Research Methodology
Prof. Soumitro Banerjee
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

Lecture - 69
Ethical Conduct in Science: Ethics in Scientific Publication, Part 01

In the last class, I said plagiarism is an academic crime, which means you cannot lift somebody else's language, phrases, text into your own thesis or your own paper. But the question might arise how can we write appropriately so that any plagiarism accusation can be avoided?

There are two ways. One, if you really need to put somebody's language in your thesis or paper -- there may be reasons for doing so -- then always put that in quotation marks, so that the reader knows upfront that it is quoted. Start the sentence with the person's name, the authors name, from which it is quoted, like Einstein said ... then quotation mark. It is not like something put in quotation mark and then you put the reference, so that the reader has to go to the reference to find out whose quotation it is. That is not the right way of quoting. Quotation means, whose quotation that should be in the beginning of the sentence.

When do we quote? (A) when it is from an authority in the field, (B) when I am trying to say something that is a part of a government policy document, say the National Education Policy, I am talking about that. Then I have to quote from there. Thirdly, there are situations when some other author has written something, and in writing that, I find that if I change anything, even a word, then it will change the meaning or it will not reflect the technical content in appropriate way. So, in that case I can quote that, but put that in a quotation mark.

But use quotations sparingly. I have cited three possible situations where we might need to quote directly. In all other situations, in scientific literature we do not quote and we use what is known as paraphrasing. That means, we state the same thing that was done, but in our own language. The problem is that, sometimes a student, when he or she is trying to state what somebody else has done, is influenced by the language of that person and as a result, unknowingly quoting happens without the quote. So that has to be avoided.

There are few ways of paraphrasing. Wherever it is possible to put a synonym in place of a word, do that. Wherever it is possible to change the structure of the sentence – where you put the subject, where you put the verb, where you put the adjective, that structure of the sentence – where if it possible to change you can change that.

You can change the parts of speech. Wherever a sentence is written in active voice, you can change it to passive voice. If it is in passive voice you can change it to active voice. If you ask me, I employ a combination of these strategies in order to paraphrase somebody else's work in my own language.

If you have written something in a paper and you want to state the same thing in another paper you should not really copy in your own work. These also be should be paraphrased; otherwise it will become self-plagiarism.

Another source of plagiarism is, when we read somebody else's paper, we take notes. A student, normally a PhD student, would normally start taking notes from the first year of his work and finally, the things that are noted will find their way into the final thesis written in the 5th year. So, unless when you are noting appropriately, there might be an error.

What happens? You should actually make it a practice of taking notes while reading a paper. Earlier people used to write. Now, since everything is on the computer, people often have lost the habit of writing. Please do cultivate the habit of writing in a note book. What is the salient point? What is the remarkable a point made in a particular paper? Take your own notes. This will come in handy when you finally write your thesis.

But when you do take note, if you write in your own language what you have learnt from that paper, it is fine. But if you like some language, some expression, some phrase from that paper, if you write it verbatim, then in your note itself put it in inverted commas, so that later you know that it was quoted. Otherwise you would tend to forget whether it was quoted or it was your own language. And then, when you write that in your final thesis, if it had been quoted, then unknowingly you commit the error of plagiarism.

So, these are the things that need to be avoided.

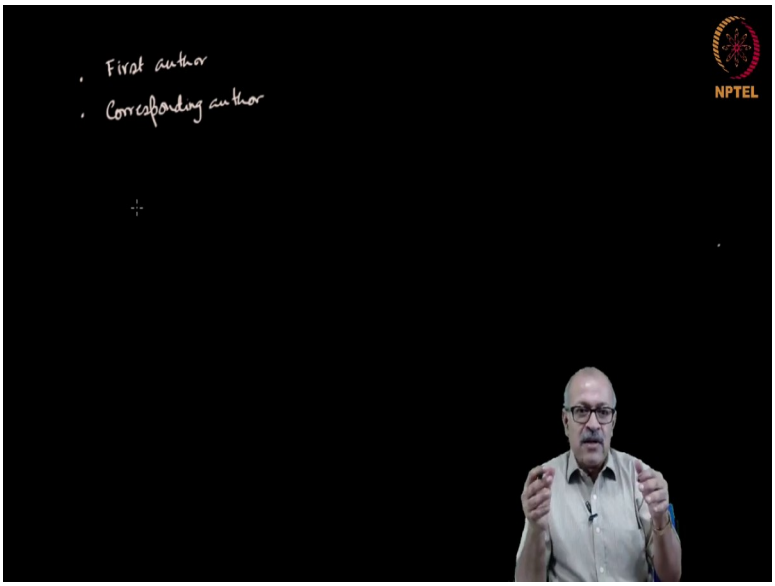
Now let us come to the question of academic authorship. Many scientific malpractices happen in academic authorship, and therefore, it is necessary to understand the ethical practices in academic authorship.

Most papers come out of a research group comprising the professor and under him or her there would be a number of research students working. In those cases, the research students will be either master's students or PhD students or postdocs. In all cases it is a joint work of the supervisor as well as the student and therefore, they will be coauthors in the paper.

If multiple students had participated in a piece of work, then there will be multiple authors. If it has been a collaboration between a group and another, there will be multiple authors. But presently let us focus on the result of a group publishing their own work in the paper.

Out of all the authors, there may be 5, 6 authors, two authors have some special position. Normally the first author is assumed to be the one who has done the bulwark of this paper, starting from generation of the idea to the actual execution of the research program. This rule is not followed in a few fields, for example, in experimental particle physics where there will be hundreds of authors. The paper itself might be three pages and the names of authors might be more than that 4 pages. In that case the standard practice is to write in a alphabetical order. So, the first author has no special significance. That is also followed in mathematics, for example. But in most other fields the first authorship carries a weight.

(Refer Slide Time: 08:20)



The next is the corresponding author. So, the important positions are the first author and the corresponding author. Who is a corresponding author? Corresponding author is one who will communicate the paper to the journal, and when the reviews come, it is the corresponding author who is responsible for responding to that, responsible for submitting the revised version of the paper. Finally, when it is published, the corresponding author's email address is there in the paper. So, that if any reader has a query, that will be addressed to the corresponding author and it is the corresponding author who is responsible for responding to that. So, for a long time the corresponding author will have a very specific role.

Normally the supervisor becomes the corresponding author, because he or she will remain in the institution for a very long time while the students will come and go. They will pass out and go somewhere else. So, this is the usual way.

When a student gets into a research group, the first few years, apart from the coursework, is spent in learning the use of the equipment, learning the programming methodology, the techniques in the field, more or less absorbing the culture of the discipline and things like that. In that stage the supervisor may ask the student to help another senior student or ask him or her to execute a program, execute a part of the research work, some procedure of the research work. So, in that sense there is a contribution of that junior student in that work.

The junior student will become a coauthor. Maybe not the first author or the corresponding author, some other position. But after the first 2 years are over, when this student actually starts doing research proper, then the student is supposed to generate the idea and to execute it. If the student generates the idea and executes it, then he or she rightfully becomes the coauthor.

But if the supervisor still has to generate the idea and the student only executes procedures according to the instruction of the supervisor, then rightly the supervisor should be the first author. I mean first author and the corresponding author. The decision as to who becomes the first author, what is the order authorship, normally is taken by a discussion within the group and a impersonal judgment of who contributed what.

But remember, even if you have done most of the work, if it is only procedural, not intellectual component, then the student should not become the first author. But everybody feels that I have done lot of work and therefore, I should be the first author. That judgment should lie with the supervisor. Ultimately the supervisor's decision is final.

And the supervisor may decide that, no, you are not the first author. I am the first author because I generated the idea and you only executed things as per my instruction. Therefore, I always advise students to learn the method of thinking, so that you can generate the idea. Learn the big picture, so that you can generate the idea, what can be done.

There are some extreme situations in which a student might generate the idea, might execute the work. And then he or she might feel that, I have done everything and therefore, I can be the solo author. It should be a single-authored paper. No. You cannot do that unless the supervisor gives consent in writing to doing that.

A PhD work is a joint work of the supervisor and the student. If the supervisor encourages a student to write a single-authored paper, permits him in writing, then you can publish a single-authored paper. Otherwise a paper always is a coauthored one, between the supervisor and the student. It is considered a scientific malpractice for a student to publish a paper without the consent of the supervisor. It is a scientific malpractice. Avoid that.

What about the data generated earlier in a group? Normally, before a student joins, other PhD students, postdocs or master's students had done work and therefore, generated data. Now that data generated by earlier students, that forms the basis of the current work. Therefore, you might need to refer to that data and in some cases you might need to put that data in your thesis.

But you should not put that data in your thesis in a way that implies that it has been generated by you. You have to very clearly demarcate that this was work of earlier students and this is what I have done. This is a part of scientific ethics.

Before a paper is published, it is the responsibility of the corresponding author to ensure that every member of the group, whoever are the coauthors, they all are on the same page regarding the scientific cohesiveness of the paper, whether or not the data lead to the conclusion that has been drawn in the paper. They have to be on the same page on that. That means, everybody should have access to every aspect of the paper and they should concur. That is necessary.

Now let us come to the issue of multi-departmental collaborative work or even multi-institutional collaborative work. Most collaborative work happen between people with complementary expertise: I am expert in something and you are expert in something else. I do not know the details of that area and you do not know the details of this area. When we put our expertise together, that might result in a wonderful work.

When that happens, and the resulting paper is to be published, then there has to be an understanding as to who will be the first author, who will be the corresponding author. There has to be understanding between the groups, because both could be the first author as well as the corresponding author. In this, it is necessary to be completely impersonal. An impersonal judgment is necessary.

The principal investigators should be responsible for ensuring that every member of the group has access to the complete data and are on the same page regarding the data, the analysis of the data, and the conclusions. All authors of the group have shared responsibility regarding the work, because if a dispute arises later, somebody challenges the result, then everybody in the authorship are responsible. One cannot say that I was only looking at that and I did not know what was happening in the other side.

At least that much responsibility one should have, so that you are sure that the data are sound. Every author in the whole group, the multi institutional group I am talking about, should have access to the complete data. You might not be a specialist in that area. But you have to have access to the complete data, so that, if you want you can cross-check the validity of the results.

Everybody, every author, should agree to be a coauthor on this paper, because if it is later proved that the conclusions are false, then it impacts the reputation of the scientist. Therefore, everybody in the paper, everybody participating in the work, should agree to being a coauthor of the paper and it is the responsibility the corresponding author to ensure that.

Early at the start of the project, there should be understanding between the groups as to who maintains the data and in what form, so that if a dispute arises later, it is easy to locate the primary data. It is important to have the primary data, because that is what ultimately needs to be referred to if a dispute arises.

Now, sometimes it happens that a supervisor is collaborating with another person in another institution, may be abroad. A part of the work he asks a student to do: maybe executing a program, maybe writing a code, maybe executing a procedure, developing a cell line, whatever it is. In that case the student might become coauthor of that paper. But the question arises: Can the student include that paper ultimately when he or she writes a thesis?

Yes, it is possible. But in that chapter of the thesis where that paper is included, it has to be very clearly stated that it was a collaborative work and this part is my contribution. It should not be presented in a way implying that the whole work is the student's contribution. That is a part of scientific ethics.

I have earlier said that in science, it is desirable to have openness. Secretiveness is not conducive to the health of science, because our predecessors, our earlier generation of scientists, made their findings freely available to mankind. That is how we have come to know about those findings, and so it is our responsibility to make our findings freely available to mankind. That is part of the ethics of science.

But certain areas demand confidentiality. If the subject matter of the research is sensitive in some sense, either for defense or for something else, then there might be an agreement between the PI and some kind of a funding agency that the result of the research will be kept confidential. In that case confidentiality has to be honored.

If the supervisor wants to patent a piece of work hoping that it will be used by some company to produce something, in that case it might not be put in a paper, and so that will remain the intellectual property of the scientist until it is sold to some company. Then it becomes the intellectual property of the company. In those cases that material, that idea, cannot be included in a research paper because research papers are open.

If some such project is going on, and a PhD student is working in such a project. Then it is somewhat problematic for the PhD student to be included in a work that demands secrecy, because ultimately, before submitting the thesis, the student will be expected to publish research papers.

Secondly, at the end of the day, he has to submit a thesis which becomes in the public domain. If the content of the work is something that cannot be made public, then that cannot be included in the PhD thesis also. That leads to various problems. So it is necessary that a PhD student works on problems that are not such confidential in nature. Confidential work should be done normally through sponsored project and through project personal who are not PhD students.

There are certain areas that might require confidentiality, for example, if the work involves human subjects. Then the identity of the human, the privacy of the person, has to be honored and so the names should not be diversed. Things that are classified, as I said, the ones that are important for the State or important for certain communities. Those information may have to be classified, not made public.

In case of research on history, or history of science, a kind of situation may arise where certain private papers, letters, diaries, and such documents come to the access of a researcher. Then one has to exercise judgment as to whether or not to make everything public. Because there are private aspects of a person’s life. That person might be dead, but if he or she were alive, might not like to make these public. So that judgment has to be exercised.

When we submit a paper to a journal, we have to transfer the copyright to the publisher of that journal, and as a result, the intellectual property is transferred to the company that is the publisher of the journal. After that, if I want to use that material that I had written in a paper, I want to use that material in a book that I might be writing, then also the permission of the publisher has to be taken.

Normally these papers are behind the pay-walls. So, if anybody wants to see the paper, then one has to pay or the institution has to pay through the library for subscription. In either way the reader has to pay in some way or other to read the paper.

Nowadays we have another model called ‘open access’ model in which the reader does not pay. It becomes free. But the writer, the author, has to pay in order to get the paper published. Many journals have now taken to this and some journals are mixed kind, both types: reader pays as well as the author pays kind. So, there are all kinds of spectrum.

Now, in the ones that are ‘reader pays’ type, which are behind the pay walls, there also there is something called the ‘principle of fair use’ that is applicable. You cannot distribute the paper directly without consent from the publisher. But photocopying a paper for personal use, or photocopying a few pages from a book for personal use: these are within the range of fair use and these are allowed. Photocopying the whole book and distributing freely is not allowed. That is against intellectual property rights.

Before I go on to another topic, let me mention that over the last few years, a new phenomenon has happened which is called ‘predatory journals’. Predatory journals are journals which offer to publish your paper practically without any review within a short time, against payment. So, they take money from you to publish your paper.

These are called predatory journals. Unethical practice. Never submit papers to such journals. You might ask: how do I know which are periodary journals? There are reputed journals, which also take payment from the authors. As I said, the open access journals are like that. The author pays. The rule of thumb is that, submit to those journals which you regularly read. You know that important papers are published in those journals. Submit your papers to those journals. Then you will be safe.