Journal of
Micro/Nanolithography, MEMS, and MOEMS

SPIEDigitalLibrary.org/jm3

How to write a good scientific paper: title, abstract, and keywords

Chris Mack

SPIE

How to write a good scientific paper: title, abstract, and keywords

This is the first in a planned series of editorials covering all aspects of good science writing.

In the era of online searches and digital libraries, the importance of a good title and abstract in a scientific paper is perhaps obvious. Yet, bad titles and poorly written abstracts are exceedingly common in the scientific and technical literature. In this editorial I'll talk about some of the common mistakes made in paper titles and abstracts, then describe a nearly foolproof approach to writing good ones. The result will be a manuscript that is more likely to be accepted by a peer-reviewed journal, and a paper that is more likely to be discovered and read by the people who should.

The purpose of a title and abstract is often described as "selling" the paper: getting someone reading the title to read the abstract, and someone reading the abstract to go further and read the paper.[1,2] I have a different viewpoint. The true purpose of the title and abstract is to get the *right* people to read your paper. Let's face it, 99 + % of the scientific papers published each year are papers that I have no need and no desire to read. But there are a few papers that I shouldn't miss—and those papers are different for me than for other readers. Thus, the purpose of the title and abstract is matchmaking: matching up a paper with the right readers—those who want and need the information contained in the paper. As the 19th century English writer and eccentric Charles Caleb Colton said, "That writer does the most who gives his reader the most information and takes from him the least time." Nothing works better than a well-written title and abstract to make sure that the wrong reader doesn't waste time on the wrong paper, and that the right reader doesn't mistakenly skip over the right paper.

The title (followed by the abstract) is the first thing a reader sees, and so should be the last thing an author writes (just after the abstract). Since the abstract should be written before the title, I'll talk about abstracts first.

1 Writing an Abstract

The most common mistake in writing an abstract is to not pay much attention to it. Authors sometimes consider the abstract as an afterthought, something that can be thrown together after the "real" manuscript is written. I've even seen abstracts that are nothing more than the first paragraph of the introduction. Needless to say, such a poor abstract is unlikely to encourage a potential reader (or a journal editor) to venture further.

The abstract should be a concise, stand-alone summary of the paper, covering the following topics:[3]

- background/motivation/context
- aim/objective(s)/problem statement
- approach/method(s)/procedure(s)/materials
- results
- conclusion(s)/implications

(You may have noticed that these topics are the typical headings of the major sections of the paper itself. This is not a coincidence.) A typical abstract is about 150–200 words (although lengths can vary depending on the journal and paper type—JM[3] has a 200-word maximum), so every word must be chosen carefully. "Concise and precise" is a common maxim. If any one of these five components is missing from the abstract, there is chance of making a poor match between reader and paper. If the abstract is too wordy, readers may give up before finding out what the paper is all about.

While I'll describe my preferred approach to the abstract in a moment, let me start by mentioning a common alternative: the newspaper lede. [The neologism "lede" means simply the "lead" (guide or beginning), and is used to distinguish between other meanings (and pronunciations) of that word.] It is conventional wisdom in the newspaper world that if you don't capture the attention of the reader in the first sentence or two, he or she will move on to another article. Thus, the lead paragraph begins with a sentence containing the main point of the piece. The second sentence contains the second most important point, etc. By the time the first paragraph is finished, the classic "who, what, where, when and why" questions have all been answered.

This newspaper lede approach can be used in the scientific abstract as well: if I have only one sentence to convince the reader to continue reading, what would I say? Then ask the same question for each succeeding sentence. There are certainly some good abstracts that have been written using this approach, but I don't like it for two reasons. First, it takes a very good writer to make the newspaper lede form of abstract work. And most of us just aren't good enough writers to pull it off. It's also easy to leave off one of the important five topics that every abstract should contain. Thus, even an extremely well-executed newspaper lede-style abstract may not do the best job of matchmaking between the paper and the reader. Second, there is a better approach: the structured abstract.

2 Structured Abstracts

For the past 25 years, structured abstracts have become required in most medical journals, though they are not very common in engineering and the physical sciences.[4] I hope this will change, since I am a big fan of the structured abstract. Simply put, the structured abstract formalizes the five topical areas mentioned above by adding subheadings and subsections (the "structure") into the abstract. While the exact structure can be modified to suit the topics of the journal (or even the specific paper), in engineering and physical sciences a five-structure format is probably best: background, aim, approach, results, conclusion. Each subsection should contain one to two sentences, answering the following questions:

Background: What issues led to this work? What is the environment that makes this work interesting or important?

Aim: What did you plan to achieve in this work? What gap is being filled?

Approach: How did you set about achieving your aims (e.g., experimental method, simulation approach, theoretical approach, combinations of these, etc.)? What did you actually do?

Results: What were the main results of the study (including numbers, if appropriate)?

Conclusions: What were your main conclusions? Why are the results important? Where will they lead?

The benefit of the structured abstract is twofold: it forces the author to include information from all five categories, and it makes these five sections easy to find and access. But while it is logical that structured abstracts will be better than unstructured abstracts, there is in fact proof that this is so. The preeminent researcher into the efficacy of structured abstracts, James Hartley, reviewed some 31 studies that had been performed by 2004 and found that these studies demonstrated the superiority of structured abstracts.[5] His review, as well as others,[6] showed that structured abstracts

- contain more information;
- are easier to read;
- are easier to search;
- facilitate peer review; and
- are preferred by readers and authors.

This is all well and good, but JM^3 does not use structured abstracts. Maybe it should, and maybe it will someday, but today it does not. The structured abstract is still important because it can be used in what I call the *structure method* of abstract writing. The method is quite simple. First, write a structured abstract. When you are finished and satisfied with the result, simply delete the subheadings and combine all the lines into one paragraph. Finally, reread this new abstract, and change sentence beginnings to increase readability and flow, if needed (though usually this will not be necessary). The result will be a well-written and effective abstract with most of the benefits of a structured abstract.

To illustrate, here is an abstract for a paper that I am just now finishing. First, I wrote a structured abstract:

Background: Photoresist development rate can be defined microscopically (the development rate at a point) or macroscopically (the propagation rate of an average resist height). In the presence of stochastic noise, these two rates will be different.

Aim: In order to properly calibrate lithography simulators, the difference between these two definitions of development rate should be quantified.

Approach: Using theoretical derivations and a stochastic (Monte Carlo) resist simulator, the propagation rate of a resist surface is characterized in the presence of stochastic variation in the resist deprotection concentration and a nonlinear development rate response.

Results: The resulting propagation rate can be more than an order of magnitude higher than for the case of no stochastic noise. Correlation in the development rate creates an effective surface inhibition over a depth into the resist of several correlation lengths.

Conclusions: The differences between microscopic and macroscopic dissolution rate can have an important effect on how development rate models should be calibrated, depending on their use in continuum or stochastic lithography simulators.

Then, deleting the subheadings and line breaks, a traditional abstract format is obtained. I added a transition clause at the front of the last sentence to make the abstract flow better, though this small change could easily have been left out.

> Photoresist development rate can be defined microscopically (the development rate at a point) or macroscopically (the propagation rate of an average resist height). In the presence of stochastic noise, these two rates will be different. In order to properly calibrate lithography simulators, the difference between these two definitions of development rate should be quantified. Using theoretical derivations and a stochastic (Monte Carlo) resist simulator, the propagation rate of a resist surface is characterized in the presence of stochastic variation in the resist deprotection concentration and a nonlinear development rate response. The resulting propagation rate can be more than an order of magnitude higher than for the case of no stochastic noise. Correlation in the development rate creates an effective surface inhibition over a depth into the resist of several correlation lengths. These results show that the differences between microscopic and macroscopic dissolution rate can have an important effect on how development rate models should be calibrated, depending on their use in continuum or stochastic lithography simulators.Note that while structured abstracts are typically longer than traditional ones, the 166-word length here is right on target for most journals. If anything, the approach and results sections could have been expanded slightly.

Additionally, the structured method of abstract writing also helps to avoid useless but all-too-common phrases like “in this paper” and “we report” or “will be discussed.” The abstract should talk about the work, not about the paper; phrases like “is discussed” turn your abstract into a table of contents rather than a summary of the paper. Do not use the first person (“I” or “we” or “the author”). Also, there is rarely a need to use phrases like “new” or “novel” in the abstract, since it is only the novel results that should be mentioned. You know that you are finished writing the abstract when there are no words that could be taken out without changing the meaning.

3 Important Additional Thoughts on Abstracts

Why should the abstract be written after the entire paper is complete? The reason is simple: if not, it is unlikely that the abstract will be accurate. A study of six highly regarded medical journals in 1999 found that about 40% of the abstracts studied contained information inconsistent with the body of the paper, or information not found in the body of the paper, or both.[7] The most likely cause of these

errors, after just plain sloppiness, would be changes made to the paper after the abstract was written. Such errors and inconsistencies can largely be avoided by leaving the abstract-writing task till after the body of the manuscript is completely finished. The structured abstract can help make the abstract more informative, but it is still up to the diligence of the authors (and journal editors and reviewers) to make sure the abstract is accurate. There is a three-part test that should be applied to your abstract when you are finished:

1. Is all of the information in the abstract consistent with what is written in the body of the paper?
2. Can all of the information found in the abstract also be found in the body of the paper?
3. Is the important information of the paper found in the abstract? Are any key words from the paper missing from the abstract?

The abstract must be self-contained, and in general should not contain citations to other papers. If a citation is required (for example, if the paper is a response to a previous publication), the full citation must be embedded in the abstract. Do not use abbreviations or acronyms in the abstract, or at the least spell them out the first time they are used. Do not refer to figures or tables from the body of the paper, or use words or descriptions that will only make sense after the full paper has been read. Trademarked terms should be avoided as well.

As for all writing, keep the audience in mind. If you are writing a spectroscopy paper for a spectroscopy journal, you can surmise that all of your readers will be spectroscopists, with a certain background knowledge assumed. A paper on that same topic for a more general optics journal may require an extra sentence in the background section to inform the reader that the topic is within the field of spectroscopy, within a certain subfield, etc.

Finally, an important goal of the abstract (and the title to be discussed next) is to make the abstract as specific as possible while still describing the full range of work reported in the body of the paper. If the paper measures only thickness uniformity of a film, the abstract should not make the more general claim that the paper measures "film uniformity." If the paper simulates the scattering properties of one-dimensional gratings (but not more general objects), an abstract that merely states that scattering simulations were performed could mislead the reader into thinking that the work was applicable to more general objects. On the other hand, if the thickness and compositional uniformity of the film were measured, saying only "thickness uniformity" in the abstract is too limiting and does not describe the full scope of the paper.

4 Titles

When the abstract is written, it is now time to write the title. Unfortunately, it is against human nature to write the title last. Instead, the title is often the first thing written, at the top of that blank document that will soon become your manuscript. It is important to consider these first words as the *working title*. When the manuscript, and the abstract, are finished, it will almost surely be necessary to revise the title.

It is probably impossible to define a universal procedure for creating a good title—there is no equivalent "structure method" for writing a title. There are some basic guidelines, however, that make use of the structured abstract to guide the creation of the title. In general, the title should reflect the aim and approach of the work. Depending on the audience (and the specificity of the journal), some of the background may have to be included. Rarely are results and conclusions even hinted at in the title. Let's look at each of these items through the use of an example.

Unlike the worlds of newspaper reporting and marketing press releases, the title of a scientific paper should describe the aim of the work, not the results. Thus, a good title might be

Impact of temperature and pressure on the compositional uniformity of sputter-deposited aluminum alloys

The following news-style title, on the other hand, is not appropriate:

Optimizing temperature and pressure improves sputter-deposited aluminum alloy films

Note that the good title is essentially a statement of the aim of the work. Often it is important to mention the approach used as well, though an experimental approach is generally assumed if it is not mentioned. If the study had been based on simulation (or some other approach), however, this would generally be included in the title:

Impact of temperature and pressure on the simulated compositional uniformity of sputter-deposited aluminum alloys

The title should be as specific as possible while still describing the full range of the work. For example, if only one aluminum alloy was being studied, that specific alloy should be mentioned in the title. If only aluminum alloys are studied, the title shouldn't say "sputter-deposited metals" or "sputter-deposited alloys." On the other hand, the title shouldn't say "aluminum alloys" if gold was also included in the study. If the title had said "uniformity" rather than "compositional uniformity," the reader could easily have believed that the paper was about thickness uniformity or some other parameter. And if only sputter deposition was studied, then leaving this information out would make the title insufficiently specific.

A conflicting goal of the title is to be as short as possible (in 2011, JM[3] titles ranged from 4 to 21 words in length, with an average length of 11.5). Specificity can often be improved through the use of more words, but a title that is too lengthy may not be read.[8] Finding the best compromise between descriptiveness and brevity is where the art of authorship comes in to play. Going back to our example, here is a title that sacrifices too much specificity to obtain brevity:

Impact of process parameters on the uniformity of aluminum alloys

A good test for your title is to answer these questions: Does the title of your manuscript, seen in isolation, give a

full yet concise and specific indication of the work reported? Would someone interested in the exact topic of your paper, reading this title, be inclined to read the abstract?

Avoid being overly clever with the title—a pun or a play on words may be great fun, but it is unlikely to help your article be found by a search engine (and can be easily misunderstood by an international audience). Titles should also be free of jargon unlikely to be understood by those not intimately familiar with the topic, and should not contain acronyms or trademarked terms. The overall goal should be a title that is clear and informative.

5 Keywords

This brings up the next topic—keywords (also called "subject terms"). We are quickly passing out of the days when most people find your article by flipping through the print version of the journal. Today, your article is unlikely to be widely read unless it comes up relatively high on a Google or Google Scholar search-results list. The first and most important thing you can do to insure that your article is found by readers looking for it is to do a good job of writing the abstract and title. Following the advice given above should help. After that, you must decide on appropriate keywords.

The important idea behind identifying the key words to be listed under the abstract as "subject terms" is simple: if you were looking for an article on exactly the topic of your manuscript, what key words would you type into a search engine in order to find it? Chances are, you would start with only two to four words or phrases. If that resulted in too many hits, or too many off-scope articles, then you would refine your search by adding one or two more phrases. These are the words or phrases (plus all of their common variants and synonyms) that should be included in the list of subject terms.

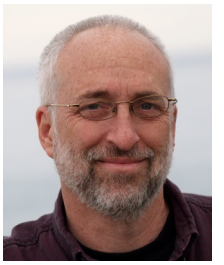
Once you have a good list of keywords, go back and look at your title and abstract. Are these keywords found in the title and abstract? If not, someone searching for your article may easily miss it. The most important keywords should be found in the title, and in the abstract several times.

6 Summary

A structured abstract is a proven way to give readers the information they need in an accessible and readable format. The structure method of abstract writing proposed here can provide many of the benefits of a structured abstract for journals (like JM^3) that do not (yet) use structured abstracts. This structure can also aid in the writing of the title, using information from the aim and approach subsections.

To be sure, I have made just about every mistake described in this editorial. But I can certainly tell that as I have adopted the structure method of abstract writing my abstracts have become far more informative and useful. And if I keep practicing my own advice, I suspect that I will continue to improve. Someday, JM^3 may adopt a formal structured abstract for all papers. Until then, I think the structure method of abstract writing described here might help you as much as it has helped me. Combined with a properly devised title, you will then have a paper that is off to a very good start.

Chris Mack
Editor-in-Chief



References

1. R. G. Driggers, "Editorial: how do you write a great abstract and why is it important?," *Opt. Eng.* **49**(6), 060101 (2010).
2. W. Rhodes, "Editorial: the abstract as a marketing tool," *Opt. Eng.* **49**(7), 070101 (2010).
3. National Information Standards Organization, "Guidelines for abstracts," *ANSII/NISO Standard* 239.14–1997 (1997).
4. R. N. Kostoff and J. Hartley, "Open letter to technical journal editors regarding structured abstracts: this letter proposes that structured abstracts be required for all technical journal articles," *J. Inform. Sci.* **28**(3), 257–261 (2002).
5. J. Hartley, "Current findings from research on structured abstracts," *J. Med. Libr. Assoc.* **92**(3), 368–371 (2004).
6. C. Zhang and X. Liu, "Review of James Hartley's research on structured abstracts," *J. Inform. Sci.* **37**(6), 570–576 (2011).
7. R. M. Pitkin, M. A. Branagan, and L. F. Burmeister, "Accuracy of data in abstracts of published research articles," *J. Am. Med. Assoc.* **281**(12), 1110–1111 (1999).
8. H. R. Jamali and M. Nikzad, "Article title type and its relation with the number of downloads and citations," *Scientometrics* **88**(2), 653–661 (2011).