- Most readers will not try to reproduce your work; therefore, it is not necessary (or wise) to provide the level of detail necessary for someone to repeat your experiments exactly.
- A detail should be included in Methods if it (a) establishes your qualifications as a researcher; (b) establishes plausibility of your approach to the problem; or (c) helps establish the investigative steps so a reader can understand your solution of the problem.
- You may have to rewrite Methods text from paper to paper, because you can't repeat your own published wording without copyright issues.

Exercises

1. Take an experiment (or observation) that you've planned or recently completed, and write two versions of Methods text to describe it. In the first version, include enough detail that another scientist could reproduce your work exactly. In the second, include only enough for readers to understand your work and find it credible. Which version seems closer to the Methods sections you see in the literature for your field?

TWELVE

The Results Section

Writing the Results seems as though it should be easy. Its contents certainly seem obvious: this is where you report the results of your experiments (or observations, or theoretical work). True enough; but this isn't a very useful way of thinking, because it emphasizes the writer's experience, not the reader's. Writers who take this perspective produce manuscripts bulging with preliminary results, multipage tables, dozens of complicated graphs, and every data point from every experiment they ever ran. The rare reader who sticks it out is left torn between bewilderment and resentment—and this isn't where you want *your* reader to be. So remember the importance of finding and telling your story (chapter 7), and strip your Results down. Show only the data the reader needs to understand and accept the answer you're presenting to your central research question. What remains will be a short and simple section in which every word, graphic, and data point contributes directly and obviously to telling your paper's story.

With content set, what remains is presentation. The Results section must tread carefully in its relationship to the Methods that precede it and the Discussion that follows. It must be organized so its most important content is easily apparent to the reader. Finally, it needs to communicate complex and heavily quantitative information while remaining easy to read.

Methods, Results, and Discussion

The strict separation of Methods, Result, and Discussion can seem forced. It might seem easier, and more natural, to write "I did this, and this was the result, and this is what it implies. Then I did that, and got these data, which I interpret this way. Finally I did this, with this result, which means the following." Actually, this *is* easier for you as the writer—but for two reasons it makes things harder for your reader (which is the more important consideration). First, relatively few readers will read Methods, Results, and Discussion with equal interest. Some skim or skip the Methods, some read Results but ignore Discussion, and some read Discussion but have little patience for detailed Results. Others may want to access one section directly: the Methods to borrow techniques, or the Results to add your data to a meta-analysis. Second, even if integrating all the sections worked well in principle, the canonical organization would remain superior simply because it *is* the canon: your readers are accustomed to this organization. When you use it, you meet their expectations and put information right where they will look for it.

Separating Results from Discussion does not mean presenting data entirely without comment. A good Results section draws your reader's attention to features of the data that you will later interpret. If you write "average dry mass was 14.2 ± 1.1 g for fertilized plants and 9.4 ± 2.3 g for unfertilized plants," you ask your reader to figure out what pattern there might be. Writing "fertilized plants grew over 50% faster than unfertilized ones (average dry mass 14.2 ± 1.1 g vs. 9.4 ± 2.3 g)" highlights a contrast that readers can relate to your story. The Results section can draw comparisons between controls and treatments, between experiments, or among experiments, observations, and theory. *Interpretation* of those comparisons, and comparisons with literature results, should generally be kept for the Discussion.

Despite my argument for separating Results and Discussion, the most common deviation from the IMRaD canon is their merger into a single section. Some journals allow this, others don't, and a few (notably *Nature* and *Science*) encourage it. If your target journal allows a choice, ask yourself the by-now familiar question: which arrangement favors the clearest communication with your reader? Sometimes, in a long paper with many sets of results, a merger reduces repetition and lets you discuss results that are fresh in the reader's mind. If you are tempted by such a merger, though, be sure of two things: first, that the length of your paper doesn't betray a failure to find your story; and second, that

the advantages of integrating Results and Discussion outweigh the cost of compromising your reader's finding system.

The separation between Methods and Results is much closer to a firewall. Your Results section may *mention* methods, but it should never *introduce* them. A reminder of where results came from can be useful: for instance, "Plants fertilized with slow-release micronutrients grew 50% better than unfertilized plants and 22% better than plants fertilized with micronutrients in aqueous solution." But such reminders should be brief and used only when a reminder of the methods is important to the reader's understanding of your results. In the other direction, you might occasionally mention a result in a Methods section, but only when it's needed for your reader to understand or accept the methods used, and it isn't discussed further. For instance, your Methods might read, "We used parametric ANOVA to test for differences between treatments. This approach is justified because our data showed no significant deviation from normality of residuals or homoscedasticity."

Carefully separating your results from your methods and discussion may leave a Results section that is startlingly short—perhaps a few paragraphs or even a few sentences (plus, of course, tables and figures). This does not mean that your work is trivial or your writing simplistic: it means that you have found an elegant approach to a well-defined research question, and presented it clearly to your grateful reader.

Organizing the Results

Even a very short Results section can benefit from careful organization, and longer ones require it. When your data and analyses are fairly simple, you do well to place the main result (the one that most directly answers your central research question) in the first paragraph. This is a "power position," in the sense that readers tend to emphasize material found there. Subsequent paragraphs can include data and analyses that support or complement your main result.

This main-result-first organization isn't feasible, however, for papers that make more complex arguments. Often, your "main" result is really a synthesis of several lines of evidence, or involves building later analyses on the results of earlier ones. In our star-formation paper, for instance, the main result (the fraction of massive protostars having close neighbors) depends on two intermediate results (protostar masses and nearest-neighbor distances), and the protostar masses in turn depend on rotational velocity measurements. Here it is more effective to work through the results from least to most complex, following the structure of your Methods section and using identical subheads. This organization brings the reader along as you build logically to your main result at the end (another power position). Your reader will appreciate a signal that they've reached that main result—something like "Finally . . . ," "Most important . . . ," or "Combining the results so far, we arrive at a test of our main hypothesis. . . ."

Perhaps the strongest organization for complex Results hybridizes the main-first and main-last techniques. Here, you work from basic to complex, but precede that with a very brief overview of the most important result. For our massive-star formation paper, we might open the Results with "Our analyses, taken together, show that only a moderate fraction of massive protostars occur in tight clusters . . ." before backing up to present the separate results that combine to establish that conclusion. The advantage of tipping one's hand at the beginning of the Results is that the reader knows in advance the full argument into which each result will fit.

Communicating Quantitative Information in Text, Tables, and Figures

Results sections are, almost without exception, numbers-dense, but without skillful handling, numbers make text dense and difficult to read. Tables and figures (I'll refer to these collectively as "graphics"), when handled well, allow efficient and reader-friendly display of numbers, and therefore will be important to almost every Results section you write. Graphics need not be quantitative and can appear anywhere in a paper; but the vast majority serve to display numbers in the Results, so I deal with them here.

Not every number belongs in a graphic, and whether placement in text, table, or figure works best depends on what it is you need to show

your reader. When you present just two or three numbers to make a point, place them directly in the text, where they won't interrupt your manuscript's flow. Using a graphic imposes a cost of navigation on your reader, who must divert attention away from the text, find and inspect the graphic, and then find their place in the text again to pick up the thread of your argument. For just a few numbers, placement in text is also much more compact than a graphic requiring legend, labels, and white space within and surrounding it. However, if you need to present more than a few numbers, placing them in your text can make your prose indigestible and the pattern obscure. Here, a graphic's ease of reading will compensate for the navigation cost of accessing it. This tradeoff is illustrated in Box 12.1.

Box 12.1 Numbers in text vs. graphics

For displaying two or three numbers, a graphic wastes page space, and asks your reader to divert attention from text to graphic and back again: "*Fertilized seeds were larger than unfertilized ones (Figure A)*". This could have been presented more effectively in text as "*Fertilized seeds were larger than unfertilized ones (1.8 ± 0.3 g vs. 0.9 ± 0.2 g)*."

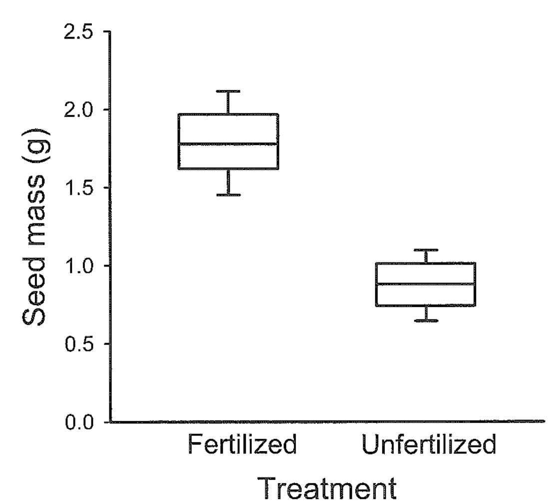


Figure A. Wasted: all this structure just to compare two numbers

(box continues next page)

(cont.)

Graphics pay dividends, though, for larger sets of numbers, because the help they give in seeing pattern outweighs the cost of diverted attention. Imagine being asked to wade through this: "*For lakebottom sediments collected at the inflow, average grain size was 1.3 mm (± 0.4 mm standard deviation). Smaller grains predominated farther out: average 0.8 mm (± 0.2 mm) at 5 m from the inflow, 0.2 mm (± 0.05 mm) at 10 m, 0.08 mm (± 0.03 mm) at 20 m, 0.012 mm (± 0.004 mm) at 40 m, and 0.002 mm (± 0.0005 mm) at 60 m.*" Ugh! The pattern is easily seen in a figure: "*Grain size of lakebottom sediments decreased away from the inflow (Figure B).*"

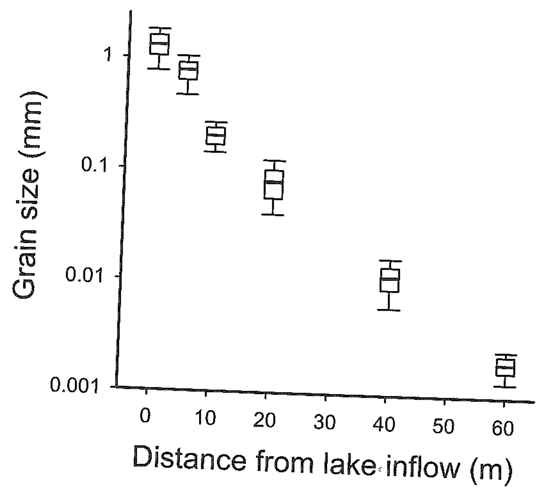


Figure B. An effective way to display pattern

Among graphics, tables excel at presenting datasets with many entries and/or variables (presuming that your story needs them). That's because adding more rows or columns to a table detracts little from the readability of what's already there. Tables are also the method of choice when your reader needs precise values of numbers. The major disadvantage of tables is that they aren't an effective way to show trends in data or relationships between variables (Box 12.2). Therefore, reserve tables for sets of numbers that stand individually (for example, a list of a mineral's properties—composition, hardness, cleavage, and so on), or for datasets where each reader will want to look up a few particular numbers (e.g., a list of molecular weights).

Box 12.2 The main weakness of the table

It's difficult for readers to see patterns or relationships among variables when you present data in a table. This dreadful example may be the worst sin I've ever committed as a writer (Heard and Remer 1997, but the blame is all mine):

TABLE 2

COEXISTENCE TIME AND CLUTCH SIZE FOR EQUAL COMPETITORS

Clutch Size, Species 1	Clutch Size, Species 2: 1	2	4	8	16
1	**134** (14)	82.3 (4.6)	49.5 (2.2)	27.3 (1.3)	16.9 (.8)
2	80.6 (5.5)	**145** (13)	60.2 (2.6)	31.5 (1.4)	17.4 (.7)
4	48.4 (2.5)	59.6 (3.0)	**182** (16)	41.0 (1.9)	19.4 (.9)
8	28.5 (1.5)	31.5 (1.5)	43.3 (2.4)	**520** (53)	26.6 (1.2)
16	16.6 (.6)	17.4 (.7)	20.3 (1.0)	26.0 (1.4)	**2,442** (213)

Note. —All parameters as in table 1, except $\alpha_{12} = \alpha_{21} = 1$. Coexistence times are as follows: bold, either species excluded at random; underlined, species 1 excludes species 2; regular font, species 2 excludes species 1. Numbers in parentheses are twice the standard errors.

I wanted to show that "coexistence times" are longer for equal (main diagonal) and larger (lower right) "clutch sizes." And I suppose you could figure that out, if you worked really hard at it—but who would? I used too many numbers with too much precision, and the underlining and boldfacing didn't help nearly as much as I thought it would. A figure would have shown the trends much more clearly.

Figures excel at communicating trends in data and, especially, relationships among variables. This is particularly true for relationships that are nonlinear: these are nearly impossible for readers to discern from a table, but readily apparent in a figure. However, figures do not display precise numbers well. They also do poorly at displaying more than a few variables at a time: adding more plots to your graph can quickly produce an indecipherable tangle of lines or forest of bars.

In summary, if you have just a few numbers to present, embed them in your text; for more, use a graphic. Among graphics, use a table to convey exact values, and a figure to illustrate trends and relationships. For any set of numbers, though, choose just one option: repeating the same information in text, table, and figure wastes your reader's attention and your publisher's resources.

Handling Numbers

Because numbers ask a lot of your reader, you should work hard to minimize those demands. Here are some ways to do that.

- **Winnow data for presentation**. Your job as a writer is not to blow the reader away with a numerical tornado; it's to present just enough data to tell your story convincingly. If several variables are strongly intercorrelated, display the most relevant one, or use a data-reduction tool such as principal components analysis. If there are several metrics that could quantify something, or several statistical analyses that could test for a pattern, present the one most applicable to your research question. If readers might question whether your results are robust to your winnowing, you can place alternative variables, analyses, or metrics in online supplements (chapter 14) or simply say, "Use of alternative metrics, such as this-and-that, yielded similar results."
- **Omit redundant numbers**. Don't inflate number content by providing several versions of what is fundamentally the same number. For example, suppose that I wrote "Among massive protostars, 13 of 25 (52%) had close neighbors, while for small protostars just 6 of 25 (24%) did." Here counts and percentages are redundant, because the denominators are the same so either can be compared easily. Along similar lines, writers reporting analyses of variance often provide degrees of freedom, sums of squares, and mean squares. But any of these three is easily calculated from the other two (MS = SS/df), so presenting two always suffices.
- **Emphasize the most important numbers**. When you do report multiple sets of numbers for a single result, let your reader know which are key and which play a supporting role. For example, imagine that you're comparing two fractions with different denominators, so it's helpful to provide percentages. There are two ways to phrase this: "25/46 vs. 23/67 (54% vs. 34%)," or "54% vs. 34% (25/46 vs. 23/67)." Use the former if you want the reader to emphasize counts and sample sizes, but use the second if you want the reader to compare percentages.

  A common emphasis problem arises in presenting statistical results. For instance, we could compare the fractions from the last paragraph, and write either "A χ^2-test showed that massive protostars were more likely to have neighbors, with $\chi^2_{(1)} = 4.47$, $P = 0.03$ (54% vs. 34%)" or "Massive protostars were more likely to have neighbors (54% vs. 34%; $\chi^2_{(1)} = 4.47$, $P = 0.03$)." The first wording emphasizes numbers associated with the statistical test, which play a supporting role but aren't what your reader picked up your paper to see. The second and much better wording emphasizes comparisons that tell your story.

- **Report only meaningful and necessary precision.** When you report a number, how much precision—"1.68234119478," "1.6823," "1.7," or "about 2"—should you provide? The answer depends in part on the number's "significant digits" and in part on what use the reader will make of it.

  Significant digits are those digits in a number that you know with reasonable confidence. These are determined by (1) the precision of the measurements that yielded the number; and (2) the propagation of uncertainty during calculation (when a number is derived from multiple measurements, each with its own uncertainty). If air currents in your lab make your balance precise only to the nearest 0.1 g, it would be silly to report that a seed had a mass of 1.6823 g; make it 1.7 g. Similarly, don't report a mean seed mass of 1.6823 g if the standard error is 0.01 g; make it 1.68 g. There are technical guides to determining significant digits (e.g., Robinson et al. 2005, Ch. 1; or http://www.hccfl.edu/media/181113/sigfigs.pdf), but as a rule of thumb, a digit is significant if you'd expect it to be consistent through multiple measurements or calculations of the same quantity. Reporting more digits than are significant makes your manuscript more difficult to read without adding any compensating information.

  Significant digits set the *maximum* precision you should report. However, the reader may not need that much precision to understand the story you are telling. If you can measure seed mass to the nearest 0.0001 g, should you report seed masses of 1.6823 vs. 0.7714 g for fertilized and unfertilized plants? Would this tell your reader more than 1.7 vs. 0.8 g? In most situations, it would not, and the rounded numbers are better.

  Statistics seem to cause many writers particular trouble with precision, probably because software tends to report test statistics and P values to many (unwarranted) decimal places. I frequently see manuscripts with statements like "seed mass did not differ significantly among treatments ($F = 0.92238674$, $P = 0.7826$)." Nearly all those digits mean nothing to the reader! "$F = 0.9$, $P = 0.8$" carries just as much information and does so more clearly. Two significant digits are normally plenty for P values, and two or three for most test statistics.

Designing Tables

There are many choices for arrangement of information in a table's rows and columns. Tables that are not carefully designed can be impenetrable. A full treatment of table design is beyond the scope of this book, but attention to a few basic principles can greatly improve communication with your reader. (For further guidance, see e.g., Tufte 2001, Council of Science Editors 2006.)

- **Arrange tables to take advantage of natural reading patterns.** Because English text is read from left to right, your readers will tend to follow the same pattern when inspecting a table. Therefore, place familiar or context-setting information in the leftmost column, and new or dependent information in columns to the right. For example, place variable names at left and their values to the right; independent variables at left and dependent variables to the right; or pretreatment conditions at left and posttreatment measurements to the right. Use similar logic to order rows from top to bottom.
- **Showcase patterns with vertical rather than horizontal display.** Readers will find it easier to see patterns if they can compare entries you've arranged vertically. Our positional notation system makes this especially true for numbers:

  1,359, 11,280, and 104,600 or

  1,359
  11,280
  104,600

  For similar reasons, tall narrow tables are easier to read than wide shallow ones. Keep in mind, though, that if you're asking your reader to see a complex pattern, a figure is probably better.
- **Format tables for easy reading.** Use design tools such as lines, white space, and indenting to draw the reader through your table's organization. Separate rows and columns well, lest you inflict on your reader an unbroken sea of numbers (or other table entries). Most journals allow lines between rows, but not between columns; use white space instead. Label rows and columns clearly, with a minimum of cryptic abbreviations, and if you want your reader to compare two rows or columns, place them beside each other.
- **Keep tables as few and small as possible.** Columns, rows, and even whole tables are so easy to create that they seem to multiply when you're not looking. However, even well-designed tables demand that your readers navigate to them and locate material within them (the latter becoming more difficult as a table grows). Keep only the essentials. Tables that need to be reproduced in landscape format or that span multiple pages are particularly difficult for readers. Finally, don't try to compress tables by using smaller fonts or removing white space; if your table won't fit on a page, compressing it will only make it unreadable.

Designing Figures

Figures come in a bewildering array of types: maps, photographs, line drawings, scatterplots, boxplots, pie charts, ternary plots, and dozens (if not hundreds) of others. Software makes it easy for you to create any of these, but offers you little help with choosing among them or designing the figure well. The overarching principle of figure design is to give your reader "the greatest number of ideas in the shortest time with the least ink in the smallest space" (Tufte 2001, 51). As with tables, a complete treatment is beyond the scope of this book, but some basic principles are worth emphasizing.

- **Use straightforward and familiar types of figure.** Makers of graphics software trumpet their packages' abilities to depict data in ever-more-dazzling ways. Bars can be grouped, stacked, color-coded, or built from icons; pie charts can be exploded; three-dimensional plots can be kriged, rotated, or heatmap-coded. But the fact that you *can* do something doesn't mean you *should*: glitzy and novel approaches to figure design are more likely to impede communication than facilitate it. Whenever possible, choose figure types that emphasize your data rather than the cleverness of your graphics software; these are usually types that are straightforward in construction and familiar to readers.

Figure types fall into three classes (with some overlap). Data reproductions are minimally processed representations of actual observations (photographs, instrument tracings, etc.). Schematics are simplified or abstracted material such as maps, flow charts, and line drawings. Data compilations display and summarize numerical data (scatter plots, bar charts, and so on).

Data reproductions should be used sparingly in written materials (although they are quite useful in illustrating talks). A photograph or instrument tracing is generally just a single datum, from which (alone) little can be inferred. Furthermore, data reproductions almost always include extraneous detail that captures reader attention but doesn't contribute to telling the story. Data reproductions can be useful as exemplars, giving readers a feel for the study system or the nature of the data. Their moderate use is conventional, furthermore, in some fields: papers in cell biology, for instance, often reproduce gels or stained tissues, and papers in in biosystematics often use photographs to illustrate features of typical specimens.

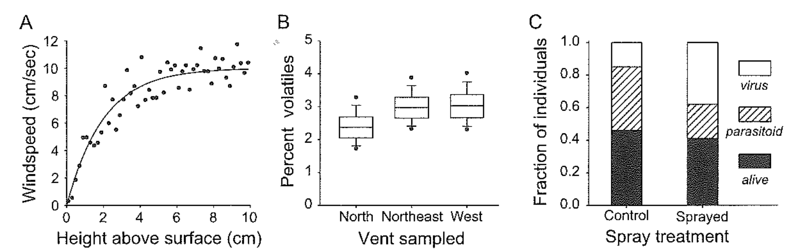
Schematics are generally used as visual support for in-text explanation. They can appear in any part of a paper: for instance, drawings of apparatus or flow charts summarizing algorithms often appear in the Methods. In the Results, schematics might include things like line drawings depicting features of fossils, or maps or diagrams of geological formations. The great advantage of schematics over data reproductions is their abstraction: they can remove detail to emphasize relevant features or extract generalizations from noisy observations. Of course, this means that they are interpretations rather than data, reflecting decisions by their makers that customize them to the story being told.

The line between data reproductions and schematics can become blurred, as when photographs are enhanced or retouched for publication. Such enhancement, when clearly disclosed, can do readers a great service: for instance, increasing the contrast of a photograph may make relevant features (such as faint stars in a telescopic view or

bands on a DNA-fragment gel) easily seen. Undisclosed, the same manipulation would be misleading at best and fraudulent at worst.

Data compilations account for the large majority of the figures we use. They come in many different types, which are best suited to highlighting different kinds of patterns. For instance (Box 12.3), scatterplots with fit lines are good for showing nonlinear relationships between two variables; box plots are good for showing comparisons between average measurements, and divided-bar charts for accurately depicting parts of a whole. Choosing an effective data-compilation type, therefore, means deciding which pattern, comparison, or other feature of the data you want your reader to emphasize. If, having decided this, you are still not sure what type to use, consult

Box 12.3 Some simple data compilations



Windspeed profile in a boundary layer

Volatile content of lava from three volcanic vents

Fates of spruce budworm larvae sprayed with water or virus suspension

These three graphs illustrate simple but effective kinds of data compilation. In each case, I've chosen a graph type that does a good job of highlighting the pattern I want the reader to see. Graph A shows wind speeds through the boundary layer above a solid surface; the scatterplot communicates variability while the fit line emphasizes the asymptotic nature of the curve. Graph B compares volatile content (water vapor, CO_2, etc.) of lava from three vents of a volcano; the boxplot makes it easy to see differences in means and spread. Graph C compares the fates of spruce budworm larvae in two virus-exposure treatments; the divided bars make it easy to see that the larvae killed by virus would mostly have been killed by parasitoids anyway. Many more graph types are available, each with its own advantages.

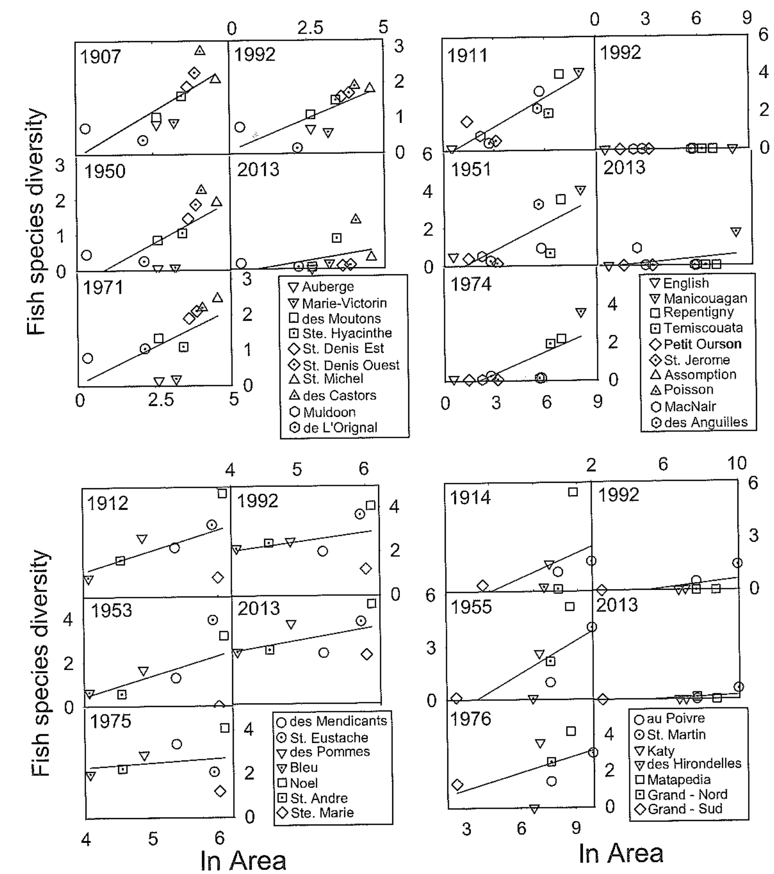
a technical guide on figure-making (Tufte 2001, Kosslyn 2006) or read the literature to find figures that effectively communicate patterns like the one you wish to show.

- **Make figures simple.** Software doesn't just make it easy to make figures—it makes it easy to make them complex, with multiple panels and many data traces (plots for different treatments, sets of observations, or variables). I recently reviewed a manuscript that included seven figures with an astonishing forty-three panels and eighty-four data traces—and for readers craving more, provided another eight figures, fifty-four panels, and sixty-eight data traces in online supplements. Another manuscript included a single figure that had twenty panels and ten different data-point symbols defined in four keys (Box 12.4). Perpetrators of figures like these don't understand that it's the writer's job, not the reader's, to extract pattern from data and relate it to the story being told. As a rule of thumb, a single figure should rarely have more than four panels (fewer is better), and a single panel should rarely have more than four data traces. If your figures are more complex than this, think seriously about breaking them up or deleting some elements (or at least moving them to online supplements where they can be more conveniently ignored).
- **Keep figures readable.** Most figures are reduced in size during publication, usually to half (or less) their sizes in your draft manuscript. To compensate, use thick lines and large fonts and symbols, and check to make sure figures are easily readable after such dramatic reduction. With electronic publication, readers *could* zoom in; but don't require them to do so, as it will disrupt their reading momentum. If using large elements makes your figure cluttered or densely packed, don't shrink the elements; simplify the figure.

When coding data points, lines, curves, and areas to a key, use symbols that are easily distinguished (such as filled and hollow circles or solid and dashed lines). Avoid subtle distinctions (triangles facing up vs. down, lines differing in thickness). Use consistent conventions across panels and figures: identical fonts, symbols, bar widths, axis scaling, and so on. Every shift in design between figures (say, filled vs. white bars in one figure and hatched vs. white in the next) asks the reader to master a new visual dialect. Provide symbol keys in the fig-

Box 12.4 Abusing your reader with a figure

This figure shows fish species diversity as a function of lake area, for several sets of pollutant-exposed lakes in Quebec, and for historical surveys over one hundred years. Or at least, it attempts to show all that. Its designer asks the reader to inspect patterns in twenty panels—five each for four sets of lakes. It's not obvious whether what's important is the overall picture (most panels show increasing trends), comparisons among panels within sets (slopes mostly decrease through time), or comparisons of corresponding panels among sets. Furthermore, the use of distinct symbols to key data points to their corresponding lakes suggests that perhaps the reader is being asked to think about each lake individually—but without being told why. (While I've disguised the real source of this figure, I haven't exaggerated its complexity.)



ure itself, not in the legend where they ask the reader to switch attention repeatedly between visual and text elements. Whenever possible, set axis scaling so that data occupy the whole plane of the figure (unless you are reserving white space for a symbol key, or unless white space results from keeping consistent axis scaling among panels or figures).

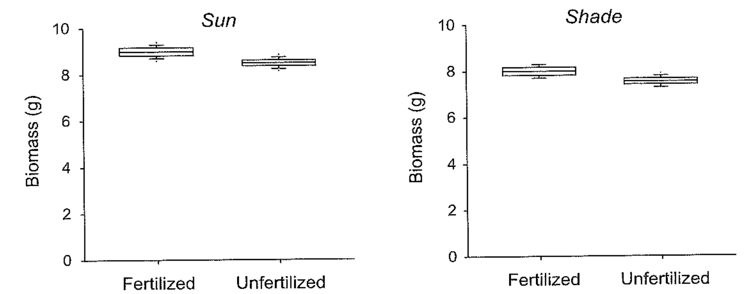
- **Minimize color.** Until quite recently, color figures were very expensive to produce and, not coincidentally, very rare. The recent trend to online publication has made color figures cheap and easy to publish—but that doesn't mean they're a good idea. While color can be useful in some situations (e.g., heat-maps for three-dimensional information, or showing multiple staining of biological tissues), it has significant disadvantages that are often overlooked. First, about three percent of the population have color-vision deficiencies, with red-green colorblindness the most common. These readers may not perceive color contrasts that you intend to carry information. Second, colors (and color contrasts) will shift depending on the frequency spectrum of ambient light and characteristics of the displaying device. Third, color figures published online are not always read there. For the foreseeable future, you should expect some readers to print your paper (most likely in black and white) before reading it. These issues make color figures advisable only when there is no other way to communicate the necessary information.
- **Don't let figures mislead readers.** Figure design always involves choices that affect readers' perception of pattern. Even an unretouched photograph is still composed and cropped to include or exclude detail that can change its message. Well-designed figures emphasize patterns the reader should see, but stop short of exaggerating patterns or implying them where none exist. This is something of a balancing act, and pulling it off requires some awareness of features of figure design that tend to mislead readers.

  A lot of misleading design is rooted in scale. Whenever possible, vertical axes should start at zero, and when two panels or figures are to be compared, they should have the same scale (Box 12.5). Any deviation from this practice should be clearly signalled to the reader.

  Fit lines are another easy way to mislead. The presence of a line through a scatterplot suggests a relationship, whether one is there or

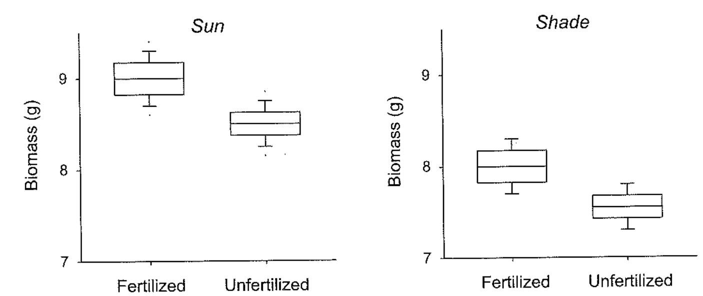
Box 12.5 Misleading figure scales

When you ask a reader to compare panels within a figure, those panels should nearly always have identical scales. To illustrate this, imagine that you've measured biomass for plants grown in sun and shade, and with and without fertilizer. This is easy to show as a boxplot:

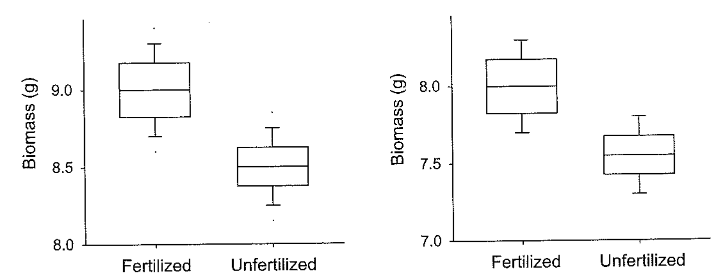


Notice that the panels' Y axes both begin at zero and have the same scale. It's easy to see that plants grew only slightly larger with fertilization and in sun.

It's tempting to think that much of the graph is wasted space—why not zoom in? Here the Y axes run from 7 to 9.5 g:



But the sun plot still has wasted space. You could start the shade plots at 7 g and the sun plots at 8 g.



(box continues next page)

(cont.)

Here's the problem: each pair of plots shows exactly the same data, but they give very different impressions of the experiment's results. The second version greatly exaggerates the differences among treatments, while the third conceals the sun/shade difference and magnifies the fertilization effect. At best these require too much reader effort: only by mentally reconstructing the first version can readers accurately assess the pattern in the data. At worst, they're misleading. Many other features of figure design can similarly distort patterns.

not. Don't show a fit line if the fit isn't statistically significant (so there is no reason to suspect a relationship at all), and don't extend one beyond the range of the underlying data. Think carefully before using lines to connect data points, especially when data points are few: this gives the impression of much greater certainty about pattern than really exists.

Many additional problems lurk in figure design. Partly because the use of graphic design to mislead is a major pillar of the advertising industry, we know a lot about human perception and misperception of visual pattern (e.g., Ware 2012). You can—and should—exploit the same knowledge to tell your story without misleading.

I've offered only a few rules of thumb for the design of both figures and tables. Fortunately, all such graphics decisions come down to the usual question: what will favor crystal-clear (and accurate) communication with your reader?

Relating Graphics to Text

Figures and tables should both stand alone and—paradoxically—be seamlessly integrated with the text. That is, a reader, once directed to a graphic, should be able to understand it without referring back to the text. At the same time, the text should make clear what the reader is looking for in the graphic, and upon returning from graphic to text the reader should easily see how the graphic's content moves the argument along.

For figures and tables to stand alone, they need both good design (above) and also helpful legends. (For no apparent reason, tables have "titles" above them, while figures have "legends" below them. Name and position are the only differences.) A legend should begin with a brief phrase identifying the key point or comparison the graphic makes (for the figure in Box 12.4, for instance, "Biomasses of fertilized and unfertilized plants in sun and shade"). The remainder of the legend provides further explanation. It should define any symbols, abbreviations, or other coding of information not shown in an in-figure key or table footnote. In multipanel figures it should explain the panels and call attention to relationships between them. It should explain statistical methods used in the graphic (specifying, for instance, whether error bars indicate one or two standard errors, or what statistical methods underlie fit lines). It may include a brief reminder of the methods behind the data displayed—just enough for a reader to understand the data without having to look back to the Methods section. All of this should rarely take more than two or three sentences.

Do not expect readers to interpret a graphic unassisted. The text should indicate what pattern they should look for, how that pattern relates to the point being made, and how to see the pattern in a complex graphic. Avoid vague references to graphics, such as "See Table 1 for activation energies in the presence of different catalysts." Instead, first identify the pattern of interest, then direct your reader to the graphic that displays it: "Activation energies were lowest on palladium catalysts (Table 1)." When a graphic is complex, steer the reader as specifically as possible to the relevant features ("Figure 1, compare leftmost bars across panels"). Finally, avoid referring to more than one graphic to make a single point. I once read a manuscript that included this sentence: "Diet overlap between species increased from 2004 to 2009 in four of six comparisons: ribbon snake–green snake, mud snake–milk snake, milk snake–ribbon snake, and milk snake–green snake (Fig. 2A-F, Figs. 3–6, Table 3)." This asks the reader to do some difficult data-analytic work, extracting and synthesizing information from four figures and a table, which the writer should have done instead.

Chapter Summary

- Results are typically presented independently of the Methods and Discussion. However, they may include brief reminders of methods used and highlight results or comparisons for later discussion.
- Numbers are demanding for readers. Their impact can be minimized by winnowing data, avoiding redundancy, emphasizing the most important numbers, and displaying only meaningful and necessary precision.
- Tables are best for presenting datasets with many entries and/or variables, or when readers need precise numeric values. They are ineffective at showing trends or relationships between variables. Figures are excellent for highlighting trends and relationships, but do not display precise numbers or more than a few variables well.
- Figures may be data reproductions, schematics, or data compilations.
- Tables and figures should be designed to give readers the most information "in the shortest time with the least ink in the smallest space" (Tufte 2001, 51).
- Text should point readers to important pattern in tables and figures, but tables and figures should nonetheless be interpretable on their own.

Exercises

1. This small dataset shows harvest dry masses (in grams) for plants grown under four treatments. By hand or using your choice of software, construct Results section elements to present these data in the following ways. For each, write down a list of design choices you made, and why you made each one.
   a. Using only text, not a graphic.
   b. Using a table (of summary values, not just repeating the raw data). Make at least two different designs.
   c. Using a figure (data compilation). Make at least two different designs.

Which way of presenting the data is least, and which is most, effective? Why?

Watered daily to flow-through		**Drought**	
Fertilized	Unfertilized	Fertilized	Unfertilized
22	21	15	19
25	21	17	11
24	19	16	14
32	24	12	17
23	25	19	16