

# Chapter 6

## Achieving Clarity and Conciseness



### 6.1 Labeling Rather Than Explaining

We love our technical terms. We’ve studied them; we learn them, and now while writing a manuscript, we finally have a chance to use them! Right? Well, actually not.

Labeling is shorthand for the full development of an idea, but people often have a different idea of exactly what that shorthand actually means. The same term often carries different meaning to people with different disciplinary backgrounds. This makes using these terms a barrier to clear communication.

Strive to explain exactly what you did. Do not label it. The more specific you are about exactly what you did, the easier it is for someone else to understand it. If a methods section reads, “For the hospital catchment area survey, we selected 20 unions, using a probability proportional to size sampling approach,” a reader may wonder, What is a probability proportional to size sampling approach? How did you apply this concept to your site? The methods section should report methods in sufficient detail so that other investigators could repeat them. So skip the label, and instead use the space to describe the steps you took to identify and enroll the population.

The three most common labeling issues in papers concern study design, sampling methods, and limitations.

Examples of the error	Alternative, better options
✗ The population of the catchment area was projected for 2008 on the basis of the 2001 Bangladesh census using population estimation by component method.	✓ We began with the 2001 Bangladesh census of subdistrict populations and applied national estimates of crude birth rate, net external migration, and national crude death rate (ref) to estimate the population in the subdistricts at the time of the assessment.

Examples of the error		Alternative, better options	
X	Confounding by wealth is a potential limitation.	✓	Households who had windows that provided cross ventilation may have been wealthier and possessed other characteristics that improved their children's health that we were unable to completely control for in the analysis.

## 6.2 Using Weak Opening Phrases for Sentences

You should try to use phrases and transitions that advance the ideas and arguments of the paper. By contrast, most of the phrases below reflect the psychological state of either the reader or the writer. Strive to write from the perspective of the ideas you are developing. You are better off having no transition than using such vacuous phrases as the examples below:

Examples of the error		Alternative, better options	
X	It was found out that...	✓	<i>Delete</i>
X	One important observation from the findings of this study was that...	✓	<i>Delete</i>
X	We conclude from our data...	✓	<i>Delete</i>
X	Moreover, our survey showed that...	✓	<i>Delete</i>
X	Therefore, this will not be an overstatement that...	✓	<i>Delete</i>
X	It is known that...	✓	<i>Delete</i>
X	It can be seen from the above table that...	✓	<i>Describe</i>
X	The explanation could be that...	✓	<i>Explain</i>

## 6.3 Using Adjectives and Qualifiers

Adjectives are words that modify a noun. Adjectives often imply substantial subjective and emotional content, both of which should be minimized in conventional scientific writing. For example, what is “important” or “large” to one person may not be “important” or “large” to another.

Qualifiers are words that modify an adjective, but do not carry a specific meaning, such as “very.” The addition of a qualifier adds to the subjectivity, as in “very important.” It is better to try to choose the best adjective, and provide justification of its use, and not to use a qualifier.

Examples of the error		Alternative, better options	
X	The outbreak caused very high mortality.	✓	56% (301/536) of people infected in the outbreak died.
X	This very large outbreak.	✓	This outbreak affected 300 school children.

Examples of the error		Alternative, better options	
X	The incidence was much higher in children <5.	✓	The incidence among children <5 was six times higher than older children.

## 6.4 Overusing Studies or Authors as Sentence Subjects

In general, when referring to other scientific work, the subject of the sentence should not be the study, or the study's author, but the core ideas or results that connect to your manuscript. Ideas and observations referenced from other studies are central to scientific reasoning. The use of a study or a study's author as the subject of a sentence risks distracting the reader from the substance that links to the author's own study. The structure of your sentences should reflect this prioritization of ideas and results over individuals and authors.

Example of the error		Alternative, better option	
X	A study by Yoruba in Tanzania suggested that 78% of the clients who presented to traditional healers were females, 95% of whom were illiterate and of a low socioeconomic group (ref).	✓	Demographic parameters are important because they may influence health-seeking behavior; a study in Tanzania, for example, found that educated mothers are more likely to discourage traditional healing practices (ref).
X	Curtis et al. have championed structured observation as the preferred approach to measuring handwashing (ref).	✓	Using structured observation to assess handwashing behavior has consistently identified lower frequencies of handwashing with soap compared with reported behaviors.

## 6.5 Using Nondescriptive Numeric or Alphabetical Labels

Study teams commonly develop some study-specific vocabulary, for example, Group A and Group B and Phase 1 and Phase 2. The study team becomes so familiar with these labels and their underlying characteristics that they use these labels in everyday conversation within the study team. Unsurprisingly, when team members start writing about the study, they commonly use these same labels.

However, such labels are inappropriate for a scientific document intended for readers outside of the study team. Such nondescriptive numeric or alphabetic labels require readers to learn the study team's private code, or not be quite sure what all the differences reported by encoded classification actually represent. The team's private code is broadly useless information not applicable to any other manuscript the reader will ever encounter. Avoid needlessly frittering away your readers' attention. Make your paper as easy to understand as possible. Use descriptive labels for each group.

This is particularly important when constructing figures. Figures should clearly label groups being compared without requiring the reader to go back into the methods section or other aspects of the paper to decode the key characteristics that distinguish the groups.

Examples of the error		Alternative, better options	
✗	At baseline, group 1 participants were somewhat less likely to own a television than group 2.	✓	At baseline, participants enrolled from Tongi were less likely to own a television than participants enrolled from Narshindi.
✗	Compared with baseline, contraceptive prevalence increased by 12% in phase 1 and by another 10% in phase 2.	✓	Compared with baseline, contraceptive prevalence increased by 12% after steps to improve supplies in clinics and by another 10% after community outreach.
✗	Category A symptoms included cough and difficulty breathing, while category B symptoms included diarrhea and vomiting.	✓	Respiratory symptoms included cough and difficulty breathing. Gastrointestinal symptoms included diarrhea and vomiting.

## 6.6 Using Respectively

The word “respectively” is an adverb meaning “in the order given.” Although it is commonly deployed in scientific writing to summarize results in few words, it is best avoided. A sentence ending with “respectively” requires the reader to go backward, reread, and mentally connect words and numbers that are physically disparate on the page. This extra effort risks interrupting the reader’s engagement with the flow of your manuscript’s narrative. Strive to make your sentences easy to read and understand without backtracking.

Examples of the error		Alternative, better options	
✗	Of the plasmodium-positive children, 17 (4%) and 9 (2%) were positive for <i>P. falciparum</i> and <i>P. vivax</i> , respectively.	✓	Of the smear-positive children, 17 (4%) had <i>P. falciparum</i> and 9 (2%) had <i>P. vivax</i> .
✗	Attack rates for any postoperative infection between the suspected outbreak period January and December 1996 and for comparison period June and December 1995 were 14% (10/72) and 6% (2/31), respectively.	✓	The attack rate for any postoperative infection between the suspected outbreak period January to December 1996 was 14% (10/72) compared with 6% (2/31) between June and December 1995.
✗	Household size differed in rural, peri-urban, and urban communities with a mean of 6.4, 4.2, and 3.2 members, respectively.	✓	Mean household size differed in rural (6.4), peri-urban (4.2), and urban (3.2) communities.

## 6.7 Using the Word Etcetera

Scientific writing is characterized by precision. “Etcetera” is not specific. This imprecision suggests that the author’s ideas have not been fully formulated or have not been fully thought through. “Etcetera” should never appear in a scientific concept paper, protocol, or manuscript.

Example of the error	Alternative, better option
<p>✗ Medical costs in the hospital included admission fees, bed rent, diagnostic tests, medicine, consultation fees, etc. Nonmedical costs included travel, food, tips, etc.</p>	<p>✓ Medical costs in the hospital included admission fees, bed rent, diagnostic tests, medicine, and consultation fees. Nonmedical costs included travel, food, and tips.</p>

## 6.8 Using a Non-English Word as an English Word

Most scientific articles aim for international readership. Words and expressions that are specific to the country where the work was conducted risk confusing international readers. In English language scientific articles, words from other languages should be *italicized and placed in parenthesis*.

Example of the error	Alternative, better option
<p>✗ We conducted a case-control study in two upazilas in Rajshahi district.</p>	<p>✓ We conducted a case-control study in two subdistricts (<i>upazilas</i>) in Rajshahi district.</p>
<p>✗ Anthropologists interviewed gaccis in each village.</p>	<p>✓ Anthropologists interviewed date palm sap collectors (<i>gaccis</i>) in each village.</p>
<p>✗ The questionnaire was translated into Bangla.</p>	<p>✓ The questionnaire was translated into Bengali.  <i>Note: Bangla</i> is not an English word. The English language word for the language spoken in Bangladesh is Bengali (not italicized). When writing about questionnaires in Latin America, scientists do not use the Spanish word for the Spanish language (<i>español</i>). They do not write that the questionnaires were translated into <i>español</i>. Instead, they write that the questionnaires were translated into Spanish. Similarly, when writing in English about work in Bangladesh, we should refer to the local language as Bengali.</p>

## 6.9 Describing Costs Only in Local Currency

International readers of scientific manuscripts are unlikely to be familiar with the value of local currency. They will have difficulty interpreting costs described in local currency.

The US dollar is the most widely used currency globally and so provides a useful metric to communicate cost to international readers. Importantly, many scientific articles aim to reach both an international scientific audience and a local audience, and so it may be useful to present costs in both currencies.

At a minimum, include an appropriate conversion (the one prevailing at the time data was collected) between the local currency and an international currency so that readers can connect the currency in your paper to amounts of money they understand.

Examples of the error		Alternative, better options	
✗	The mean monthly household income in the study communities was 4916 Pakistan rupees.	✓	The mean monthly household income in the study communities was US\$ 370 (4916 Pakistan rupees).
✗	The cost per fully treated patient at a government hospital in Mexico City was 3200 MXN.	✓	The cost per fully treated patient at a government hospital in Mexico City was US\$ 162 (3200 MXN). <i>Note:</i> In the methods section include the exchange rate (1 US\$ = 19.8 Mexican pesos (MXN)) and the date the currency was converted on (June 7, 2021).

## 6.10 Using the Term “Developing Country”

The term “developing country” is nonstandard and imprecise. All countries are developing. Belgium was a different country in 2020 than it was in 2000. It had a higher income and a greater number of internet connections. It was developing. Belgium will look different in 2044 than it did in 2020. It will develop further. Although the term “developing” historically connotes industrial development, there is no standard definition that can be consistently applied to classify countries as a “developing country” or not.

By contrast, the World Bank has clear standards for characterizing countries by income level. It assigns the world’s economies to four income groups—low, lower-middle, upper-middle, and high-income countries. The classifications are updated each year on July 1 based on gross national income (GNI) per capita in current USD (using the World Bank Atlas method exchange rates) of the previous year.

## 6.11 Using the Term “Socioeconomic Status” as a Synonym for Wealth

When referring to income or poverty/wealth among persons, households, or communities, many writers mistakenly use the term socioeconomic status. If the available measurements are strictly measurements of wealth or income, for example, household assets, then use terms that refer to this narrower concept. For example, wealth, income, or poverty level. Socioeconomic status and wealth are not synonyms. The concept of socioeconomic status captures more than just wealth. It encompasses income, education, and profession and also includes the idea of social class. Restrict the use of the term socioeconomic status only when the available data supports this broader conceptualization.

## 6.12 Using a Technical Term in Its Nontechnical Sense

Several technical scientific terms carry a less specific meaning when used in general speech. To avoid confusing the reader, avoid using technical terms in their nontechnical sense.

### 6.12.1 Using the Term “Random” in Its Nontechnical Sense

The term “random” has a specific technical meaning within science. Random selection, for example, implies that the entire population is enumerated and that a process, such as a lottery or a random number generator, is used to select individuals from among the entire population. In a scientific manuscript, the word “random” should only be used within this specific context.

In common speech, the word “random” is often used as a synonym for “haphazard.” For example, “I was walking down the street and selected a restaurant for lunch at random.” To a scientist, this was not random selection of a restaurant. Rather, the choice of lunch location was based on convenience.

Example of the error	Alternative, better option
<p>✗ In-depth interviews were conducted with 10 randomly selected key informants working in health centers.</p>	<p>✓ We conducted in-depth interviews among 10 key informants we identified who worked in health centers in the study communities.</p>

### 6.12.2 Using the Term “Reliable” in Its Nontechnical Sense

The term “reliable” has a specific technical scientific meaning that is narrower than its meaning in common speech. Within science, “reliability” refers to whether repeated measurements of the same phenomenon are similar. A blood test is reliable if it provides the same result on repeated testing of the same sample. The synonym for “reliability” in this technical sense is “repeatability.” To avoid confusing your scientific reader, the words “reliable” and “reliability” should only be used in their strict technical sense in any scientific document.

Example of the error		Alternative, better option	
X	The self-reported data may not be reliable.	✓	The self-reported data may not be valid.
X	The direct observations were conducted to cross-check the responses and ensure reliability of the data collected in the self-administered survey.	✓	We cross-checked the findings from the self-administered survey by comparing them with results from direct observation.

### 6.12.3 Using the Term “Significant” in Its Nontechnical Sense

The term “significance” has a specific technical meaning in quantitative scientific writing. Specifically, it refers to statistical associations that are less likely than would be expected by chance. Conventionally, these are associations with a probability of occurring by chance of less than 5%. Many thoughtful commentators on scientific writing are critical of this narrow dichotomous conceptualization that divides all results into “significant” or “not significant” (see Error 7.1.1). Despite this criticism, when scientific readers see the term “significant” in a scientific manuscript, they will assume the author is referring to statistical significance. Avoid using the term in a different context or you risk confusing the reader.

Example of the error		Alternative, better option	
X	A significant number of respondents could not identify common signs of H5N1 in poultry (Table 2).	✓	Most respondents could not identify common signs of H5N1 in poultry (Table 2).
X	Backyard poultry can be a significant source of high-quality protein for rural low-income families.	✓	Backyard poultry can be an important source of high-quality protein for rural low-income families.

### 6.12.4 Using the Term “Valid” in Its Nontechnical Sense

The term “valid” has two related technical meanings in quantitative scientific writing. When used to describe a measurement, it implies that the measurement reflects the underlying phenomenon of interest and is not an artifact of the instrument being used for measurement or other cause of inaccuracy.



When used to describe a scientific inference, the term *valid* implies that the inference is sound given the results and the way the data were collected. The term is used more loosely in general communication. To avoid confusing readers in scientific manuscripts, only use the term in its technical sense.

Example of the error		Alternative, better option	
X	Preventing nosocomial transmission of tuberculosis is especially valid in Bangladesh because of its high tuberculosis burden.	✓	Preventing nosocomial transmission of tuberculosis is important in Bangladesh because of its high tuberculosis burden.
X	The similarity of results from the repeated assessment of the samples suggests that the assay is valid.	✓	The similarity of results from the repeated assessment of the samples suggests that the assay is reliable.

### 6.12.5 Using the Term “Incidence” Incorrectly

Epidemiologists define incidence as the number of new cases of illness that occur in a specified population in a specified time. For example, the incidence of hepatitis B in the population was 23 cases per 10,000 persons per year. The numerator for incidence is a count of new cases (or new events). The denominator is person-time, that is, a measure that captures both population size and time. Because time is in the denominator, incidence is always a rate. Thus, the second word of the phrase “incidence rate” is redundant.

Prevalence, by contrast, is the number of cases in a population. It includes both new cases and old cases. For example, there may be 400 cases of hepatitis B in the same population of 10,000 people. Most of these cases are old cases. The prevalence of hepatitis B in the population is 4%.

Reporting incidence as an unqualified percentage is incorrect because it does not communicate the time frame that the new cases occurred.

Example of the error		Alternative, better option	
X	We followed a cohort of live poultry market workers in Bangladesh to determine the seroprevalence and incidence rate of seroconversion of antibodies to H5N1 virus.	✓	We followed a cohort of live poultry market workers in Bangladesh to determine the seroprevalence and incidence of seroconversion of antibodies to H5N1 virus.
X	The incidence of diabetes among Marin County residents, 5%, is the lowest in the state.	✓	The prevalence of diabetes among Marin County residents, 5%, is the lowest in the state.

### 6.12.6 Using the Term “Correlated” Incorrectly

In statistics, the term “correlated” implies that there is a statistical relationship between two continuous variables. In common speech, the term correlation often implies any sort of statistical association. In scientific writing, only use correlated when it is technically accurate.

Example of the error		Alternative, better option	
✗	Cross-sectional quantitative studies have found that higher trust was correlated with increased compliance with Ebola control measures.	✓	Cross-sectional quantitative studies have found that higher trust was associated with increased compliance with Ebola control measures.
✗	Consistent correlations were noted between ethnicity and a variety of health outcomes.	✓	Ethnicity was associated with a variety of health outcomes.

### 6.13 Using the Term “Documented”

The word “document” is a noun. English often turns nouns into verbs, but not always with good results. To “document” means to make a document, that is, to write something down. So if you write down on a piece of paper the phrase “the earth is flat,” then, strictly speaking, you have documented that the earth is flat. Creating a document is unrelated to the validity of an assertion. Therefore, we should not use this verb to communicate scientific validity of a statement.

Example of the error		Alternative, better option	
✗	Studies in Bangladesh, India, and Malaysia also documented neutralizing antibodies against Nipah virus in <i>Pteropus</i> bats.	✓	Studies in Bangladesh, India, and Malaysia also identified neutralizing antibodies against Nipah virus in <i>Pteropus</i> bats.

### 6.14 Framing an Argument in Terms of Need

Quite often, arguments in draft scientific papers are framed in terms of needs. The underlying message is that we “need” to do something. Usually, the authors are asking the reader, the government, or society more generally to care about the issue in the same way that the authors care about the issue and follow the specific advice of the authors.

In a scientific manuscript, it is reasonable to talk about a need for water, oxygen, and food for survival, but it is less appropriate to assert a need for health-care reform or a need for social change. The problem with this language is that it disguises the

goals and aspirations of the authors in terms of a need when the issue of what constitutes a legitimate need is an open question for individuals, for society, and for science.

Scientific writing is most persuasive when it demonstrates the connection between a set of conditions and consequences. Rather than framing arguments in terms of needs, the same ideas should be described as steps that are required to achieve a particular outcome. Importantly, the outcome should be specifically stated.

Examples of the error		Alternative, better options	
X	There is a need to standardize and expedite the assignment of causes of death, thereby enhancing a timely process of appropriate decision-making.	✓	If the assignment of causes of death could be standardized, appropriate decision-making based on these data could be expedited.
X	A low-cost, accurate approach to characterize handwashing behavior is needed.	✓	A low-cost, accurate approach to characterize handwashing behavior would improve the assessment of handwashing promotion programs.

## 6.15 Using the Term “Illiterate” as a Synonym for “No Formal Education”

Although we often use the word “illiterate” as a synonym for “no formal education,” these terms are not synonymous. Literacy can be evaluated by asking people if they can read or write and validated using specific literacy tests. People may have attended school for some years and still not be able to read or write. People who have not attended formal schooling are unlikely to be able to read and write, but it is more precise to characterize their lack of education rather than their literacy skills. The term illiterate also sometimes carries a condescending tone and so risks communicating a lack of respect for one’s study subjects.

Examples of the error		Alternative, better options	
X	The age range of program beneficiaries was 18–65 years old, and over 25% who took part in activities were illiterate.	✓	The age range of program beneficiaries was 18–65 years old, and over 25% who took part in activities had less than 4 years of schooling.
X	Educated mothers were 2.3 times more likely to have a handwashing station with soap and water than illiterate mothers.	✓	Educated mothers were 2.3 times more likely to have a handwashing station with soap and water than those with no schooling.

## 6.16 Using the Word “Challenging” as a Synonym for “Difficult”

We often use the word difficult to describe public health problems or solutions. The word is appropriate as major problems are characteristically complex and defy simple solutions. The word challenging is often used as a synonym for difficult, but challenging carries a different connotation. The root noun of the adjective challenging is challenge. The connotation is that the situation is testing us; that by engaging in this issue our capacity to take on new issues and to grow to address these issues is revealed. When a situation is difficult, motivational coaches encourage us to see this difficulty as a personal challenge so that we can strive to overcome it.

This implicit motivational jargon is out of place in scientific writing that values precise description. The substitution of challenging as a synonym for difficult is so overused that it risks sounding insincere. If the situation is difficult, then call it difficult. If you want to challenge a group, in an editorial or in the discussion section, then do so explicitly.

Examples of the error		Alternative, better options	
✗	We will explore challenges in implementation, as well as find out what factors motivate children to participate.	✓	We will explore difficulties in implementation, as well as find out what factors motivate children to participate.
✗	In these impoverished contexts, changing child feeding behavior is challenging.	✓	Poverty is a major barrier to improving child-feeding behavior.
✗	These modest findings highlight the challenges of maintaining high-quality implementation of interventions at scale.	✓	These modest findings highlight the difficulties of maintaining high-quality implementation of interventions at scale.

## 6.17 Describing a Laboratory Test Result as Positive

Scientific communication is characterized by specificity and nuance. It avoids unqualified generalizations. Scientific thinking eschews narrow dichotomies, such as stating that an intervention was a success or failure. Instead, a scientific approach is more likely to identify aspects that achieved objectives and aspects that did not.

Scientific writing should bring this framing to our description of laboratory results. No laboratory test is 100% sensitive and 100% specific. A laboratory test provides additional information that scientists can interpret. When describing laboratory results, use sufficiently precise language so that readers can interpret the meaning without having to jump back to the methods section to review which laboratory tests were conducted and how they were interpreted.

Examples of the error		Alternative, better options	
✗	Out of 23 samples tested for different respiratory viruses, 21 were positive for respiratory syncytial virus.	✓	Out of 23 samples tested for different respiratory viruses, 21 had detectable RNA for respiratory syncytial virus.
✗	From the surveillance database, we identified 209 influenza-positive patients during May to October 2010.	✓	From the surveillance database, we identified 209 laboratory-confirmed influenza patients during May to October 2010.
✗	Among the 123 people tested, six were positive for Nipah.	✓	Among the 123 people tested, six had IgM antibodies against Nipah virus.

### 6.18 Using Increase or Decrease in the Absence of a Time Trend

The words increase or decrease imply a change in quantity over time. They should not be used when comparing two groups during the same time interval.

Example of the error		Alternative, better option	
✗	Children < 5 had an increased risk of infection compared with school-aged children.	✓	Children < 5 had a higher risk of infection compared with school-aged children.
✗	Children in the nutrition intervention group had a decreased prevalence of anemia compared with controls.	✓	Children in the nutrition intervention group were less likely to have anemia than controls.

The words increase and decrease can be used appropriately when evolution over time has occurred. For example, the incidence of anemia decreased between 2003 and 2015.

### 6.19 Describing a Test as a Gold Standard

The phrase “gold standard” has a precise meaning in economic history, but this overused phrase is too imprecise for scientific communication. Most commonly, authors use it in a context when a laboratory test yields few false positives. However, in many situations, the errors generated by false-negative tests are as equally misleading and harmful as false-positive tests.

The term “gold standard test” implies an argument from authority that the authors used the best test. Arguments of authority are received skeptically by scientists (Error 2.3.3). All tests have advantages and disadvantages. All tests require thoughtful interpretation. When discussing the use of a specific test or a comparison between two tests, communication is improved by specifying the particular characteristics that are being compared and contrasted.

	Example of the error		Alternative, better option
✗	Enteric fever surveillance is often based on blood culture, the current gold standard diagnostic test for enteric fever.	✓	Enteric fever surveillance is often based on blood culture because the absence of false-positive results provides confidence that each identified case is a confirmed case.
✗	We compared the seroprevalence of the total reported positive tests in the area to understand the level of underreporting from gold standard RT-PCR testing.	✓	We compared the seroprevalence of IgG antibodies against SARS-CoV-2 to the government reports of respiratory specimens from residents of these communities who had SARS-CoV-2 RNA detected by RT-PCR testing.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

