Writing a Research Paper, and Submitting it for Publication

Peter Hall University of Melbourne

Introduction

Undertaking research can be very challenging, to the point of exasperation. Just because you enjoy a field doesn't mean that you'll enjoy doing research there.

For example, reading about history can be a lot of fun, but doing rigorous research in history may be a different matter entirely. You have to read, and absorb, lots of information in highly condensed, difficult to understand research papers, and you need to back up all your assertions with cold, hard facts.

If you can't do these things then your enthusiasm for research may evaporate quickly.

How can we possibly read those complex, highly condensed papers? Is there a trick to it?

And how do we assess the originality of our new ideas, relative to those that have already been published? This requires a literature search, but how do we do that?

Introduction (continued)

How can we create the right environment for doing the hard thinking that is essential to making progress in research?

How do we write a research paper? Do we simply start with an Abstract and an Introduction, and keep going until we've finished the list of references?

What are the dos and don'ts when submitting a paper for publication?

How do we deal with disappointment, for example when we find that our method doesn't work as well as we thought, or that a referee thinks our work is not all that interesting?

We'll try to address all these issues.

Originality

Originality, novelty, innovation, creativity, nonconformity—call it what you will, originality is one of the main ingredients of excellent research. It is more important than technical correctness.

Try not just to find a research problem in your field that no-one else has tackled. Try to look in a genuinely new direction, one that nobody has considered before.

Of course, this can be very difficult, and you may not manage it more than a few times in your career. But it should be your aim, the main goal of your scientific work.

In particular, it should be your main goal when you look for a problem on which to work, even it you don't always manage to achieve it.

Locating the Literature

Generally your research problem, or topic, will be associated closely with work in several other papers that you already know about. It is important that you determine who else has worked in the area, and what they have done. There are several ways of doing this.

First, use a search engine such as Google to locate items that are referred to online. In this search you would normally use carefully chosen keywords from the problem or topic you plan to address.

Using the titles of the papers you already know is not necessarily a good approach—the titles may be too narrow, and restrict your search unnecessarily.

Locating the Literature (continued)

Secondly, use a citation index, such as Google Scholar, the Web of Science, MathSciNet or Scopus, to find articles that have cited the papers that you feel are related to your topic.

In statistics, colleagues and I have found Scopus to be less complete, even for very recent work, than the Web of Science. For example, up to the end of last year Scopus had completely missed all papers published in the Annals of Statistics in 2012, and it tends also to overlook volumes that are indicated by a letter as well as a number (e.g. volume 1A).

Be aware that a good literature search is not a linear, or even a monotone, process! In particular, you may find work of others that is uncomfortably close to the work you are planning to do, thus leading you to take a new direction. This means that you will have to reconsider the next step. Be adaptive, and move from one research problem to another as you learn more!

Treat the discovery of closely related work as a compliment—your choices of topic and problem are clearly good, because other people also are interested in them.

Locating the Literature (continued, 2)

Thus, the contributions of others may take you in new directions, along paths that you had not contemplated before you knew the literature better.

Allow your literature review to lead you to quite new topics, if that turns out to be the most appropriate way to go.

I remember discussing some of these issues with a colleague, G.S. Watson, many years ago. We agreed that there is an optimal time at which one should search the literature for related work.

Searching too early means that you might be too influenced by work of other authors, and that the originality of your work might therefore be less. Searching too late might mean that you discover relatively late that your proposed approach is too close to one that has been considered already, and that you have wasted time developing it.

Reading the Literature

Reading scientific literature, particularly literature that is very technical, can be particularly difficult. There is a great deal to absorb, and in most journals, pressures on space mean that the ideas in a paper are described in the minimum number of words, so that there is little opportunity for the authors to convey intuition or other helpful information.

To overcome these problems you should read a scientific paper in layers, rather than in depth in a single sitting. The process is like picking at your meal—first you eat the things you like, i.e. you skim through the paper, noting mainly the things that you can understand fairly quickly. Then, unless there is a pressing need to understand the paper further, you put it to one side and get on with the next article.

You may be surprised to find that your brain actually works quietly on the more difficult material, while you are getting on with other tasks. Later, when you return to the paper, it will likely make more sense to you than it did at first, even though you haven't consciously tried to work on it.

Reading the Literature (continued)

Even if you do need to understand the details of the paper, rather than just get an overview, it can be much more profitable (and also less painful) to read the relevant parts of the paper in several sittings, with each sitting probing a little more deeply than the previous one.

Learning to read complex, highly compressed scientific literature may be one of the most challenging things you have to master in research. It can be particularly difficult.

Discovering the Environments that are Most Conducive for Difficult Tasks

Try to discover the reading environments that are best suited to you. For me, they are environments with as few distractions as possible. Aeroplane flights are one; I usually don't like the movies, so I may as well read a difficult research paper. The only other attraction is reading a book, or sleeping! If you can read while you are commuting to work, that might be the right place to read challenging papers.

Another place is the gym—if, like me, you are a regular gym user, consider taking some of your work to the gym, to read between sets or on the exercise bike. I often take along a paper, or a document that I have to read. (For example, I proof-read these transparencies in the gym.)

Discovering Good Places to Think

Those same distraction-free environments, where you can't be interrupted by email, or surf the web, or answer the phone (leave your mobile phone behind as much as possible!), and where no-one will knock on your door, are also good places just to think.

They provide opportunities for you to think through those aspects of your research that are vexing you, and to make progress. Thinking doesn't have to be accompanied by writing; indeed, writing can be a real distraction.

In short, you need to ensure that you have enough quality time for just plain thinking. That time is critical to good research. This is particularly important if your social life, or family life, tend to be overwhelming.

Writing a Paper

Research Papers often tend to be written backwards, rather than forwards. In particular, the Introduction is often written last. (Preparing this talk was similar—I wrote the first two transparencies, headed "Introduction," only after all the others were done.)

There are sometimes exceptions to this "rule," particularly if the direction of your research is changing as you learn more.

For example, sometimes it can be helpful to write the introduction to a paper before other aspects of your work are finished, because doing so helps you to organise your current thoughts and lines of reasoning.

Writing a Paper (continued)

However, be prepared to scrap this introduction completely at some future date, when your research takes another turn and you discover that most (or all!) of the things you thought were worth doing turn out to be uninteresting.

There is nothing wrong with that—it is the nature of research. You haven't made a mistake; it's a normal way of making progress.

More generally, to write a very good paper it is normal to go through many drafts, and to write and rewrite the paper many times.

Producing a Theorem

If you are doing theoretical work then the steps of developing the technical assumptions you need, drafting the theorem, and preparing its proof, are also usually done in an order quite different from that in which they will be presented.

In particular, you would have an idea of what you wanted to prove, and why it was true, so you might start by developing a proof which captured, in some sense, your intuitive reasoning. Along the way you would start drafting regularity conditions, and then, when everything started to fall into place, you might draft the statement of the theorem.

You will almost never start by stating the theorem or writing down assumptions.

Getting it all Wrong

Ben Selinger, a retired ANU chemist with whom I have worked, once confided to me that only when everything seems hopeless, and he has discovered that all his preconceptions and arguments are wrong, does he know that he is on the verge of discovering how everything works.

Those words also capture very well my own experience, on many occasions, of the process of research. The best research is far from a linear process.

Many other scientists, and even non-scientists, share this view. For example, the eminent applied mathematician and theoretical physicist, Freeman Dyson, has written that:

You can't possibly get a good technology going without an enormous number of failures. It's a universal rule. If you look at bicycles, there were thousands of weird models built and tried before they found the one that really worked. You could never design a bicycle theoretically. Even now, after we've been building them for 100 years, it's very difficult to understand just why a bicycle works—it's even difficult to formulate it as a mathematical problem. But just by trial and error, we found out how to do it, and the error was essential.

Getting it all Wrong (continued)

Thomas J. Watson, who as IBM's Chief Executive built IBM into one of the world's most influential companies, once said, "If you want to succeed, double your failure rate."

Robert Lefkowitz, a Nobel Prize-winning chemist, has argued that "Science is 99 percent failure, and that's an optimistic view."

Winston Churchill characterised success as "going from failure to failure without losing enthusiasm."

Failure is an intrinsic part of the scientific process. It teaches us new things about our field, and that is always good. When everything we are doing is going wrong, and we have to understand why and fix the problems, we are learning more than when everything is working out as we expected.

You should remember these words when you discover that your ideas don't work as well as you had anticipated, or that someone else has already had the idea that you thought was yours, or when a journal rejects your work.

The Antonym of Originality: Plagiarism

This page and subsequent pages are based on ethical principles for the *Annals of Statistics*, which I helped to draft. See:

`http://www.imstat.org/aos/principles.html`.

All submitted material must be your, and your co-authors', own work, original to the submission, unless specifically stated as having another origin, which must be identified clearly.

Here "original to the submission" relates not just to published work, but to unpublished work that you or your co-authors have submitted, or are submitting, or plan to submit for publication.

The Antonym of Originality: Plagiarism (continued)

This applies not just to the original research on which the submission is based, for example numerical studies or theoretical work.

It applies also to items such as your literature review, to summaries of known aspects of the matters being addressed in the submission (e.g. previously published accounts of known properties of some of the methods studied in the submission), etc.

It is also not permissible for you to copy material from your own previously published work, or work accepted for publication, or your currently submitted work.

This applies even when that work is only an introduction, or a literature review. (Consider, for example, the copyright issues involved.)

Several leading statistical journals now use "plagiarism checkers" to ensure that authors abide by these rules.

Submitting a Paper

The corresponding author is the one who actually submits the paper.

This person should be an author who is familiar with the correct procedures for obtaining permissions from authors, for submitting a paper, and more generally for ensuring that correct procedures are adhered to.

In particular, if a relatively inexperienced author is the corresponding author then more experienced colleagues, for example other co-authors or colleagues, should ensure that that the corresponding author is fully aware of their responsibilities.

The corresponding author should ensure that all permissions, for example for using real data analysed in the paper, have been obtained.

Submitting a Paper (Continued)

All authors of a paper must be fully informed of its submission, and should have in their possession the submitted version of the paper, indicated as such.

In particular, all authors should receive copies of all relevant correspondence relating to the submission, and relating to any subsequent revision of the paper. The corresponding author should be willing and able to provide evidence to an editor of the fact that all authors have agreed to the submission of the work, and are fully aware of the contents of the submission.

Mentoring programs should be in place at the authors' institutions to ensure that all authors are fully aware of these issues. However, if those programs are absent then senior colleagues must accept responsibility for making the authors aware.

Corresponding With Editors

The process of corresponding with editors, after first a decision has been reached on the suitability of the paper for a journal, can require a reasonable amount of experience. Therefore it is suggested that an inexperienced author approach a more experienced colleague for advice.

Depending on the nature of the decision, the journal may not permit the editorial decision to be reviewed. For example, the *Annals of Statistics* usually is not willing to revisit a decision of "reject"; on the other hand, the category of "reject with resubmission," as the name implies, allows the author to resubmit, as a new paper, a revision that has been made in response to reviewer reports.

When corresponding with an editor about your work, be as brief and succinct as possible.

Authors should be professional and cordial in all their communications with editors.