Studies in the theory of scientific research. Part 1. Theory of information transfer by scientific publications: Basic rules to optimise the effect of scientific communication[1]

Kao Lin, Shiu Sisako, and Josef Horák[2]
Institute of Chemical Technology, Prague

(Original received: April 1 1969)
Translation received: October 20, 2009

This paper addresses a number of elements involved in the process of sharing information and knowledge, i.e. needs, transfer, and disclosure in scientific environment. It concerns standards and conventions, which science communicators should learn as they develop into true scientists. In other words, scientists beginners should learn how to present their work to the scientific community; how to select information and put it into a desirable format. The most important medium for knowledge production and exchange between scientists are peer reviewed journals. Unfortunately, in chemical literature only a negligible attention is paid to teaching how to communicate science successfully. Therefore, the aim of this paper is to give beginners the basics of how to write a high quality scientific publication.

Paper title

The title is the most important part of the paper. Once the subject is known, one can make up the title, taking into account certain important factors. Beginners might think that the title must exactly reflect the subject. However, some degree of freedom is allowed. It is an absolute imperative to make up a flashy title to arouse reader's curiosity and get them go to the article, convinced that it is an extraordinary masterpiece by a respectable author from an institution of world renown. Therefore, the author should fill the title with as many fashionable terms as possible, such as optimisation, model, simulation, computer, structure, relationship, quantum, mechanism, system, process, and so on, including all derivatives of these terms and their combinations. To extend this list, please check a recent issue of *Chemical Abstracts*.

The following story illustrates how to coin a flamboyant title. A chemist employed in the cosmetic industry requested a collaborative lab to conduct a minor test to see if a hair lotion released acids if treated by an alkali at room temperature. Collaborative lab scientists carried out several experiments with as much care as befitted the negligible amount of money offered for the job. No sooner had the study been completed than the absent-minded colleague from the cosmetic industry apologised for sending a wrong bottle with ethyl acetate instead of hair lotion. To make up for the lost valuable time, the lab staff decided to publish the obtained experimental results. Now, an inexperienced researcher would entitle this work "A study of ethyl acetate hydrolysis by an NaOH solution". However, this simple title would not arouse any interest even in a junior student, let alone an old research warhorse. An experienced scientist would think along the following lines: obviously, living organisms contain all kinds of chemical compounds, including esters. In addition, they live at temperatures between 0 °C and 50 °C and are mostly made of water. Hence, the title "A study of the mechanisms governing metabolic reactions in biochemical models at conditions analogous to physiological. Part 1. Solvolysis of compounds containing an ester functional group: Hydrolysis of ethyl acetate in the presence of sodium hydroxide at a higher concentration range".

It is unnecessary to discuss the blatant difference between the first and the second title. The fact

Dedicated to Professor Josef Horák on the occasion of his 75th birthday.

[1] Translated and adapted from Czech by Ivan Wichterle, Institute of Chemical Process Fundamentals, Prague, E-mail: wi@icpf.cas.cz, and Grozdana Bogdanić, INA d.d., Zagreb, E-mail: grozdana.bogdanic@ina.hr. English language revision by Dado Čakalo, Institute for Medical Research and Occupational Health, Zagreb, E-mail: dcakalo@imi.hr.
This paper was published 40 years ago in *Chemické Listy* **63** (1969) 414–417. It is surprising how it still holds true after such a long time. Therefore, we asked Professor J. Horák for the permission to translate it into English.
One should not overlook a little detail that the manuscript was submitted on 1 April (see the attached facsimile). The meaning of the first fictional author name is clear to everyone. The second name, Shiu Sisako, looks like a genuine Oriental name, but its literal translation from colloquial Czech is "I'm sewing myself a jacket" (= šiju si sako). It is worth mentioning that the paper is quoted in the Web of Science and the authors Kao Lin and Shiu Sisako have forever been included in the database of world scientists. The reasons for including the two "authors" and for the long title are clarified in the paper.

[2] E-mail: Josef.Horak@vscht.cz

that this work has been presented as the first part of an extensive series is not at all binding for the author. Future circumstances could force the author to switch to yet another more important subject. If the paper is not written in English, one should also take care that the title gets a duly impressive translation for *Chemical Abstracts*.

Introduction

The introduction is a road map for the paper. It should capture reader's interest and lure them to read on for the rest of the text. It should convey a lot of information to let the reader know what the subject is and why it is important. It should also argue why the work has been carried out and indicate whether the work is theoretical or practical, even if it is sometimes difficult to decide between the two. Some mentors suggest that the choice should depend on content. For example, if the paper deals with solving a theoretical problem and provides mathematical models and framework, then the work should be classified as theoretical. If the results can be used in practice then the paper is user-oriented. However, this approach has a downside, as it cannot help to classify papers dealing with nothing. This is why we suggest that papers should be classified based on what they *do not* involve. This means that a paper could be classified as theoretical if the results are not applicable in practice. Alternatively, it could be classified as practical if theory is completely ignored. Thus, a paper could be described as involving application of theory in practice, if it deals with neither theoretical nor practical problems. However, this classification does not apply for theoretical papers, which might also be used in practice. Fortunately, virtually no such case has been recorded in chemical literature, and the proposed classification stands unchallenged.

It is assumed that the author has at least some idea about what to write, but the central idea cannot come out of the blue. It needs to be introduced in a way that makes sense to a reader. The introduction should give reasons why a problem has been studied. The reasons presented in the introduction do not have to correspond to the actual reasons; moreover, this is not even advisable. A topic must be carefully selected to avoid troublesome questions. Therefore, one should limit research to the ordinary and easily available chemicals, which can be borrowed or obtained in a way that does not require author's own effort to synthesise them. The same approach applies for the method and equipment used. Evidently, these reasons must not be revealed in the introduction. The first thing to do is to come up with as fascinating a subject as possible. Naturally, the proposed subject would have something to do with the experiments. It should end with a compelling story, an attractive quotation, an interesting question, or with a stirring example to give an idea why this topic matters and to invite the reader to follow through this interesting discovery. The reader must be convinced that the paper treats an enormously important subject, which nobody else has addressed before. Then a discreet note may follow suggesting that despite all adversity, the author has figured out a simple and elegant solution to the problem.

The following example illustrates how to write an excellent introduction; for instance, a part of a research programme concerns hydrogenation. The focus is on hydrogenation of benzene because a bottle of benzene was just at hand on the laboratory shelf. The catalyst used was once received as a free sample, and the authors have slightly adjusted for measurement a distillation apparatus borrowed from a colleague from the lab next door. Various mixtures of benzene and hydrogen are then put into this "reactor", and a variety of resulting data are plotted in a series of assorted diagrams. Generally, the obtained results are good-for-nothing, but this has never put in doubt whether to publish or not. The answer is always positive.

The significance of the results could be duly emphasised in the introduction by saying something like: "Large-scale industrial reactors are rather complex systems possessing parameters which cannot be easily generalised. Consequently, only a fraction of gathered data could be used to design new units. Computer models, which are able to treat a process with many degrees of freedom and which involve computer graphics, almost invariably blur the presentation of otherwise reliable-looking data that in fact do not entirely correspond to reality. Since data from large-scale reactors (or from computer simulations for that matter) cannot provide dependable input for process design due to inflexible handling and control, obviously the best solution is to apply laboratory equipment and its subsequent scale-up. Laboratory model reactors have proven to be very convenient for conducting experiments, for validating the properties of materials and for modelling relationships." This should do to present the problem and stress its importance. A well written introduction should not fail to make the reader uneasy. It should be a challenge to make the reader aware that until now he has been unaware of such a serious issue. In addition, he should feel ashamed for overlooking the possibility to solve the problem earlier. "This is why all experiments were carried out using the laboratory-scale reactor, taking into account all the difficulties ensuing from its complex behaviour." This statement is

very true for the 'reactor' under study and shows that the problems mentioned above have completely been overcome due to author's clever approach. Then the article could continue as follows: "Even though investigations carried out in laboratory reactors are fundamental for the design of industrial processes, they have not received due attention. The aim of this study was to establish the basic phenomena using a simplified model reactor and to contribute to practical applications of the chemical reactor theory".

Experimental part

It is relatively easy to compile the experimental part. The best way to do it is to find an article in literature that looks convincing enough and then copy a part of it with slight modifications. The text must be well structured to make the impression that every single detail is included and that the work was carried out with utmost precision. For example, an inexperienced researcher would say: "The temperature was measured using a mercury thermometer". By contrast, an experienced scientist would write this: "The temperature of the reaction mixture was measured using a mercury thermometer 4.22 mm in diameter, and mercury column length of 122.74 mm at 20 °C. The thermometer was calibrated using the standard Zeiss thermometer with certified accuracy of 0.01 °C. It was placed in a bulb so that the distance from the bottom was 23 mm. Comparative experiments confirmed that the position of the thermometer did not influence temperature reading. Temperature was read 73 seconds after immersing the thermometer."

However, one serious problem has to be considered when writing the experimental part. There is always a risk that somebody would like to reproduce our results, no matter how senseless the subject of our work is. In order to eliminate suspicion that the work was carried out irresponsibly, the description in the experimental part must create a strong feeling that the amount of information is sufficient to reproduce the experiment, yet a key item must be omitted. A thorough analysis of the problem should tell which information is to be denied. This may require certain experience, which is largely gained by reading current patent literature.

Theoretical part

Mathematical procedures are sometimes very useful to correlate and evaluate results, but they primarily serve as decoration. Partial derivatives, multiple integrals, vectors, tensors, infinite series, and limits are particularly impressive. This is why the author should use the most complicated mathematical means available. Here is an example; an unskilled scientist would write: "Chemical reaction can be described by the kinetic rate equation of the first order". An expert researcher, however, would find the most complicated and the most general equation in a chemical engineering textbook describing an industrial chemical reactor and accompany it with the following text: "A mathematical model that can describe the behaviour of the whole experimental equipment is absolutely needed in order to evaluate measured data. The equation (1) can be derived from the differential balance and is valid for non-isothermal, non-adiabatic, imperfectly stirred, and non-steady-state flow reactor in which activity coefficients, diffusion coefficients, density, specific heat, reaction enthalpy, activation energy, and thermal conductivity depend on composition, temperature, and pressure." Follows the equation, which should occupy at least half a page, and then the text can continue as follows: "It was necessary to make certain simplifications since the relation (1) cannot be resolved analytically, and the numerical solution is not feasible even with the use of the most powerful computers". The author may then simplify the equation based on assumptions like both input and output flows are negligible, mixture behaviour is ideal, the reaction runs isothermally, etc. Finally, after several pages describing these reductions, the rate equation of the first order would be obtained with the following comment: "After having implemented simplifying solutions, the equation (1) was successfully converted into an analytically integrable expression. As a result of integration, the equation (32) was obtained and used for evaluation of experimental data."

The advantage of the procedure described above is obvious. The author has managed to introduce a cornucopia of equations, several dozens of symbols, two pages of description, one page of simplifications, and in addition, he has brought into play a lot of trendy terms such as non-ideal, isothermal, unstable, activation energy, etc. A quick look at this article must convince everyone that it was written by a most erudite author, who is an excellent specialist in the field. In addition, the vast amount of equations will result in tremendous respect of most chemists (even the authors' co-workers), that will discourage further reading. As a result, the author will retain reader's interest and respect, and his reputation will remain spotless.

Discussion

The Results section, which usually follows the Experimental section, brings the only factual infor-

mation, leaving no room to discuss the significance of the results obtained. Therefore, as a rule, every paper must include a discussion in order to show how profound and erudite the author actually is. In addition, the discussion can criticise other authors' findings, as needed.

However, a certain inconvenience arises for the author of a paper in which nothing is solved and in which there is nothing to discuss. In such a case, it is recommended to repeat a piece of information from tables and diagrams of the experimental section. Sticking to ethyl acetate hydrolysis, here is an example: "An interesting conclusion can be drawn from experimental results, namely that the concentration of ethyl acetate decreases monotonously and asymptotically approaches zero after a sufficiently long time. By contrast, the concentration of sodium acetate shows an increase." The discussion is more elegant (and more assertive of the author's intellect) if an extreme could be found on any suitable curve in a diagram. For this purpose, the author should make every effort to "discover" variables which show, say, a maximum in a plot. A suitable variable for the hydrogenation reaction could be the product of the concentration of the initial compound (benzene) and the concentration of the final compound (cyclohexane). Plotting this variable against time or degree of conversion would make it possible to draw the diagram needed. This in turn could lead to an outstanding conclusion such as the following: "The dependence of the product of benzene and cyclohexane concentrations on time is rather remarkable. At the beginning, this parameter increases until it reaches the maximum and then it starts to decrease. A similar extreme was observed for its dependence on the degree of conversion."

The crown of the discussion is to note that the results obtained are different from those of other authors. For example, "In contrast to Chuchata, who has pointed out that hydrogenation of benzene is a rather slow reaction, we have found that this reaction takes place rapidly". Of course, there will be no mention that Chuchata studied hydrogenation in the liquid phase at 20 °C using a platinum-based catalyst, while this study concerns the same reaction in the vapour phase at 150 °C using nickel as catalyst.

In discussion, great care must be taken to avoid any statement which could be used against the author later on; basically, the author should avoid expressing his own opinion. The safest way to do that is to compile opinions expressed in papers of renowned authors. Selected texts could be pasted together with a little modification of assertive into tentative sentences. The author need not be afraid that individual parts would be recognised as a fraud, particularly if the article is not written in English. The danger that plagiarism will be discovered is minimal since the discussion is frequently copied from foreign language sources, and language proficiency of the beginners is limited.

Order of authors

In practice, a beginner scientist hardly ever has a say in the ordering of authors. Nevertheless, we will discuss this problem just in case such opportunity should present itself. To get it all right, it is necessary to explain the symbolic meaning and the importance of name ordering.

Almost every young scientist has the privilege to start his career under the supervision of an older and experienced scientist, who provides guidance and interesting and plausible research topics. This father figure deserves symbolic acknowledgement; therefore his name should appear first and the young scientist's second. Naturally, the young scientist is not happy with this situation and wants to move his name up. In turn, the older scientist does not intend to provide guidance to the young co-worker for life. As time goes on, he will realise that he can no longer cope with the harsh reality of experiments, and that his invented tasks are entirely unfeasible. He also feels that he can no longer dignify an increasing amount of tiring questions by the young scientist. He therefore decides to hand the guidance over to the nearest subaltern. The main task of the subaltern is to design more plausible studies and to guide the younger scientist in such a way that makes the senior scientist believe that this is all of his own making. This situation usually reflects on the author order; the pseudo head (i.e. the subaltern) occupies the first place while the senior scientist takes the last.

All this clearly shows that only an inexperienced young scientist would aspire to the first place in the list of authors. Since authors need not to prove their physical existence by ID, it would be easy to add two fictitious names in front. It would be particularly impressive if these names sounded Arabic, Indian, or otherwise exotic, implying that the lab is often sought after by foreigners who seek valuable experience.

ACKNOWLEDGMENT

In addition to the know-how acquired when writing my own papers, I availed myself of many other authors to whom I owe thanks and an apology for not quoting them in this paper.

VÝZKUM TEORIE VĚDECKÉHO VÝZKUMU I TEORIE SDÍLENÍ INFORMACÍ PROSTŘEDNICTVÍM VĚDECKÝCH PUBLIKACÍ I ELEMENTÁRNÍ ZÁKONITOSTI OPTIMALISACE EFEKTU VĚDECKÝCH SDĚLENÍ

KAO LIN, SHIU SISAKO a JOSEF HORÁK
Vysoká škola chemicko-technologická, Praha

Došlo dne 1. 4. 1969

Při současném rozvoji metodiky zpracování vědeckých poznatků a integrací světové vědy nabývá na důležitosti výzkum teorie sdílení informací o vědeckých poznatcích. Přesto je v chemické literatuře věnováno málo pozornosti teorii sestavování vědeckých publikací, které jsou hlavní formou sdílení výsledků. Cílem této práce je podat začínajícím vědeckým pracovníkům informace o základech teorie psaní publikací a možnostech jejich aplikace při vlastní vědecké práci.

Název práce

Název je jednou z nejdůležitějších částí práce. U začínajících vědeckých pracovníků může někdy vzniknout chybný dojem, že hlavním úkolem práce je stručně a výstižně informovat o obsahu a cíli práce. To není pravda. Hlavním úkolem názvu je vyvolat v čtenáři dojem, že jde o mimořádně významné dílo moderní koncepce vypracované vzdělaným a váženým autorem na ústavu světového formátu. Proto v názvu používáme co nejvíce slov, jež jsou módní. Takovými slovy jsou např. optimální, model, analogie, výpočetní technika, počítač, struktura, relace, kvantový, mechanismus, systém atd. včetně všech odvozenin a kombinací. Seznam slov obohatíme prolistováním několika posledních sešitů Chemical Abstracts. Postup sestavování názvu ilustruje příklad.

Přítel z jednoho závodu nás požádá, abychom za nevelký obnos v tuzemské valutě zjistili, zda určitá vodička na vlasy neuvolňuje kyseliny působením roztoku hydroxidu při teplotě místnosti. S pečlivostí odpovídající zmíněnému obnosu provedeme několik pokusů. Když jsme hotovi, volá nám přítel a omlouvá se, že nám omylem místo vzorku poslal lahvičku ethylacetátu. Rozhodneme se výsledky zužitkovat alespoň opublikováním. Nezkušený pracovník by práci nadepsal: „Studie hydrolysy ethylacetátu roztokem hydroxidu sodného". Tento název nevzbudí ovšem respekt ani u posluchače druhého ročníku, natož u ostřílených vědeckých pracovníků. Zkušený vědec postupuje proto např. takto. Uvědomí si, že v živých organismech je spousta všelijakých sloučenin, že tam tedy mohou být i estery, dále, že živé organismy existují při teplotách od 0° do 50 °C a že je v nich převážně voda. Práci pak nazve: „Studie mechanismů metabolických reakcí biochemických modelů za podmínek analogických podmínkám fysiologickým I. Solvolysy látek s esterovou funkční skupinou I. Hydrolysa ethylacetátu za přítomnosti hydroxidu sodného v oblasti vyšších koncentrací."

Rozdíl v působnosti prvého a druhého názvu není třeba diskutovat. To, že práci uvedeme jako součást velkého seriálu nás nikterak nezavazuje publikovat další sdělení. Je pochopitelné, že nás vždy okolnosti mohou přinutit přejít na jiný, ještě významnější úkol. Aktivnímu začínajícímu pracovníku ponecháváme v úvaze, zda se efekt názvu nezvýší tím, že římské jednotky nahradí-

The facsimile of the title page published in Chemické Listy (1969)