focus on what needs to be written in those sections. I will explain the goal of literature reviews in each section to help you decide what aspects you should focus on in each section.

Experimental design

Because experimental design is often not taught as a separate subject, many mistakes occur:

- Start by defining your research hypothesis: what problem are you investigating?
- What evidence do you need to support that hypothesis? (List the primary parameters you will measure to provide that evidence.)
- What additional parameters must you measure to confirm the results for the primary parameters? (This is called "triangulation".)
- What evidence will exclude alternative explanations? (If there are two or more possible explanations for a result, what experiments will tell you which explanation is correct?)
- What measurements (data) are necessary for each parameter? (For example: a reliable mean value versus a calculated index value.)
- The research literature will teach you the common measurements, but sometimes you must invent your own measurements.

Notes: Students are often expected to learn how to do research from reading journal papers or watching their professors. Even a skillfully written paper will not be published if the experimental design is inadequate. Here are some ways you can avoid the most common mistakes. To design a successful experiment, you must start by clearly defining your research problem. This will lead you to one or more hypotheses. Each hypothesis can be tested by measuring certain parameters, so you must decide what those parameters are (that is, you must list the evidence required to support your hypothesis). For each primary parameter, it may be necessary to identify additional parameters that will provide more evidence to support the results for that parameter. This is often called "triangulation", which is named after the positioning technique where you measure the angles to a position from two different points, and use trigonometry to calculate the coordinates of that position. In research, this means that looking at a phenomenon from two directions is more likely to help you see all sides of the phenomenon, and therefore to ensure that you are seeing it clearly. Because many problems have two or more different explanations, you should also collect evidence that will help you understand whether the other explanations are possibly correct, or are probably wrong. Once you know the parameters you must study, you can then decide the optimal way to measure them. Previous research will tell you what kinds of measurements other scientists accept, but sometimes you must develop your own measurements (e.g., if there are problems with existing measurements).

Experimental design

The statistical design must also be good:

- Design an approach that will let you measure all the parameters you identified.
- Compare your approach with the approach used by previous researchers: Have you missed any important steps?
- If you are not confident, follow the Methods in a published article. However, modify their experimental design and methods to account for your specific situation.
- Learn the sample sizes and the resulting standard deviations or standard errors reported by previous researchers.
- Use this information to calculate the minimum sample size that will provide statistically significant results.
  Example: http://www.itl.nist.gov/div898/handbook/prc/section2/prc222.htm
  Example: http://www.itl.nist.gov/div898/handbook/prc/section2/prc242.htm

Notes: Once you know what you must measure, you can define a series of steps that will let you obtain all the necessary data. When you are starting out as a researcher, you may not be certain that you are solving all possible problems with your data-collection strategy, so you should examine published research: because it has been published, this means that at least two journal reviewers and the journal's editor believe they have solved all the problems with that data collection strategy. Following that proven strategy increases the likelihood that you will not forget any steps in your own data. Next, learn what sample sizes previous researchers have used and the kinds of error terms (standard error or standard deviation) that those sample sizes produced. You can use those sample sizes to tell you whether you must increase your own sample size or whether their sample size is adequate. If you can obtain this information from several studies, you can predict the size of the error terms you will encounter in your own research; if you're not certain, use the largest sample size that was reported. There are also tools, such as the descriptions at NIST, that will let you calculate (predict!) the sample size you will need to obtain a statistically significant result. There is no guarantee that you will succeed,

but following this approach greatly reduces the likelihood that you will fail, and will collect only non-significant results. Many experiments fail because the researcher guessed at the sample size instead of trying to predict it based on existing data.

Sections of a manuscript

In the Abstract or Summary:

- Write this section last, not first.
- Summarize your study (research problem) and its results in around 200 to 250 words.
- Summarize your methods. Only describe the details if your paper is about the methods, not the results of using those methods.
- Present the main results.
- Briefly discuss the implications (consequences) of your results.
- No literature citations are allowed!

Notes: Many authors write their Abstract first because they think they already know what they will say. However, you will often discover that you did not fully understand your data until you presented it and began looking for relationships between your different results. This leads to changes in your conclusions, and that causes two problems if you wrote your Abstract first: you may end up with different results or conclusions in the Abstract and in the Results and Discussion section, and you have wasted time because now you have to rewrite the Abstract. An Abstract must be brief; most journals limit it to 250 words, and many have even shorter limits. As a result, you can only focus on the most important points. Unless your paper describes the development of a new methodology, all you can do is provide the most basic details of your method. Focus on the problem you tried to solve, and the results that solve that problem: only the most important results, not all results. If you still have enough space, discuss what your results mean: Did you support or contradict your hypothesis? What additional research will be necessary? Do your results have important consequences for other researchers, farmers and foresters, the government, or the general public? Note that you cannot provide literature citations in an Abstract because it is often published separately from the rest of your paper, including the References section (e.g., in "Current Contents" or "Biology Abstracts"). Thus, readers have no way to know what paper you are citing.

Sections of a manuscript

In the Introduction:

- Define the problem you studied by reviewing the work of previous researchers who studied this topic.
- What have they learned about this problem?
- What information is not yet known?
- Based on the gaps between what is known and what is unknown, what new information will you provide?

Notes: Readers want to know what new information you will be contributing to the scientific literature. Thus, you must remind them of what is already known and what information is still unknown. Your goal will be to provide information that is unknown, and to do so, you must convince the reader that you understand what unknown information relates to the research problem. Your goal will be to focus on specific knowledge gaps (things that are not known) to provide a deeper understanding of the problem and to guide future researchers to fill in any remaining gaps.

Sections of a manuscript

Goals of your literature review in the Introduction:

- Summarize the current state of knowledge.
- What research problems are you trying to solve? What unknown things did you try to learn?
- What approaches have you adopted from this previous work? Which ones have you avoided?

Notes: When you summarize the current state of knowledge, you convince the reader that you are familiar with existing explanations for a phenomenon. But not all aspects of the phenomenon are fully understood, and readers want to know which of these aspects you will focus on. The many years of previous research on a problem reveal certain effective ways to study a problem, and also certain approaches that don't work particularly well or that not all researchers accept as valid. For generally accepted approaches, you will not have to work hard to persuade the reader that your results are valid. For approaches that are not universally accepted, you may need to take certain steps (e.g., triangulation) to support the results produced by those methods.

Sections of a manuscript

In the Methods section:

- Describe how you studied the problem: provide enough detail that readers can repeat your research *exactly*.
- The goals of replication: to confirm that your results are valid, to provide similar results for other situations.
- Demonstrate that you are using the most advanced or powerful methods that are currently available.
- Describe the overall approach in general terms; cite a published paper for the details.
- For new procedures that you developed in this research, provide complete details.
- Write your Results section first, then use this as a checklist for the Methods section: every result = a procedure that you must describe in the Methods.

Notes: The goal of the Methods section is to let any reader completely repeat your experiment, and confirm that your results are real by obtaining similar results. Alternatively, the goal may be to obtain similar results for another species or country or context so that readers can determine whether your results are generally applicable, or only apply to a specific situation. You must also demonstrate that you are using the most appropriate methods to perform your research: these methods must have been proven (by previous research) to produce reliable results. You should describe your overall approach in sufficiently general terms that readers will understand the basic approach that you used to solve the research problem. Where the procedure is well known, you can often cite a published paper that provides details for readers who are not already familiar with those details. If you have developed a new methodology, you must provide enough details that readers will not have to guess at what you did. One very useful approach to writing is to write your Results section before you write the Methods: for every result that you present in this section, you must add a section in the Methods that describes the procedure that you followed to obtain that result.

Sections of a manuscript

In the Results section, summarize the key data you collected:

- If you create a separate Discussion section, do not interpret the results.

- Do not repeat all your data: summarize (e.g., mean ± SD, range) and compare.
- Finish analyzing your data *before* you begin writing: emphasize results that test your research hypotheses
- Only present a meaningful level of precision: high precision may be mathematically significant but not meaningful to the reader.

Notes: If you are writing a separate Results section, most journals will not let you discuss the meaning of the data you collected. Instead, you only present the key data. You should not repeat all the data that you presented; instead, summarize it to present values that let you compare, such as the means for two experimental treatments or two different study sites. Readers who are interested in the details will read your tables and figures to learn those details. It's important to finish analyzing your data before you write the Results section. This seems obvious, but many researchers write while they are analyzing the raw data, and that makes it hard to concentrate on the overall hypothesis and whether your results support it. The key data that you must report is based on your research hypotheses: what evidence supports each hypothesis, and what evidence contradicts each hypothesis? Note that when you present data, it's tempting to produce a large number of decimal places for all values because this makes the data appear to be highly precise. There are clear rules for determining the number of mathematically significant digits you can present, and these can guide the level of precision you can report in a table. But in the Results, it is often is more important to focus on the results that are meaningful. Example: Most journals only accept 1 decimal place of precision for percentages. Example: When the resolution of a satellite image is 300 m, it is unlikely that results precise to more than 1 km^2 are meaningful.

Sections of a manuscript

In the Discussion and Conclusions:

- Repeat only your key findings: ones that support or contradict your hypotheses.
- Only describe the methods if they explain why your results differed from previous results or if they had limitations.
- Do your results solve the problem you studied? If not, what future research is necessary?
- Do your results confirm or contradict previous research? Explain why.
- What new knowledge are you adding to the literature?
- What do your results mean for the reader? Summarize the implications (consequences) of your results (e.g., for government policy, for other researchers).

Notes: As is the case in the Results section, the goal is to focus on the most important things you learned in your research. These are the results that support or contradict your hypothesis. It may be appropriate to discuss the methods you used (briefly!) if they can explain why your results differ from previous results; for example, the methods may not be appropriate (or may be more appropriate) for your species or region, and factors related to your specific situation may cause those methods to produce different results. It's important to tell the reader whether your results confirm previous research, and if they don't, you must provide a plausible explanation for the cause of the difference to convince the reader that the differences have a cause; don't let readers ask themselves whether you are just a careless and imprecise researcher! Based on your findings, what does this mean for the reader? Emphasize what new information you are providing: you may be filling a gap in the existing knowledge, reporting an entirely new result, or simply confirming previous results, possibly in a new context (e.g., in China). State clearly whether you solved the research problem that you were studying. If you only partially solved the problem, tell the reader what you plan to do in the future to provide any

missing knowledge. If there are implications of your results for the world, tell the reader what those implications are.

Sections of a manuscript

Goal of the literature review in the Results and Discussion:

- What did previous studies find?
- Did you confirm or contradict previous results?
- Explain any contradictions. Sometimes they cannot be easily explained, and further research will be needed.

Notes: When you review the literature in these sections, your goal is to explain how your results fit into the previously existing body of knowledge (i.e., the published research literature). When you discuss the findings of previous studies, you are telling the reader whether you are confirming those results for your specific, unique situation, or whether there are any contradictions. Where there are contradictions, you will form new hypotheses to explain the possible causes of those contradictions. In particular, you will cite papers that provide evidence that your hypotheses (explanations) are correct. Where no evidence exists, this is a sign that you may need to do additional research to provide evidence that will support or contradict your explanations.

Literature review

Other important points about your literature review:

- Always read the original paper, and read the whole paper. Many errors enter the literature when someone cites a paper they have not read, or reads only the Abstract.
- Cite English-language papers to prove that you understand the world research situation.
- If you expect a specific researcher to review your paper, cite their recent research.
- To compare your results with previous Chinese studies, it is always acceptable to cite Chinese papers.
- You must also cite papers that present results from similar studies outside China.
- Remember that English readers can't read Chinese. Wherever possible, cite English papers in addition to Chinese papers.

Notes: I have briefly explained the goals of a literature review for each section of a manuscript, but there are many details that apply to each of these sections of the paper. First, you must always try to read the original paper. Many serious errors enter the research literature because someone made a mistake in describing the results of another researcher's work; if you read the original paper, you will not repeat this mistake. Steven Jay Gould, a famous evolution researcher, has written many essays about these problems. One of my former colleagues had a paper rejected by a journal because the reviewer cited contradictory evidence from an Abstract that did not match the full text of the manuscript (this is why I told you earlier that you should always write the Abstract last, not first). Because you are writing in China, you must always cite English papers from around the world to prove that you understand the world research situation, not just the situation in China. This same rule applies for researchers from outside China, of course. It is particularly important to cite recent papers by researchers who are likely to be selected by a journal to review your paper. This is because human nature (psychology) makes people more willing to accept your paper if you show that you understand and agree with their own beliefs. Citing Chinese papers is a good strategy if you must compare your results with previous research in China. But China is only one part of the world, and the conditions in China (e.g., climate, vegetation species) are very different from conditions in

other parts of the world. Readers want to know whether your results are only valid in China, or are likely to be valid in other parts of the world. My last point is that citing Chinese papers creates a problem when you are writing for an English journal: very few readers of an English journal speak Chinese, so they will not be able to read any Chinese papers that you cite. Providing an English paper that describes similar procedures or similar results lets the reader fully understand your research.