SEVEN

Finding and Telling Your Story

The very first step in writing any scientific paper is to answer a simple question: what, exactly, will it be about? That is, what's your story?

Your "Story"

That your paper will be about your results—data from your experiments, proof of a theorem, observations with a new telescope—is obvious, but for two reasons, it's also unhelpful. First, it's rare for your results to consist of a single analysis from a single experiment that worked as you planned it and yielded the outcome you expected. Instead, you'll normally sit down to write with a mountain of data from multiple experiments and observations, some obviously critical to your research question, some obviously tangential, and an awful lot somewhere in between. Some results may be unexpected and some apparently conflicting. For each dataset you may have several alternative analyses. You may even find yourself piecing together datasets collected for different reasons by different people long before your own involvement, to test a hypothesis that was never considered when the data were collected. In other words, it may be obvious that your paper is about your results—but *which* results, and why?

Second, saying that your paper is about your results places the focus on the writer ("*your* results"), which is not where it should be. Good writing is oriented to the needs of the reader. You should ask not "what should I write about?" but instead "what does my reader need to hear about?" The distinction is subtle but important, and it leads you to identifying the story you have to tell.

What does it mean for your paper to tell a story? Successful fiction or drama sets up and then resolves some interesting question in a reader's mind, by exposing compelling characters to a well-defined plot (well, except for *Waiting for Godot*). A scientific paper does the same. It has characters: the rocks, chemicals, equations, or other entities that you studied. It has a plot: the methods you applied to your characters and the results you obtained from them. Most importantly, it raises and answers an interesting question.

The central question and its answer are the purest distillation of your story. A clear central question gives your paper a single, obvious direction. Every element of your paper can then work together to draw the reader inescapably to your question's answer. You should be able to state question and answer in a sentence or two—perhaps not in a way useful for a reader, who would need definitions, context, and so on; but in one that defines the story for you. Whenever you are not sure whether a dataset, analysis, figure, table, or anything else belongs in the manuscript, referring to this two-sentence mini-summary should give you the answer: would including it help tell the story or distract from it?

For a concrete example (one we'll return to repeatedly over the next several chapters), imagine that you're an astronomer interested in the formation of massive stars (Box 7.1). You have used the new Atacama

Box 7.1 Star formation

Briefly, stars form as clouds of interstellar dust and gas collapse under their own gravity. As a dense core (a protostar) forms, gravitational potential energy is converted to heat. The heated material emits radiation, supplying an outward force (radiation pressure) that increasingly opposes the gravitational pull on remaining matter in the cloud. Matter accretes to the protostar until it has become large and hot enough for radiation pressure to balance gravitational pull.

This process is well understood for small stars, but not for larger ones (more than about ten times the mass of our sun). The simplest models suggest that radiation pressure should become too strong for further accretion before the protostar can reach such masses. One possibility is that radiation doesn't escape equally in all directions, but rather in concentrated jets that clear away material in one direction but allow accretion from others (Banerjee and Pudritz 2007). Another is that massive stars form at the center of star clusters, where accretion can be driven by the combined gravity of many protostars (Bonnell and Bate 2006). The question is interesting because massive stars are rare, only about 0.2% of all newly formed stars (Chabrier 2003), but we don't know why.

The rival models of massive-star formation make different predictions about the appearance of massive protostars and their distribution in space, which are testable with sufficient observational data. Fortunately, gas clouds containing protostars are fairly common, with at least 6,000 large clouds and many smaller ones in our galaxy alone (Sanders et al. 1985). Regions of massive-star formation close enough for detailed imaging include the Orion, Eagle, and Carina Nebulae (1,300, 7,000, and 8,000 light years from Earth).

Large Millimeter/Submillimeter Array (ALMA; a radiotelescope complex in Chile) to image the Eagle, Orion, and Carina Nebulae, where new massive stars are being formed. ALMA is more sensitive and has higher resolution than any other millimeter-wave telescope, and you have reams of imagery, spectral data, and so on. But what story will you tell? Here are two-sentence summaries of some ways you might present your ALMA data. These are framed informally, because they aren't intended for anyone but the writer.

Very bad:

ALMA provided the most detailed images ever made of the Orion, Eagle, and Carina Nebulae. I explain how ALMA works, and show some of the images.

This mini-summary focuses on the research done, rather than the question answered. It makes no reference to star formation (or to any question at all). It forces the reader to do the job of analyzing and interpreting the data, but most readers will simply move on.

There is, of course, always more than one story that could be told with a set of data. If the paper was intended to report on ALMA's capabilities, rather than to test models of star formation, this mini-summary wouldn't be as bad.

Bad:

I present images of star-forming regions in the Orion, Eagle, and Carina nebulae, taken at 8 wavelengths from 0.4 to 10 mm. They show many protostars, some in groups.

This summary mentions star formation, but asks no specific question about it. It defines methods and offers data that are presumably relevant, but the job of figuring out how is left to the reader.

Better, but not good:

I outline what we know about star formation and stellar evolution. I present ALMA images of massive protostars at various stages of stellar evolution.

Here the writer makes some attempt to establish the relevance of the data, but "star formation and stellar evolution" is far too broad for a single paper, and no specific question is offered. The images presumably tell the reader something about massive-star formation, but it's not clear exactly what.

Very good:

If massive stars form with gravitational assists from their neighbors, then massive protostars should always appear among other protostars. However, solitary massive protostars are common in ALMA images of the Orion, Eagle, and Carina nebulae.

This mini-summary has a clear central question (do massive protostars form via combined gravity of neighbors?), and offers an answer (no). The question drives the function of each part of the manuscript: the Introduction will set up the cluster-gravity hypothesis and specify what's needed to test it; the Methods will explain how we can measure protostar masses and distinguish between solitary protostars and sets of neighbors; the Results will tabulate frequencies of solitary and neighbored protostars; and the Discussion will interpret the Results as a test of the cluster-assist hypothesis.

Perhaps the case for having a clear story is so obvious that it's hard to picture yourself sitting down to write without one. However, it's remarkable how often I see draft manuscripts that betray their authors' uncertainty about the topic. Among the symptoms: a long and convoluted title, or a short but vague one; data presented but never analyzed; figures or tables that don't relate to any hypothesis raised in the Introduction or that don't contribute to the Discussion; or topics appearing in the Discussion that weren't broached earlier in the manuscript. The writers of such manuscripts have omitted a critical step in their work: they have not found their stories.

Finding and Planning Your Story

Finding and planning your story means accomplishing three things. First, you must identify your central question and its answer (the two-sentence mini-summary from the last section). Write this down, even though no reader will ever see it; the writing will force you to be explicit. Second, you must decide which information, data, analyses, and interpretation belong in the paper, and which are better reserved for another manuscript or abandoned to a dusty row of old notebooks. Third, once you've identified the content, you must decide the order in which to present it. Of these, the first point is finding a story to tell, and the second and third are planning the best way to tell it.

Because the path from concept through execution to data analysis and interpretation is rarely simple, finding your story uses hindsight as much as foresight. Writers who don't realize this often cling to reporting everything they did, in the order they did it, including experiments that turned out to be blind alleys and observations that seemed relevant at the proposal stage but became immaterial to the conclusions eventually drawn. Experienced writers, in contrast, take full advantage of hindsight and work to determine the story they want the reader to hear (whether or not it's the one they had in mind when they began the work), and what information the reader needs to understand that story.

Do not hope to find your story by sitting down at your keyboard and beginning to type the first line of your Introduction. Instead, begin by brainstorming possible content and then selecting and organizing the content that defines your story and tells it effectively. There are many techniques for finding and planning your story; what follows is a toolbox from which you can select. Many writers apply more than one of these techniques to each writing project.

Figure 7.1. My wordstack for this chapter.

cohesive story. . . . "thesis"
?titles - shortest summary of story
outline
story about how I thought I didn't do outlines
when to outline – when story is ready; vs. to find the story
head and subheads as coarse outline
topic sentences as detailed outline
what goes in and doesn't, and what order
    not everything you did
    not in the order you did it
concept map: non-linear
intermediate step to outline
despite HTML, basic form still linear
wordstack/idea pile
    first step
    accumulate pre-writing
?Cahill – "pitch"
fail to consider story: leads to overlong, poorly organized MS
IMRaD structure
online supplements
"retroactive storytelling"
avoid lock-in
simple clear direction
selling the story
    not "I was interested in"
    not "no studies have examined"
    not "increase our understanding of"
contrast "writing backwards" (Magnusson 1996)
figure shuffling

Wordstacks. A wordstack (Figure 7.1) is an unsorted list of points you think might be useful ingredients in your manuscript. Each point can be a single word or short phrase indicating a relevant fact, idea, or topic, or can be a roughly sketched graphic. Your wordstack might have some hierarchical structure (some points having subpoints) if this is immediately obvious, but don't force it.

I illustrate wordstacks and the next two techniques with examples for this chapter because doing so makes clear the relationship between the wordstack and the finished product. Notice that it isn't a very close one. A wordstack is a venue for brainstorming, and it's more important to get ideas down than to worry about their being fully developed, related to each other, or in logical order. For the same reason, don't worry if you're not sure an entry really belongs: note your skepticism (I use question marks), but don't remove the entry. The point of a wordstack is not to organize your material or fully define your story, but rather to display the raw material from which you can draw. In moving from wordstack to finished chapter, I included one point I'd marked as questionable (pitch) but left out another (titles). I also left out material I hadn't initially questioned (a story about my outlining habits) and included material not in the wordstack (the story summary). All this is normal: the wordstack is a tool, not a product.

Concept maps. A concept map (Figure 7.2) is a tool for exploring relationships among concepts (Novak and Cañas 2008). It consists of a set of nodes connected by lines. Each node is a concept: a word or phrase denoting an idea, a thing, or a property of one of these. Each line connects two nodes and is labelled to indicate the relationship between those concepts. Typically, most nodes will be nouns or adjectives, while most line labels will be verbs.

To construct a concept map, first identify (provisionally) the most general concept involved in the topic of your manuscript. For this chapter, I chose "clarity"; for our astronomy example, one might choose "star formation." Then add concepts from your wordstack or others that suggest themselves as you build the concept map. As you add each concept, think about which others it relates to, and how; indicate these relationships with labelled connecting lines. I added "appropriate content" to my concept map, and indicated that clarity depends on determining this; I later broke "appropriate content" down by indicating that it consists of "content items" that are "few" and "ordered." Of course, with sufficient imagination one could connect every concept to every other one, so part of the job is to choose the relationships that you want your reader to think about. Place more specific, narrowly defined concepts

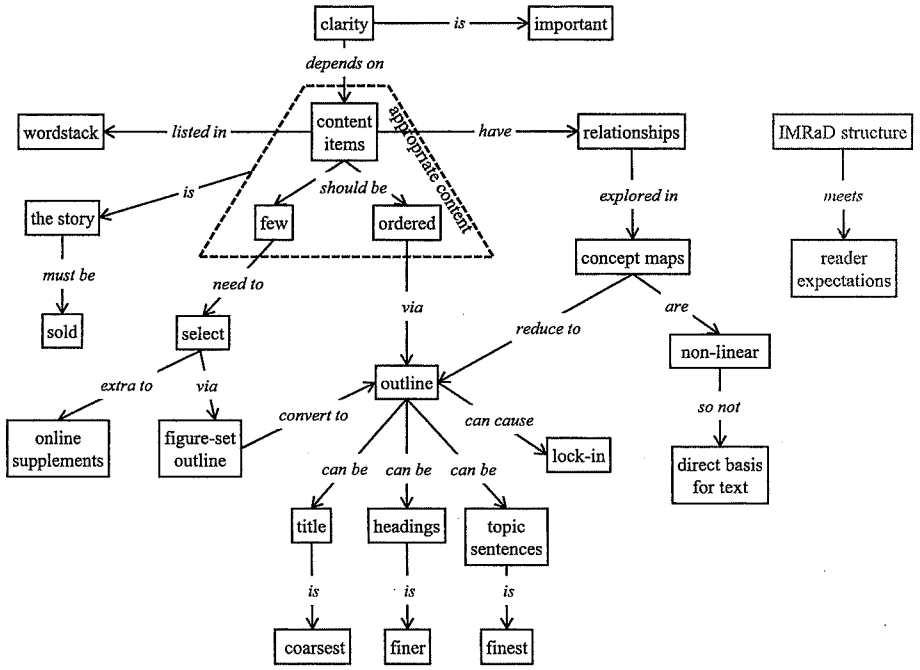

Figure 7.2. My concept map for this chapter. Concepts appear in roman font, with connections in italic. Concepts in gray failed to connect to the core structure, and were omitted from the chapter.

lower down, so that your concept map is roughly hierarchical. You will likely rearrange your concept map repeatedly as it grows (perhaps using CmapTools software: http://cmap.ihmc.us).

The completed concept map depicts your decisions about the logical relationships among material you might include in your manuscript. It doesn't fully define your story, because it's likely to show too much complexity, but your story is implicit in the main flow of connections among concepts. You may find that your concept map includes nodes or sets of nodes that are not connected to the rest of the map (for example, the gray nodes in Figure 7.2). These suggest material that doesn't belong. Your most important material is likely to appear near the center top of your concept map, with strong connections to other concepts—as it does in my example (the trapezoidal supernode in Figure 7.2).

An important property of concept maps is that they can, and usually do, incorporate branching, loops, and other network structures (as in

the three paths from "content items" to "outline" in Figure 7.2). This is a strength of the technique for initial exploration of your material, but a shortcoming of its use for translating material to text. We've spent several decades experimenting with crosslinked hypertext on the World Wide Web, but a linear sequence of ideas is still the most effective way to communicate complex information. It's the writer's job to determine *which* linear order makes the material clearest for the reader. Converting a concept map into a manuscript, then, requires an intermediate step to linearize it. The best way to do that is to construct an outline.

Outlines. An outline (Figure 7.3) is an ordered list of topics or points that summarizes the intended content of your manuscript. It is the result of selecting and ordering topics from the larger set available in a wordstack, or of linearizing a concept map. An outline has an intended 1:1 mapping onto a completed draft, such that expansion of each outline point, in turn, into a section of text is sufficient to produce the draft.

Figure 7.3. My outline for this chapter.

1. The concept and importance of your "story"
    - 1.1 what's a "story" (simple, clear direction)
    - 1.2 importance of story
2. Planning the story
    - 2.1.1 wordstack
    - 2.1.2 concept map
    - 2.1.2.1 use in learning
    - 2.1.2.2 use in writing
    - 2.1.3 outlines
    - 2.1.4 story summary
    - 2.1.5 subhead outline
    - 2.1.6 topic-sentence outline
    - 2.1.7 figure-set outline
    - 2.1.8 title
3. Hindsight storytelling
4. Revising the outline
5. Online supplements
6. Selling the story

(That this mapping is *intended* recognizes that reorganization during writing is common; see "Revising the outline," below, and notice the differences between the outline in Figure 7.3 and the chapter you are reading.) In addition to identifying and ordering included topics, outlines usually indicate hierarchical organization of subtopics into topics.

While there may be as many ways to outline as there are writers who outline, three particularly useful approaches are the story summary, the subhead outline, and the topic-sentence outline.

- **The story summary.** A story summary consists of answers to the following nine queries about your work and your story:

    1. What is the central question?
    2. Why is this question important?
    3. What data are needed to answer this question?
    4. What methods are used to get those data?
    5. What analysis must be applied for the data to answer the central question?
    6. What data were obtained?
    7. What were the results of the analyses?
    8. How did the analyses answer the central question?
    9. What does this answer tell us about the broader field?

(If you're a theoretician, to apply the story summary to your work, think of "data" defined broadly as the outcome of models.)

Think of the nine queries as a form with nine fields to be filled in. Queries 1–3 and 8–9 should be answered with a single sentence each. Queries 4–7 may need a bit more if multiple experiments, datasets, or analyses must be combined to answer the central question, but a few short sentences should suffice. For our star formation example, for instance, query 4 might need three elements:

> *4a. Rotational velocities from red/blue shifts of ^{13}C emission line.*
> *4b. Protostar masses from plot of radius vs. rotational velocity.*
> *4c. Nearest-neighbor distances from annual parallax[1] of each protostar.*

[1] Annual parallax (the shift in direction to an object viewed from opposite sides of Earth's orbit) indicates distance from Earth to the object.

When the story summary is complete, you've outlined a whole paper: answers 1–3 for your Introduction; 4–5 for your Methods; 6–7 for your Results, and 8–9 for your Discussion.

A story summary doesn't provide a complete manuscript outline. You will likely expand it later into a subhead or topic-sentence outline. Its value lies in forcing attention on your story and on what material the reader needs to understand it. For example, query 3 doesn't refer to all the data you have gathered; it asks about the data needed to answer your central question (a different and usually much smaller set).

- **The subhead outline.** A subhead outline is made up of phrases or other entries intended for use as headings and subheads in the completed manuscript. For a standard scientific paper, the top-level headings are nearly always Introduction, Methods, Results, and Discussion (chapter 8), although they will differ for other writing forms. Each top-level heading can receive subheads dividing it into logically distinct subtopics, and these can in turn be divided further. The manuscript is then written around the subheads, as blocks of text are inserted under each.

  Think carefully if you find yourself tempted to use more than three levels of headings (such as, for our star-formation example, "*Methods—Protostar neighbor distances—Parallax measurement*"). The function of subheads is to communicate organization to the reader, and if the organization of your paper into major topics is complex enough to require fourth-level subheads, it's probably too complex! Material below third-level heads should usually be simple enough for a reader to be guided through by normal paragraph and sentence structure (chapters 17–18). If you feel the urge to elaborate your outline further, that's fine—but do it by a topic-sentence outline instead.
- **The topic-sentence outline.** A topic-sentence outline has one entry for each intended paragraph of the completed manuscript. Each entry is a complete sentence that expresses the topic of its paragraph (and is thus suitable for use as the first sentence of that paragraph). Topic-sentence outlines are more detailed than subhead outlines, in two ways. First, they are more finely resolved. Second, they specify the material to be written more completely, because they consist of logical

statements that summarize points to be made, rather than just naming topics. Returning to our star-formation example, a topic-sentence outline for the third-level head "*Methods—Protostar neighbor distances—Parallax measurement*" might consist of two sentences:

> Parallax of each protostar is measured by change in right ascension and declination[2] between March and September sightings.
>
> Right ascension and declination are measured relative to distant galaxies, for which parallax is negligible.

As you write, you'll expand each topic sentence into a paragraph giving further detail: in this case, how the measurements are made, which reference galaxies are used, and so on. A topic-sentence outline is complete when it contains enough detail for this expansion to seem straightforward to you. While this is an imprecise standard, it isn't very important to know when your topic-sentence outline is "complete"—because the process of expanding it simply grades into writing of the manuscript itself.

Figure shuffling. Figure shuffling is an alternative to outlining that focuses on data and analyses as the elements that define your story. Before writing, you probably spend lots of time doing exploratory data analysis: plotting relationships between different pairs of variables, summarizing data in tables, running alternative statistical tests, and so on. You may have dozens of rough figures and tables at hand; some belong in your manuscript and some don't. Figure shuffling involves pinning these up on a wall and winnowing and shuffling them to produce a set of reasonable size representing the story you want to tell.

When you're done, every figure and table you've selected should be essential to telling your story. If one isn't, take it out. How many should remain depends, of course, on the work and on conventions in your field, but I think you should be reluctant to include more than about five figures and four tables in a single paper. If you need more, you may not have found your story. You will also have decided on the order in which

[2] Right ascension and declination are astronomical coordinates that express the direction from Earth to a viewed object; they are analogous to longitude and latitude.

the figures and tables make the most sense. Shuffling figures without thinking about methods allows you to ignore the order in which you did the work (irrelevant!) and focus on the order that best communicates the results.

Figure shuffling may seem to put the cart before the horse, because it doesn't start by identifying your central question. Instead, it finds the answer, and the question is implicit in that. Figure shuffling can be thought of as a way to "write backwards" (Magnusson 1996). When you write backwards, you begin by identifying your most important conclusion. Then you write your results, including only those necessary to support the conclusion. Figure shuffling accomplishes these two steps. Next, you write the methods necessary to obtain the results, and then the discussion that sets your conclusion in a broader context. Finally, you write the introduction, setting up the central question to which your conclusion is the answer. This technique can be very effective in finding your story, because it divorces the presentation from *what you started out to study* and *what you did* and focuses it instead on *what conclusion the reader should take from the work.*

Wordstacks, concept maps, various kinds of outlines, and figure shuffling can complement each other, but few writers use them all. My own mainstays are wordstacks, subhead outlines, and topic-sentence outlines; constructing those (in that order) usually suffices for me to find and plan my story. Experience will show you the set of techniques that helps you find and plan yours.

A Caution on Leaving Things Out

Leaving things out is critical to finding and telling your story. However, it's very important to think carefully about what it means to present "only those results necessary to support your conclusion." This does *not* mean omitting results that conflict with your conclusion, which, of course, is unethical. It does mean omitting results that are not relevant to your conclusion, or are redundant with others that suffice to support it. These are some of the most important judgements you can make as a

scientist, and there is no simple prescription for making them other than experience and careful thought.

Revising the Outline

Finding and planning your story is essential, but danger lurks in outlining, concept mapping, and the like nonetheless. Remember that these are tools, not straitjackets. As you flesh out an outline into a manuscript, you may feel something not working—perhaps a topic appearing early in your outline doesn't seem to fit comfortably there anymore, or a new topic is clamoring for inclusion. Should you stick to your outline, which represents the story that you've planned, or change it?

The answer, of course, is that it depends. On one hand, there's no point using a story-planning technique if you're going to completely ignore the plan you've made. On the other, the thinking you do as you write can change your interpretation of your data, or otherwise alter the story you're telling. Getting locked into the outline fossilizes your thinking and closes off avenues for improvement.

The happy medium is to think of your outline as providing an explicit criterion against which you can measure a potential change. Ask yourself critically: does adding new material to the outline improve the story? If you can explain why it does so, then include it; if not, it's distracting you from your planned story. Does reordering topics or deleting an outline point let you tell the story better? If and only if you're sure that it does, make the change. For example, in writing this chapter from the outline in Figure 7.3, I deleted my intended section 2.1.2.1 on concept maps in learning, after realizing it was irrelevant to my story. I moved section 5 on online supplements into chapter 14, where it fit better. Section 3 on hindsight storytelling turned out, once I'd written most of it, to fit better near the beginning of section 2 on planning the story. In each case, I thought carefully about which organization suited my story better and then made the change. This kind of reluctant willingness to change the outline makes story-planning a dynamic process that continues throughout writing—but a self-guiding one, as the current plan always provides a benchmark for assessing revisions.

Selling Your Story

Finding and planning your story is part of your job, as a writer, to work toward effortless reading. But your job doesn't end there. Remember that your work competes for readers' attention with an ocean of published material. A story that reads effortlessly improves your competitive position, but you also need to tell readers why they should spend their time reading *your* work, rather than somebody else's. You need to sell your story.

You may be uncomfortable with this. We are told from childhood not to brag, and many scientists are introverts at heart. This discomfort is likely responsible for such timid offerings as "I was interested in studying X," "further studies are required to increase our understanding of X," or "no studies have reported results of experiment/observation X." Many conceivable studies could pass these tests and yet be of little interest to readers. It's tempting to think that your good science will speak for itself, and readers will know why it's important. But few will invest the effort to read through a paper unless its importance is established explicitly right up front. So sell you must.

It's crucial to understand that what matters is your central question's importance to your reader, not to you. Perhaps you were motivated to study massive-star formation by your intrinsic fascination with the physics of collapsing gas clouds. That's perfectly fine as a reason to go to your office every day, but it won't get you an audience. Instead, you need to connect the narrow subject of your manuscript to more general issues that people care about. For instance, nobody knows why massive stars are rare, but they play an important role in the universe: only in such stars can fusion produce elements heavier than carbon, and only through their supernova explosions are such elements available for incorporation into smaller stars, planets, and people. You could explain that your study of massive-star formation will help us understand the process that made the universe suitable for life.

There are alternative ways to sell any story, and the approach you choose determines the set of interested readers. This, in turn, determines the journals for which your manuscript is appropriate. Manu-

scripts with larger sets of interested readers tend to be accepted by higher-impact journals and to accumulate more citations. Among common and effective ways to sell a story are pitches like these:

- "There's a controversy in the literature over issue X, and I present the kind of data needed to resolve it."
- "The fact that we don't know X hinders our efforts to understand issue Y, which is central to a developing subdiscipline."
- "Our lack of understanding of thing X impedes our efforts to solve economic problem Y."
- "We need to know more about thing X because it's a model system widely used to investigate problems in field Y."
- "I have discovered thing X, which suggests a way to make progress toward difficult-to-reach goal Y."

Any manuscript can be pitched in different ways—those above, and more. Some pitches will excite more readers than others, but just as important, a chosen pitch may excite a particular audience or demonstrate a fit with a particular journal. You can think of these different pitches as slightly different definitions of your story, and you can put them to work to best sell what you have to say to the audience you'd like to say it to.

Chapter Summary

- A paper has a story, with "characters" and a "plot," and it raises and answers an interesting question.
- Tools for finding and planning your story include the two-sentence mini-summary, wordstacks, concept maps, figure shuffling, and outlining. Outlines may be story summaries, subhead outlines, or topic-sentence outlines.
- Telling your story isn't enough; you must sell it, too. This means showing how your work solves a problem, or answers a question, that matters to readers.

Exercises

1. For a paper you've recently read, write a mini-summary, a concept map, and a story-summary outline. Can you suggest an alternative way of organizing the same content, and does your alternative tell a better story or a different one?
2. For a writing project you've recently started or plan to start soon, write a mini-summary. Next, make a wordstack, a concept map, and a subhead outline. If you wrote a new mini-summary, would it differ as a result of the story-planning process?
3. For the project outlined in (2)—or if possible, for a classmate's or colleague's—write three different selling pitches, each no more than two sentences.