Revising for reader ease 1: order of information and consistent terms

Key points:

- Introduce new information by relating it to the overall story
  - Only matters if paragraphs are unified and all logical steps are included
  - To revise order of information, make the subject of each sentence something previously mentioned
    - May involve omitting or combining sentences
    - Makes subjects of sentences relate to those of earlier sentences
- Unrelated subjects indicates you might need to revise
  - Can revise by trying to make subjects related
- Use the same term for the same thing
  - If you must use a different term, introduce it by relating it to one you've used recently

The overall goal of these lessons is to help you produce clear scientific writing. For your writing to be clear, it must be written in such a way that most readers will be able to follow it; that is, readers should be able to see how each idea fits into the overall argument or story as it's introduced. (Sometimes this quality is called "coherence" or "flow," but we won't bother with such terms since they have a lot of meanings). The key phrase there is "as it's introduced"—not at the end of the paragraph or section, but when the new thing appears. If your reader has to wade through a sentence (or several) to see how the new piece of information is related to the topic of the paragraph, he or she will likely be confused in the mean time about why you introduced the new thing.

Order of information

You've probably experienced this temporary confusion as you've read a paper before, but to illustrate how to fix this problem in your own writing, here's an example:

Example 1 version 1: new players first, context later

Over the past few years, candidate cancer stem cells have been identified in a variety of human malignancies including leukemias and a number of solid tumors such as glioblastomas, medulloblastomas and carcinomas [11]–[24]. Breast cancer is the first human carcinoma for which a putative cancer stem cell subpopulation has been isolated [25]. Using *in vitro*-separated tumorigenic cells from malignant human breast cancer-derived pleural effusions, Al Hajj and colleagues isolated a cell population characterized by high CD44 expression and low or undetectable levels of CD24 ($CD44^{+}CD24^{-/low}$) [25]. These cells were highly tumorigenic when injected into immunocompromised NOD/SCID mice and shared classic features of normal stem cells, including the capacity for self-renewal and generation of heterogeneous progeny [25]. The stem/progenitor cell phenotype of these cells was further refined by the Daidone group, who were able to grow mammospheres from single-cell suspensions obtained from the dissociation of primary breast tumors [3]. Mammospheres are non-adherent spherical cell clusters obtained in selective culture conditions, that have been shown to be enriched in

mammary stem/progenitor cells [26]. The vast majority of cells in culture were $CD44^{+}CD24^{-/low}$, and 10 to 20% of these retained the ability to self-renew [3].
From Morel A-P, Lièvre M, Thomas C, Hinkal G, Ansieau S, et al. 2008 Generation of Breast Cancer Stem Cells through Epithelial-Mesenchymal Transition. PLoS ONE 3(8): e2888.

Analyzing this paragraph will reveal some prerequisites for making writing easy to follow and lead us to ways to do that. The example may or may not seem to you like it needs improving; most readers would likely understand the point of the paragraph and how each piece relates to it. The reasons it's comprehensible (as a whole, by the end of the paragraph) are that it's all about one thing (isolation of breast cancer stem cells), the pieces of the story are logically ordered (general to specific, then by chronology), and all the pieces necessary to the point are included. If you have a piece of writing that lacks these qualities, then working on the order of information within sentences won't make the text any easier to follow—you first need to go back and work on development and organization. Once you've completed those steps of revision, though, you will likely need to work on making the draft flow smoothly, so the process we'll use to fix the example will be useful.

→ *Can only revise unified, organized paragraphs for reader ease*

Though you likely understand the paragraph, it's difficult to follow as you make your way through it. For example, the authors narrow down from all sorts of cancer in the first sentence to breast cancer in the second, and the reason is unclear until you reach the end of the sentence. The paragraph then jumps into the details of a specific study, which you might guess has to do with cancer stem cells, but the authors don't confirm that until the fourth sentence. Further, the details about the marker expression of the cells don't seem particularly relevant until the end of the paragraph, when we see that it matches that of the other study. The fifth sentence says that it provides further evidence for the previously mentioned cells' stem-ness, but it's not clear how the information supports the characterization until a sentence later (again). Reading the paragraph thus feels like a series of two steps forward and one step back.

→ *If new things are introduced without context, the writing is hard to follow*

Revise by making the subject something previously mentioned

The example would be much easier to read if each sentence felt like a single step forward, but how do you make that happen? Let's consider why we get that "two steps" feeling—new things are introduced before their relation to the story is explained. Reversing that, or putting the relation to the story before the new thing, would keep the reader from feeling like he or she missed something. Let's apply that to the example:

Example 1 version 2: context as the subject

Over the past few years, candidate cancer stem cells have been identified in a variety of human malignancies including leukemias and a number of solid tumors such as glioblastomas, medulloblastomas and carcinomas [11]–[24]. The first human carcinoma for which a

putative cancer stem cell subpopulation has been isolated is breast cancer [25]. This isolation was accomplished by Al Hajj and colleagues, using *in vitro*-separated tumorigenic cells from malignant human breast cancer-derived pleural effusions. The cells they isolated shared classic features of normal stem cells, including the capacity for self-renewal and generation of heterogeneous progeny, and were highly tumorigenic when injected into immunocompromised NOD/SCID mice [25]. The stem/progenitor cell phenotype of these cells was further refined by the Daidone group, who used single-cell suspensions obtained from the dissociation of primary breast tumors to grow non-adherent cell clusters enriched in mammary stem/ progenitor cells, i.e. mammospheres [3], [26]. The vast majority of these cells in culture had the same surface protein expression profile as those from the Al Hajj study: $CD44^{+}CD24^{-/low}$, and 10 to 20% of these retained the ability to self-renew [3].

These revisions don't follow an obvious formula because each change addresses a distinct issue that keeps the reader from following easily. The second sentence already contained multiple things that were already mentioned, so one of them became the subject of the sentence; "putative cancer stem cell subpopulation" could have worked. Since "carcinoma" was at the end of the previous sentence, that seemed like a smoother progression. Likewise, the third sentence contained multiple references back, but beginning with the phrase that contains "breast cancer" would have required more rearranging and might have sounded more awkward. Other rearrangements involved relationships across multiple sentences. In the original, the need to mention mammospheres was obscure until the following sentence, so I rearranged so that the term that had appeared before ("stem cell") introduced it. Similarly, since the relevance of "$CD44^{+}CD24^{-/low}$" wasn't clear until the last sentence, and explaining its relevance in the third sentence would require referring to a study not yet described, I moved the mention of it regarding the first study to the last sentence. These revisions demonstrate the variety of ways you can move the relationship (or the previously used term) to precede a new one—sometimes it's just a matter of restructuring individual sentences, and sometimes it involves combining them.

- → *If multiple pieces of context could serve as the subject, choose the one that refers to the previous sentence (or the end of the previous sentence*
- → *Introduce a new thing when it's relevant to the story*

Subjects of sentences are related

If this way of figuring out how to revise seems loose or open-ended, you might consider another approach that focuses on the structure of each sentence. Since we're trying to get a paragraph that tells a clear story, we can focus on what the story is about in each sentence—the subject. If the story progresses logically, the subject of each sentence will be related to the subject of the one before. This doesn't mean that all, or even most, sentences in a paragraph should have the same subject—that would mean there is no story, just a list of descriptions of the same thing. To see whether our revision made it easier to follow the paragraph, let's compare the subjects of the sentences:

Example 1 analysis: subjects of easy-to-follow paragraph are related to one another

candidate cancer stem cells	candidate cancer stem cells
Breast cancer	The first human carcinoma for which a putative cancer stem cell subpopulation has been isolated
Al Hajj and colleagues	This isolation
These cells	The cells they isolated
The stem/progenitor cell phenotype of these cells	The stem/progenitor cell phenotype of these cells
Mammospheres	
The vast majority of these cells	The vast majority of these cells

In the original, some of the subjects had no clear relation to any of the others, while in our revision, every subject clearly relates to that of the preceding sentence. Thus, revising to make the subjects of sentences related would yield a draft that reads more smoothly, but this would not make the relevance of new pieces of information in the later parts of sentences immediately clear. This approach would best be combined with identifying every new element and placing it in the position that makes the most sense.

→ *To revise for reader ease, can also find subjects related to those of previous sentences*

Revision isn't the only way to come up with writing that a reader can follow easily; you can also do this as you create the draft. If you remain aware of when you introduce new information, it's easier to do so in a way that's reader-friendly. Rather than letting new terms and players plop in whenever you think of them, consider how you can frame them in terms of something you've already discussed. Composing each sentence in this careful manner will reduce the amount of revision you'll need, but getting the most polished text you can will require re-considering how you introduced each new piece. This will especially help with those pieces that relate to multiple parts of the story, but whose relevance is most clear in only some of them (e.g. the marker expression in the example).

→ *As you draft, introduce new terms in relation to what you've already written*

Consistent terms

Helping readers follow your writing not only involves careful placement of where you explain the relationships of new information to your story, but also careful wording of those relationships. The words you choose to refer back to previously discussed material matter because some more clearly show the relationship than others. This is true not only for introducing new things, but also for other instances in which you intend for the reader to connect different parts of the text. Consider this (altered) example:

Example 2 version 1: relationships among terms unexplained

Numerous studies have described the orexigenic action of anandamide and 2-arachidon-

oylglycerol (1↓; 2↓; 3↓; 4↓).Based on these findings, an abundance of synthetic compounds also have been synthesized to interfere with cannabinoid CB_1 transmission in attempts to exploit the therapeutic potential offered by targeting this diverse neurotransmitter system. For example, rimonabant has acute central effects on appetite and continuing actions on body weight probably via peripheral interaction with lipid mobilization pathways in white adipose tissue and with cellular glucose uptake systems (5↓; 6↓). However, studies assessing the behavioral satiety sequence after either rimonabant or its derivative, AM251, reductions in feeding have been associated with off-target actions (probably opioid mediated) leading to excessive scratching (7↓,8↓).
From Dodd GT et al., "The Peptide Hemopressin Acts through CB_1 Cannabinoid Receptors to Reduce Food Intake in Rats and Mice," *J Neurosci* May, 2010, 30(21):7369-7376.

This paragraph likely left you feeling as though you missed something—you can see what the authors are going for, but you don't quite get there with them. For example, the second sentence indicates that it follows from the first ("based on these findings"), but it's hard to see how—what does "cannabinoid CB1 transmission" have to do with "anandamide and 2-arachidonoylglycerol"? Though the context suggests that rimonabant would be expected to reduce feeding, we would feel a lot more certain of our understanding if we knew what "orexigenic" meant. From this reading, we see the importance of which terms the author uses to how easy it is to follow a text—it would be much easier to get the point of this paragraph (and be sure we get it) if the actions of the drugs were described in the same terms throughout.

→ *Use the same words to describe related ideas*

Contrast the above version with the original (with a few minor revisions to better connect terms):

Example 2 version 2: terms related to one another

Numerous studies have shown that lipid-based *endogenous CB_1 agonists*, such as anandamide and 2-arachidonoylglycerol, *increase feeding behavior* and appetite (1↓; 2↓; 3↓; 4↓). An abundance of synthetic compounds also have been synthesized to interfere with cannabinoid CB_1 transmission in attempts to *develop therapies for obesity that would decrease feeding*. For example, *the CB_1 antagonist* rimonabant has acute central effects on appetite and continuing actions on body weight probably via peripheral interaction with lipid mobilization pathways in white adipose tissue and with cellular glucose uptake systems (5↓; 6↓). However, in this and in previous studies assessing the behavioral satiety sequence after either rimonabant or its derivative, AM251, reductions in feeding have been associated with off-target actions (probably opioid mediated) leading to excessive scratching (7↓,8↓).

The clearer version always refers to the drugs described in terms of their action on CB1 receptors and always describes their effects on behavior in terms of feeding and appetite. Since the terms are the same, it's easier to see the connections among the sentences—we know for certain that rimonabant is one of the drugs mentioned in the second sentence, which now announces the direction in which they affect

feeding, so we can see that the "reductions in feeding" in the last sentence are definitely expected.

→ *Relate new terms to those previously used*

Using the same terms throughout your paper likely seems either unnecessary, or to conflict with general guidelines for good writing. First, if you've used the term before (and the example is from the discussion, so the authors may well have done so), you might expect the reader to know what the term means and thus see the connections to related terms. However, if the reader isn't familiar with your field, reading the definition once probably isn't enough for her to automatically interpret the term every time she reads it later. Thus, the reader might have to go back and find the definition, which interrupts her progress through the text. Further, even those who know all of your terms will find it easier to see connections that are explicitly stated by identical terms than those implied by related terms. A second possible objection to using the same terms is that this would reduce the diversity of your vocabulary, and you've probably been encouraged in the past to vary the words you use in a paper to make your writing more interesting. However, using multiple terms suggests to your reader that each term means something slightly different; he assumes that you use your terms precisely. If you don't actually mean something different, this complexity distracts from the ideas you're trying to convey (the reader spends unnecessary energy trying to determine how the instance in which you use "orexigenic" differs from that in which you use "increase feeding"), which are complex enough.

→ *Making your writing easy to follow is more important than variety in word choice*