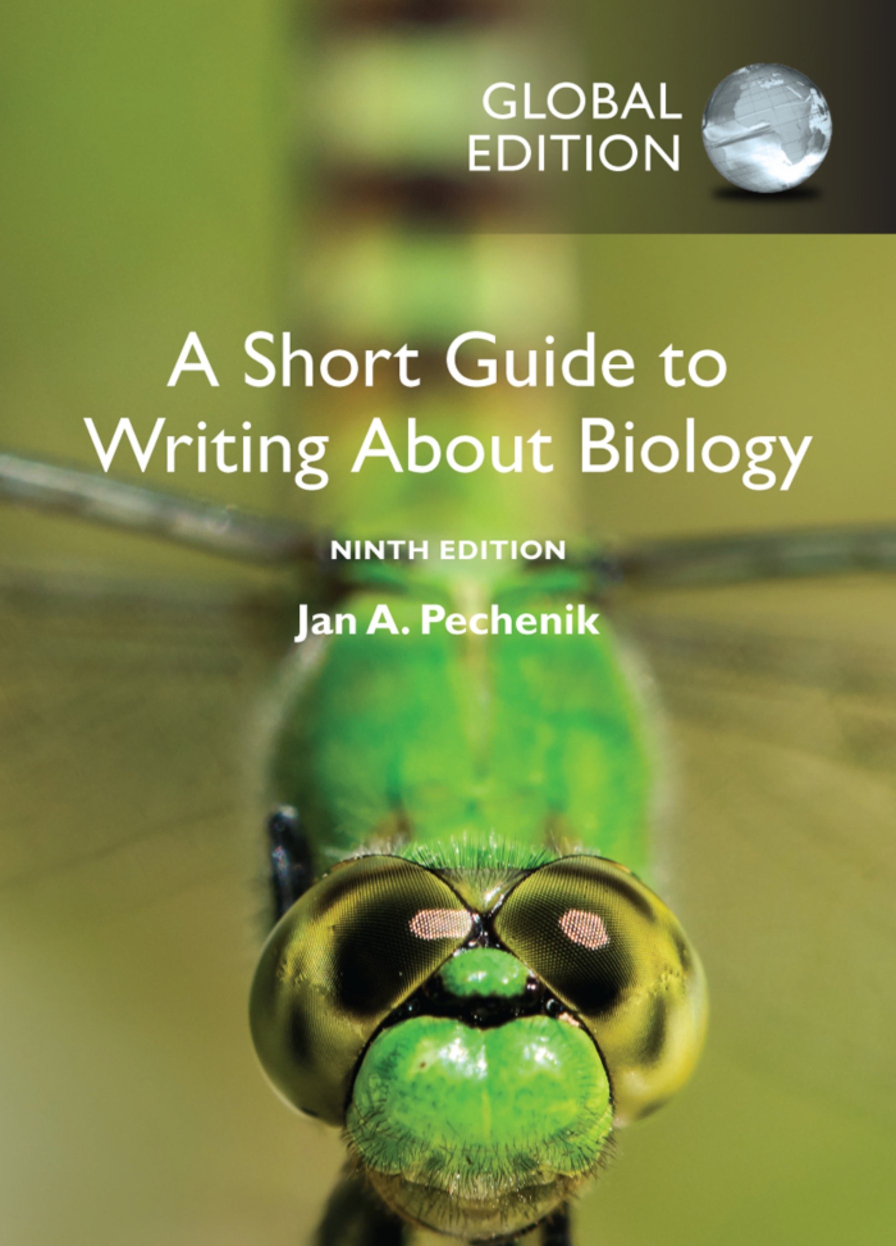

GLOBAL EDITION

A Short Guide to Writing About Biology

NINTH EDITION

Jan A. Pechenik

ALWAYS LEARNING

PEARSON

A Short Guide to Writing about Biology

NINTH EDITION

GLOBAL EDITION

JAN A. PECHENIK
Biology Department
Tufts University

PEARSON

Boston Columbus Indianapolis New York San Francisco
Amsterdam Cape Town Dubai London Madrid Milan Munich
Paris Montréal Toronto Delhi Mexico City São Paulo Sydney
Hong Kong Seoul Singapore Taipei Tokyo

Individual studies often lack significant power to detect a meaningful difference, even when one exists, because of necessarily small sample sizes; in consequence, many researchers are now doing **meta-analyses**, in which results of many individual studies, usually done by many different researchers, are combined into a single study. The results are still examined by hypothesis testing.

Sometimes you will run across a very different approach to evaluating data, one that does not involve hypothesis testing. Instead, the experimenters compare their data with those predicted by selected mathematical models and then evaluate how well the data do or do not fit the models. This general approach is referred to as **Bayesian**, and a typical way of evaluating the data is called **Maximum Likelihood**. Again, you don't need to understand the details of how the calculations were made in order to understand the point of the study: The data either do or do not conform to the predictions of the model.

READING ABOUT STATISTICS

When you read the results of statistical analyses in a published research paper, you don't need to understand how the particular calculations were made in order to understand what the results mean. **A small *p*-value means that if the null hypothesis were true, we would find a difference as large as the one found in that study very rarely.** For example, $p < 0.001$ means that such a large difference would be expected by chance fewer than 1 time in 1,000 repeats of the study if H_0 were true. Thus, the null hypothesis is probably wrong. **In contrast**, **a large *p*-value means that the researchers obtained a difference of about the size expected if the null hypothesis were true.** For example, $p = 0.28$ means that the results would turn up at least 28 times out of 100 repeats of the experiment if H_0 were true. Thus, the researchers can't reject H_0 with confidence, especially if the statistical power of the test is low; instead, they must retain it for the time being as part of a work in progress.

WRITING ABOUT STATISTICS

If you have analyzed your data using appropriate statistical procedures, the products of your heavy labor are readily and unceremoniously incorporated into the Results section of your report to support any major trends that you see in your data, as in the following 3 examples. **Note that in Example 1, the author reports the sample size ($n = 30$ caterpillars), the test statistic and its value, the number of degrees of freedom associated with the test, and the size of the associated *p*-value.** Note also that the writer focuses on the biology being studied—that is, on the biological question being addressed—rather than on the statistics themselves.

EXAMPLE 1

> The 30 caterpillars reared on the mustard-flavored diet and subsequently given a choice of foods showed a statistically significant preference for the mustard diet ($\chi^2 = 17.3$; *d.f.* = 1; $p < 0.05$). For the 30 caterpillars reared on the quinine-flavored diet, however, there was no significant influence of previous experience on food choice ($\chi^2 = 0.12$; *d.f.* = 1; $p > 0.10$).

In this example, H_0 states that prior experience does not influence the subsequent choice of food by caterpillars; $p < 0.05$ means that if we were to conduct the same experiment 100 times and H_0 were true, such a high value for χ^2 would be expected to occur in fewer than 5 of those 100 studies. In other words, the probability of making the mistake of rejecting H_0 when it is, in fact, true is less than 5%. You can therefore feel reasonably safe in rejecting H_0 in favor of the alternative: that prior experience did influence subsequent food selection by caterpillars reared on the mustard diet.

Different results were obtained, however, for the caterpillars reared on the quinine-flavored diet; $p > 0.10$ means that if the experiment were repeated 100 times and H_0 were true, you would expect to calculate such a small value of χ^2 in at least 10 of the 100 trials. In other words, the probability of getting this χ^2 value with H_0 being true is rather high; certainly, the χ^2 value that was calculated is not unusual enough for you to mistrust H_0 and run the risk of rejecting the null hypothesis when it might, in fact, be true. **This does not mean that H_0 is true, only that we do not have enough evidence to reject it.** The quinine-flavored diet may have altered dietary preference—the variability in response may simply have been too great for us to perceive the effect with the small number of caterpillars used in our study. It is also possible that diet affected food choices in ways we did not test for. Thus, the wording in Example 1 was carefully chosen to say no more than is safe to say.

EXAMPLE 2

> Over the first 10 days of observation, seedlings receiving the nutrient supplement did not grow significantly faster than those receiving only water (2-sample *t*-test, $t = 1.12$; *d.f.* = 47; $p = 0.29$).

In this second example, the writer has chosen to report the actual *p*-value rather than writing $p > 0.10$. The null hypothesis (H_0) states that the nutrient supplement does not influence plant growth; $p = 0.29$ means that if the experiment were repeated 100 times and H_0 were true, you would expect to calculate such a low value of *t* in 29 of the 100 trials. As before, you have obtained a value of *t* that would be fairly common if H_0 were true, so you have no convincing reason to reject H_0. It is, of course, possible that H_0 is actually false and the nutrients really do promote seedling growth, and that you just happened on an unusual set of samples that gave a misleadingly small t-value. If such is the case, repeating the experiment should produce different results and larger t-values. But with only the data before you, you cannot reject H_0 with confidence.

EXAMPLE 3

> The hermit crabs in our sample ($N = 12$) showed a significant relationship between their wet weights and the size of the shells they occupied in the field ($r^2 = 0.477$; test for zero slope: $F = 9.104$; $d.f. = 1, 12$; $p = 0.013$).

In this case, hermit crab weight and shell size were correlated (the slope of the line relating the 2 variables differs significantly from zero), even though variation in hermit crab weight accounted for only 47.7% of the variation in the sizes of the shells occupied (as shown by the r^2 term).

Note carefully how the word *significant* was used in the preceding examples, and that the writers put the focus clearly on the biology; they said little about the statistics themselves. Statistics are used only to support any claims you wish to make about your results, as in examples 1–3; resist the temptation to ramble on about how the statistics were calculated, how brilliant you are to have figured out which calculations to make, or how awful it was to make those calculations.

Compare the first 3 examples with the following 2 examples.

EXAMPLE 4

Wrong:

> A chi-square value of 6.25 with a p of 0.0124 revealed that the physical condition of the shell had a significant influence on shell choice by the 15 hermit crabs used in our study.

Corrected:

> The physical condition of the shell had a significant influence on shell choice by the hermit crabs ($\chi^2 = 6.25$; $d.f. = 1$; $N = 15$ hermit crabs, $p = 0.0124$).

EXAMPLE 5

Wrong:

> A Student's 2-sample t-test was used to determine the significance of the difference in mean interaction times. The data were not significant (i.e., we found no significant results in our experiment), and there was no difference in the mean contact time between hermit crabs in the presence or absence of predators.

Analysis of Example 5:

The first sentence, of course, should not be there at all. With the second sentence, the student has fallen into the common trap of confusing the results of a significance test with the value of the data, or of the study itself ("The data were not significant … "). There is nothing wrong with the data. **Don't apologize for results that support H_0. Failure to discredit H_0 does not mean that your experiment was a failure.** If your sample size was small and the amount of natural variability was high, then it is hard to discredit a null hypothesis even when H_0 is wrong. And, of course, H_0 might actually be correct.

The student in this example then uses the results of a significance test to make a definitive pronouncement ("there was no difference in…"), by omitting the word *significant* from the sentence. As discussed previously, statistics do not prove things; they only indicate degrees of likelihood. "No significant difference" is not the same as "No difference." Finally, the student provides no statistical support for the final statement made.

Corrected:

The presence of a predator did not significantly affect the amount of time that hermit crabs spent interacting with each other (2-sample *t*-test, $t = 1.012$; *d.f.* $= 12$; $p = 0.332$).

For more detailed advice about discussing "negative" results (i.e., those that fail to reject H_0) see Chapter 8, on writing a Discussion section (p. 188).

As mentioned previously (pp. 80–82), it is becoming common to report not just *p*-values but also statistical power and effect sizes to help interpret the reported *p*-values. Examples are shown on pages 80–81. Ask your instructor whether he or she expects you to include such information in your reports.

SUMMARY

1. Individuals typically differ in size, physiology, and behavior, making it difficult to definitively characterize the average (mean) trait without large sample sizes.
2. Variation in traits is typically characterized by variance, standard deviation (SD), or standard error of the mean (SEM)—the larger the numbers, the greater the amount of variation among measurements—and by confidence intervals (CIs).
3. Null hypotheses (H_0) assume that treatments have no effect, that there are no differences between groups, and that there is no correlation between the traits of interest; the null hypothesis is rejected only when statistical testing makes it seem very unlikely.

4. Rejecting the null hypothesis does not prove that it is wrong; accepting the null hypothesis does not prove that it is correct.
5. When incorporating the results of statistical tests into reports, emphasize the biological result, not the statistics; that is, use statistics in the same way that you use references to support statements of fact or opinion:

    The nutrient supplement had no significant effect on the rate of tomato seedling growth (2-sample *t*-test, $t = 1.06$, *d.f.* $= 47$, $p > 0.10$).

    The nutrient supplement caused significantly more rapid growth of tomato seedlings (2-sample *t*-test, $t = 2.75$, *d.f.* $= 47$, $p < 0.01$).

    The nutrient supplement caused the tomato seedlings to grow nearly 12% faster than the controls (2-sample *t*-test, $t = 2.75$, *d.f.* $= 47$, $p = 0.009$).

6

REVISING

Something that looks like a bad sentence can be the germ of a good one.
LUDWIG WITTGENSTEIN

What a very difficult thing it is to write correctly.
CHARLES DARWIN

No one will ever criticize you for having written too clearly.
NAJ A. KÍNEHCÉP

Easy reading is damned hard writing.
NATHANIEL HAWTHORNE

What Lies Ahead? In This Chapter, You Will Learn

- The importance of completing a first draft early enough to allow at least several days for revision
- How to revise for different sorts of problems in multiple passes through the manuscript
- To always begin by revising for content and for the logical flow of information and ideas
- How to give constructive criticism on another student's draft and how to interpret criticism of your own drafts

Much of this book concerns the reading, note taking, thinking, synthesizing, and organizing that permit you to capture your thoughts in a first draft, from which they can't escape. This chapter concerns the revising that must follow: Now, you must examine the first draft critically and diagnose and treat the patient as necessary.

You can't do a thorough job of revising in a single pass. Once you fix the major problems (often relating to the organization of ideas), a whole new set of problems bubbles up to the surface. They then become the next round of major problems that need attention, and so on.

This chapter presents revision as a multistep process. Draft by draft, the product gets better. I typically revise my own writing at least 4 or 5 times before letting anyone else see it, and several more times after it has been reviewed by others.