Effective Bio*molecular* Reading & Writing

Valerie Matarese

2017

Matarese Valerie Ann

First printing

Effective Biomolecular Reading & Writing

ISBN (print) 979-12-200-1879-1

Published by Matarese Valerie Ann, Via Roma 10, I-31020 Vidor (TV), Italy.

Printed and bound in Italy.

Table of Contents

Tables

Figures

Boxes

Preface

This handbook is intended as a support text for "Effective Biomedical Writing," an intensive practical course designed for doctoral students and early career researchers. It is written for young scholars who use English as an additional language, especially those unfamiliar with the speech of a native English speaker. Accordingly it presents the content of the lectures, but not all the examples and exercises used in class. The language used is American English but the text addresses issues of writing in British English as well, with explanations of differences in terminology and usage between the two languages.

Course participants are recommended to read the relevant lecture notes both *before* each lesson (to familiarize themselves with the vocabulary) and *afterward* (to reinforce the concepts). They are also asked to bring this handbook to every lesson. The handbook will also serve as a reference for course participants after the course is over. For this reason, it is organized in two parts, the first dealing with the issues and ethics of research publication and the second describing methods for writing scientific articles. This is not the same organization as that of the class work, as each day of the course addresses topics from both parts of the handbook. The accompanying program indicates which topics are dealt with on each day.

This volume is a work in progress. It has developed from several editions of the course given in different formats over 10 years. I gratefully acknowledge the comments and feedback of course participants, both those of prior editions and of the present, for sharing their feedback on the effectiveness of the explanations and for providing teaching examples from their writings. I am also indebted to my international colleagues who have shared their own approaches to teaching research writing through their publications, their presentations at professional meetings, and their contributions to *Supporting Research Writing*, a book I edited and published in 2013. This handbook nonetheless still needs improvement. As researchers do when drafting research articles, I will revise, and revise and revise!

Valerie Matarese
Vidor, February 2017

Part I. Publication issues and ethics

I.1 Aims and challenges of publishing research

A researcher's main unit of output is the research article. We produce other things, of course, such as review articles, letters to the editor, peer review reports, patents on new technologies, software algorithms, data sets, and teaching materials, but it is the research article that, at least today, counts the most towards our status and career. This is changing, thankfully, but for now our focus is the research article (also called "research paper").

Why do researchers publish? Many of us will say that, by publishing, we **advance knowledge** by adding new findings to the scientific knowledge base. We also leave a **permanent record** of our work, so that we can get **due credit**: we lengthen our CVs, we hope to be invited to talk about our work at meetings, our superiors take note and promote us to higher academic positions, and so on. By publishing, we also get **feedback**, as the article is critiqued in letters to the editor and, increasingly, in comments posted online. And, hopefully, the study gets **independent verification** by being repeated by others, who confirm the findings. Only once a study is replicated and validated in this way can it really advance knowledge—this is a tenet of the scientific method.

By the act of publishing, we communicate with a broad group of readers. In the narrowest sense, we communicate with researchers working in the same field as us, namely our colleagues and competitors. They form a minority of all our readers, however, and can be thought to represent the innermost of a set of concentric circles of readers (Figure I.1). Broadening out to the next circle, we communicate with researchers in other fields, some closely related but many more working in distant fields wishing to understand the broader picture of biological phenomena. Our readers are not only scientists but also persons whose work intersects with research, for example grant funders and university administrators who make decisions that affect our careers. The literature is also read by science writers, journalists, health policy-makers, patients, activists, and even the average person (perhaps to understand a disease). The outer circle of readers has different

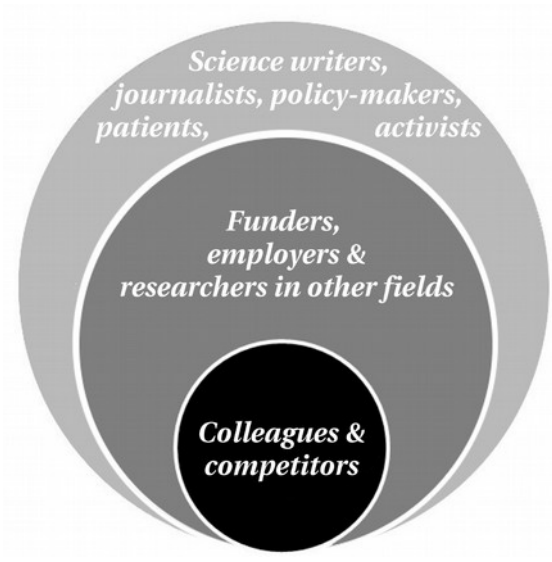

Figure I.1: Concentric circles of readers

information needs and reads research articles differently from how the inner circle reads. If we address all these readers when writing, our work will have greater impact and the visibility of our research efforts will be optimized.

So what makes a good research article? To start, good research: Our study addressed a scientific topic that was timely and important. We formulated a clear study question and designed a series of experiments able to provide a reliable answer. Our research materials were of high quality, and the individual methods were appropriate and powerful enough to provide answers. The experiments were skillfully executed, and the data were accurately collected, analyzed, and interpreted.

That's half the work to producing a good research article. **Equally important is good reporting**: The Introduction section clearly states the context of the research and justifies why the study was done, helping readers understand the study's **relevance**. The materials and methods are accurately described in enough detail for the research to be repeated and, ideally, confirmed by multiple teams of other researchers; this supports the study's **reproducibility**. The results are fully documented (in figures and tables) and clearly explained, providing the **evidence**. The conclusions are based on the reported evidence: this is our **scientific advance**. And, overall, the writing is clear (easily and quickly understood), precise (not vague or ambiguous), original (not copied nor recycled), efficient (not unnecessarily verbose), and plain (formal but not sophisticated). The text has a forward flow of information, the arguments are sound, and the prose is reader friendly: this is our **scientific voice** emerging from the page and communicating with our readers.

Many published research articles lack these features. Even though journals choose articles to publish on the basis of a positive peer review, not all journals set high-enough standards for acceptance. We may therefore find ourselves reading articles with ambiguous wording or insufficient methodological details. We may judge a study to be underpowered or have an imperfect statistical analysis. We may even repeat the reported experiments in our own lab but not get the same results, concluding that part or all of the study is irreproducible.

On rare occasions, readers (or the authors themselves) report the problems they encounter to the journal that published the questionable article, possibly resulting in that article's **retraction** (withdrawal from the literature). Surveys[1] have shown that the main reasons for retraction are fraud (data manipulation and fabrication), plagiarism (of text or data), and honest scientific error. Retracted articles make up only a small fraction, about 0.01 %, of all articles indexed in Medline, while the fraction of articles reporting irreproducible research is estimated to exceed 50 %.[2] This irreproducibility corresponds to an enormous amount of **research waste** (waste of both efforts and funding). Therefore, the problem of irreproducibility is getting much attention in the scientific community and general press.

1 See, for example, Decullier et al. (BMC Res Notes. 2013;6:238), Fang et al. (Proc Natl Acad Sci U S A. 2012;109(42):17028-33), and Wager and Williams (J Med Ethics. 2011;37(9):567-70)

2 See Freedman et al. (PLoS Biol. 2015;13(6):e1002165) and references therein

In most cases, **irreproducible research** is not the result of fraud, plagiarism, or other forms of misconduct. Rather, it is the product of poor scientific practice or poor reporting. Studies may be poorly designed: researchers may choose to use study models or methods that are inappropriate for the study question, or their controls may be inadequate and their statistical tests unsuited for the type of data. Their chemical reagents and biological samples may be contaminated, impure, degraded, or mislabeled, or they may be uncharacterized or unvalidated for the specific use in the study. The methods may not be executed to perfection, with uncontrolled environmental variables and instrumental error interfering with the collection of reliable data, and sample mix-up and spreadsheet errors corrupting the data analysis. Moreover, the investigator's objectivity in analyzing the data may be clouded due to a fervent belief in the tested hypothesis. Finally, in the published report, the methods may be described in an incomplete or inaccurate manner, and the results may be only a selection of the "best" data that fit a hypothesis but that do not tell the whole story.

Journals, grant funders, and universities are now working to help researchers resolve these problems of irreproducibility. Among the initiatives we are starting to see are: removal of word limits on the methods section of research articles, to allow full reporting; development of checklists to help authors include all the necessary information for a study's replication (and to guide peer reviewers in assessing reports for reproducibility); encouragement of the sharing of data, materials, and software; and the establishment of best practices for reporting biological reagents such as antibodies and cell lines. These changes aim to develop a more rigorous publication culture, requiring researchers to be ever more familiar with the requirements of "publishability" in terms of both **language skills** and **skills in communicating scientific-technical content**. Therefore, this course dedicates equal time to issues of writing for publication and to methods for presenting content in a research article according to current best practices.