Procedures for Detecting Errors in Chemical Literature*

AARON ADDELSTON and URIEL J. GOLDSMITH
Winthrop Laboratories, Special Chemicals Department,
90 Park Avenue, New York, New York 10016

Received May 16, 1966

Errors in fundamental journals and in secondary publications are corrected only in small part in lists of errata. Careless editing by editors and authors permits most of these errors. Frequently encountered potential sources of error are listed.

For this Symposium on Error Control in the Chemical Literature, I propose as our motto Messala's curse from Shakespeare's Julius Caesar (*1*):

> "O hateful Error, Melancholy's child!
> Why dost thou show to the apt thoughts of men
> The things that are not? O Error, soon conceived,
> Thou never com'st unto a happy birth,
> But kill'st the mother that engendered thee."

It is my unpleasant task to charge that our chemical literature is in danger of being encumbered, if not flooded with error, and to plead with authors, editors, referees, and proofreaders to halt the tide. "To err is human, to forgive divine;" (*2*) but we are scientists and not theologians. We must strive for accuracy. We cannot afford to have our reading progress delayed by ambiguities, contradictory structures and names, or indecipherable mathematical equations. We just cannot afford the time required to analyze what the author meant to write.

I humbly admit that I am not fit to cast the first stone. A catalog of chemical products (*3*) which I prepared and edited last year contains three incorrect names, two incorrect structures, and two incorrect melting points.

American Chemical Society and other fundamental journals, textbooks, chemical encyclopedias and dictionaries, and trade papers all print numerous errors, only a small portion of which are subsequently "corrected" in lists of errata. The extent of these errata is indicated by the following summaries:

J. Am. Chem. Soc. Dec. 1964, 44 errata, some going back to 1950.
J. Am. Chem. Soc., Dec. 1965, 27 errata, one going back to 1933.
J. Org. Chem., Dec. 1964, 45 articles with 91 errors.
J. Org. Chem., Dec. 1965, 41 articles with 75 errors.
J. Chem. Soc. (London), Dec. 1962, 36 errata.
J. Phys. Chem., Dec. 1963, 13 errata.
Chem. Zentr., Dec. 1964, about 500 errata going back to 1950.

Although most journals do compile and publish errata, at least two important publications, *Science* and the *Journal of the American Medical Association*, do not gather these corrections in one place, but scatter them in odd spots throughout the year wherever dummying provides room. Consequently, the editors of *Chemical Abstracts* may easily overlook these because they properly argue that they cannot afford a "systematic effort to peruse each page of each issue of the over 10,000 journals currently monitored" for buried errata.

Unlike *Chemisches Zentralblatt*, *Chemical Abstracts* does not publish its own errata in one place. A systematic effort is made to pick up such corrections before the article is abstracted. According to James L. Wood, head librarian of Chemical Abstracts Service (*4*):

> "During the indexing of abstracts, the indexers note errors in concepts or compounds, structures, formulas, and the correct information is entered into the *CA* index with a note, 'Incorrectly so-and-so in *CA* (Original).' Corrections for errors called to our attention after an abstract has been printed are handled in *CA* indexes by an entry reading, 'Reference should be to so-and-so.' We have no mechanism for additions and corrections which are in the originals but that are not indexed."

An analysis of the types of errors reported in the last issue of 1964 and 1965 of *Journal of the American Chemical Society* and *Journal of Organic Chemistry* reveals that in each volume there was at least one article which appeared with a whole page of the author's manuscript missing. In one case this involved a long table of data, in another a series of 24 structural diagrams, and in a third, a complicated flow sheet. The most frequent errors were in the following categories:

Structure diagrams	64
Roman letters	15
Chemical names	20
Terms in mathematical equations	26
Numerical data	49
(counting as only one error a long table of incorrectly calculated dipole moments)	
Empirical formulas	7

Other reported errors include: misidentification of α and β and *cis-trans* isomers, wrong Greek letter, yields, quantities of reactants, identities of reactants, reaction temperatures, experimental procedures, biological species, misplaced or omitted footnote superscripts, wrong reference citations, omitted references, incorrect spelling of authors' names, incorrect spelling of nonchemical words, incorrect acknowledgment, and even one correction in date of receipt

*Presented before the Division of Chemical Literature, Symposium on Error Control in the Chemical Literature, 151st National Meeting of the American Chemical Society, Pittsburgh, Pa., March 23, 1966.

of manuscript. To my surprise, there were no instances of omission of the word "not," but there were five analogous reversals of meaning: an opposite direction of an arrow in a flow sheet, omission of the angular sign for "less than," "proton donor" for "proton acceptor," "stable" for "unstable," and "directly proportional" for "inversely proportional." There were nine reinterpretations of conclusions, based on new theories or new data.

In an effort to determine the origin of these errors, a letter was sent to the author of each erratum in *Journal of Organic Chemistry* for December 1964. According to some 34 replies, the original error was made in:

The holograph (handwritten manuscript or drawing)	10
Typed copy	9
Typesetting the copy	13

In one instance the author reported that a correction made in a proof sheet was not made by the printer; in another, the author said he never received a proof for correction; in one case, the author was not aware of a contemporary publication with conflicting data until after publication.

The error was discovered by:

The authors	18 times
Journal editors	2 times
Journal readers	7 times
Abstractors	1 time
Colleagues	1 time
Index Chemicus	1 time

Presumably referees, editors, and proofreaders caught many other errors before publication, but no data are available.

In an effort to determine how seriously librarians consider these errata, a letter was sent to the libraries of 14 prominent pharmaceutical manufacturers asking how these errata are handled. All answered. Of the 14, six *never* check the errata, four check them in some journals, and four claim to check them in all. Most of those who took action at all preferred to paste the erratum on the original journal page or to enter the correction directly in ink. Only two made a cross reference on the original journal page without actually entering the correction. The problems of these librarians may be summed up in contrasting replies from two pharmaceutical companies of top rank as judged by caliber of scientific staffs and contributions to medical progress.

One wrote: "We have no time for this important job. We made corrections only if they are specifically called to our attention." The other wrote: "I feel strongly about errata control, especially in the medical field where it can be life-saving... I think publishers have an obligation to record them in uniform locations and at least to record them in tables of contents... We don't make a page by page search as is needed in many cases, but we do look in the places where errata are regularly found, and no matter whether current or overlooked in an old volume, we take care of it."

I sometimes feel that authors, referees, and editors are not as conscientious as they should be about avoiding (or even correcting) errors.

In the January 5, 1966 issue of the *Journal of the American Chemical Society*, there appears an article (*5*) so replete with editorial and typographical errors that I characterized it in a letter to the editor as "a disgrace to the Journal's standards." This article gives elaborate mathematical treatment to presumably precise kinetic and isotope ratio data, but the starting materials were used without analysis or purification as received from commercial sources, and there are no analytical data for the synthetic products. Among other errors noted and submitted were:

1. Two incorrect structural diagrams (a misplaced NH_2 group) and a suddenly saturated benzene ring in a flow sheet where the author used no reducing agent.
2. An incorrect chemical name, using the misleading prefix "thiol," implying a free SH group, instead of "thio."
3. Omission of preparative directions for this compound (an obscure, and possibly new compound), where elaborate directions are given for every other compound used, many of which are well known.
4. A calculation of number of moles of reagent which was off by a factor of 10.
5. Multiple confusion of what was printed as l-cystine with *l*-cystine or L-cystine.
6. Directions to add a reactant portionwise to the liquid "as fast as the *metal* reacted" in a procedure where no metal is mentioned.
7. Boldface paragraph titles giving ordinary names of amino acid derivatives when the material made was actually an isotope-enriched product. This might have been forgiven until there appeared a boldface paragraph title for S-Benzyl-cysteine-N^{15} where inspection of the detailed procedure failed to show that any of the reactants had any N^{15}.
8. Many other minor errors of usage such as "ammonia chloride" and "until it was just alkaline *of* phenolphthalein."

In my letter to the editor, I pointed out that the existence of so many errors made me doubtful of the author's elaborate mathematical treatment and kinetic hypotheses, which I was incompetent to check.

The following is the full text of the reply from the editor:

"Thank you for your letter of January 20. We can only say that the manuscript was refereed by two thoroughly competent people who are in fact leaders in the general field of the mechanism of hydrolysis and aminolysis of esters. Apparently their examinations were primarily limited to matters of scientific content and they did not pay much attention to the less important details. I am passing along a copy of your letter to our Production Office in the hope that they may profit from your keen observations and perhaps improve proofreading in the future."

In my opinion, with all due respect, the editor has missed the point. This multiplicity of errors not only wastes the reader's time but taxes his credulity.

A careful reading of the published errata does little to increase one's faith in the printed word or datum. Many errors actually appear in the errata!

In December 1963, McBee *et al.* (*6*) suggested that in their article earlier that year (*7*), in the abstract, "octachlorocyclopentane" should read "octachloropentene."

Thus they restored the missing double bond, and decyclized the compound. In December 1964 (*8*), they had to point out that they meant "octachlorocyclopentene" all along.

In December 1964, Brace (*9*) recommended that for his compound named 1-trifluoromethyl-4-iodo-8-pentylnonane (*10*), the published empirical formula, $C_{11}H_{28}F_3I$, should read $C_{25}H_{22}F_3I$. Both empirical formulas in this so-called correction are wrong! The originally published formula was $C_{17}H_{28}F_3I$ and it should have been corrected to $C_{15}H_{28}F_3I$. This is a halogen-substituted alkane. The formula can be checked easily by replacing the three fluorines and one iodine with four hydrogens to get the expected $C_{15}H_{32}$ (C_nH_{2n+2}).

Curious to understand how these errors occurred, I decided to check the author's line formula which was

$$CF_3CH_2CH(C_5H_{11})CH_2CHI(CH_2)_4CH_3$$

This corresponded with the empirical formula $C_{15}H_{28}F_3I$, but not to the name. Using this unusual system of nomenclature, which seeks to preserve the identity of the trifluoromethyl group, and characterizes the compound as a nonane derivative, the name should be: 1-trifluoromethyl-4-iodo-2-pentylnonane. The Geneva name should be: 6-(2,2,2-trifluoroethyl)-8-iodotridecane (Figure 1).

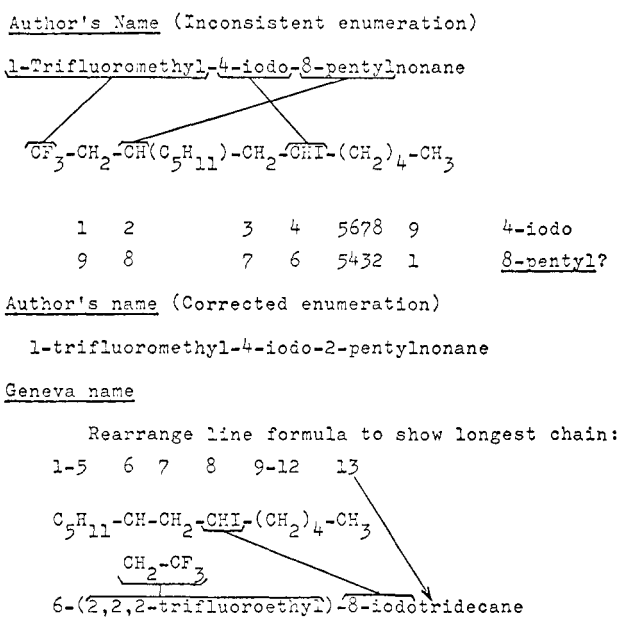

Figure 1. How an incorrect compound name is generated.

In view of this discrepancy, we decided to determine how *CA* handled the problem. The abstract itself [*CA* **59**: 15185 (1963)] does not mention the compound. The *CA* Formula Index for 1963 has the correct empirical formula, $C_{15}H_{28}F_3I$, but gives the name as 1,1,1-trifluoro-5-iodo-9-methyltetradecane (Figure 2). This turns out to be the Geneva name corresponding to the *incorrect* 1-trifluoromethyl-4-iodo-8-pentylnonane. Thus we have a classic example of how two typographical errors (C_{17} for C_{15}) and 8-pentyl for 2-pentyl have been compounded and complicated.

Even eminent texts have errors. Some time ago I sent

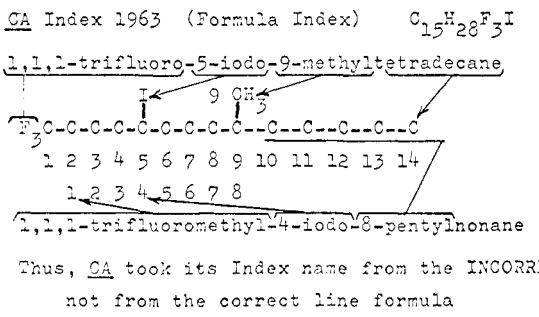

Figure 2. How an incorrect index name is generated.

a list of some 22 errors or questions in the first 156 pages of Fieser's "Topics in Organic Chemistry" (*11*). Dr. Fieser was gracious enough to correct some of my own misinterpretations, but admitted some 10 typographical errors, and promised to give further study to another few which he could not adjudicate off hand. Dr. Fieser makes a genuine effort to avoid such errors. He writes: "I have had several expert chemists read our manuscripts before typesetting and have engaged professionals to read galley proof, but errors inevitably come through. When a new book is prescribed for use in one of our courses, I offer a prize of $1.00 for each error discovered in order that the first reprinting can be corrected as fully as possible."

Chemical Biological Activities recently sent its subscribers a set of corrections to issues 1 through 13. Unfortunately in at least eight cases, the correction is wrong. Either the old error is repeated, or a new one is introduced. In one case the correction was supposed to delete a nonexistent "HCl," but the new diagram omits the more significant 3-chloro ring substituent.

How does one detect errors in the manuscript or the proof?

Rule 1. Trust nothing!

Rule 2. Wherever a structural diagram appears, check it against the name, against the empirical formula, against the sense of the flow sheet. Write out a chemical code from the diagrammed structure and check code against the name.

Rule 3. Wherever a Roman number appears, link it with both the structure and name, and see if they agree. My advice to authors and editors alike is to abandon the Roman numbers because our unfamiliarity with them, except in the pages of *Journal of the American Chemical Society* and *Journal of Organic Chemistry*, introduces a new source for error, and to use boldface Arabic numbers instead. It is easier to make a mistake in XXIX*b* than in 29*b*. Such compound numbers should always be repeated in the experimental section.

Rule 4. In tables, check column headings against the text and footnotes for sense; where graphs appear, check units for abscissas, ordinates, and names of curves for agreement with table headings.

Rule 5. Worry about minus signs. If one appears in an equation or table of data, maybe there should be more.

Rule 6. If, in a table of data, any number differs by a factor of 10 from its companions, don't trust it unless the author makes a point of it.

Rule 7. Watch out for the following frequent errors:

a. O for C and vice versa in the first member of a side chain.

b. Cyclopentyl or cyclohexyl rings with just a vertical double bond implying a methylene group where C=O is meant.

c. Careless usage of the solid, dotted, and wavy lines to show known or unknown stereo bonds. Once a paper uses a dotted or wavy line, one would expect that the solid lines mean something, but this is not always the case.

d. Beware of the 5-position in steroids. The hydrogen at this position can be either α(dotted) or β(solid). Authors often fail to indicate which it is, even where this can be determined from flow sheet or context.

e. Trade and nonproprietary trivial names often imply more than they say. Where an author says Demerol he means meperidine hydrochloride. Biochemical literature is loaded with articles referring to epinephrine where epinephrine bitartrate is meant. Recent issues of *CBAC* have made a false distinction between 2-isopropylamino-1-(2-naphthyl)-ethanol, Registry No. 54808, and its hydrochloride, No. 51025, under the false assumption that the trade name Nethalide refers to the base. It is the hydrochloride.

f. It is easy to confuse pyrrolidino with pyrrolidono.

g. In phenothiazines, the U. S. 3-substituent would often be called a 2-substituent in European literature. This confusing designation will be carried over to *CA* if the name appears in the title, but may be corrected in the *CA* Index. It will not be corrected in *CBAC*.

h. Italic lower case *p* is often confused with Greek rho, and rho with sigma.

i. In dosage statements, check milligrams *vs.* micrograms.

This list could be expanded almost indefinitely. At the risk of appearing to be overly cynical, I repeat Rule 1. Trust nothing.

If I have seemed unduly harsh in my criticism of current practice, I revert to humanistic literature and cite Wordsworth's "Ode to Duty (*11*).

"O, Duty. If that name thou love
. . . check the erring and reprove."

To which there should, of course, be an erratum. Dear Mr. Wordsworth:

"Love does not rhyme with reprove."

LITERATURE CITED

(1) Shakespeare, Wm., "Julius Caesar," Act 5, Scene 3, line 67.
(2) Pope, Alexander, "Essay on Criticism," line 525.
(3) Descriptive Price List, Winthrop Laboratories, Special Chemicals Dept., April, 1965.
(4) Private communication.
(5) Wagner, E. S., R. E. Davis, *J. Am. Chem. Soc.*, **88**, 7 (1966).
(6) McBee, E. T., W. L. Dilling, H. P. Brandlin, *J. Org. Chem.* **28**, 3595 (1963).
(7) *Ibid.*, p. 2255.
(8) *Ibid.*, **29**, 3744 (1964).
(9) Brace, N. O., *J. Org. Chem.*, **29**, 3744 (1964).
(10) *Ibid.*, **28**, 3102 (1964).
(11) Fieser, L. F., Fieser, M., "Topics in Organic Chemistry," Reinhold Publishing Corp., New York, 1950.
(12) Wordsworth, William, "Ode to Duty."

Procedures for Assessing Errors*

KENNETH L. COE, DON P. LEITER, Jr., HARRY L. MORGAN, and FERD R. WETSEL
The Chemical Abstracts Service, The Ohio State University, Columbus, Ohio 43210

Received May 26, 1966

User confidence in the validity of published chemical literature is directly related to the accuracy with which it is presented. The nature, types, and sources of errors which find their way into the primary and secondary journal literature are analyzed. A description of the various methods by which the effects of such errors can be minimized through abstracting and indexing techniques is presented.

The Chemical Abstracts Service (CAS) recognizes that many errors are introduced into primary and secondary journals in the publication processes. Rather than dwell on the present status, which will be reviewed briefly, we will discuss CAS plans for minimizing these errors through various abstracting, indexing, and computer-handling techniques.

At CAS, abstracts are prepared, edited, and indexed by subject and language experts. This processing locates and resolves many of the detectable errors or ambiguous statements which occur in the primary literature, and/or which occur in the abstracting and indexing operations. However, all such errors are not found even in these rigorous review and edit stages.

Since errors in any part of a paper or in the prepared abstracts—title, bibliographic reference, author names, body of the abstract, body of the paper, or even in the

* Presented before the Division of Chemical Literature, Symposium on Error Control in the Chemical Literature, 151st National Meeting of the American Chemical Society, Pittsburgh, Pa., March 23, 1966.