Scientific writing for geologists
- The very basics -

Steffen Büttner
Rhodes University – Department of Geology

This is a document that is aimed to support geology students in writing scientific documents such as project or field reports, or theses and publication manuscripts. The current version is not yet complete and many comments and recommendations will be added in the forthcoming versions that will be posted. Students and colleagues are encouraged to provide me with feedback, suggestions, corrections or additional material. The current version is biased towards problems I have come across while reading student field reports or during manuscript review. These include recurring typos and misspellings, strategic mistakes in the organisation of the report and formal mistakes, such as poor illustration and documentation of data. However, this guide may help students to avoid common mistakes in the organisation and writing of other scientific texts such as Honours theses or, at a higher postgraduate level, first manuscripts for publication.

1. Organisational matters

Before starting to write a scientific text it is necessary to have a very good idea what the purpose and contents, and the focus of the text is meant to be. Documents that are designed while they are written end up as incoherent and pointless. It is very important to align the Introduction and descriptive part to the purpose of the text. Think about what you want to say before you conceptualise the document. For instance, in the Discussion you only can discuss what you have a) presented to the reader as the purpose of the study in the Introduction and b) what you have backed up with observations and data in the descriptive part.

Do not make the mistake to throw at the reader everything you know, have seen, or read somewhere. Select carefully the relevant information that the reader needs to know in order to follow your argument and that you need to prove your point. This will shorten your document and make it more readable.

Abstract

Abstracts concisely point out the most important findings of a study and highlight why they are important or what they prove. An Abstract is NOT a piece of text that announces the content of the report/the study/the paper. Tell the reader what you have found out, not what your text contains. The Abstract is the last section you will write. Keep the overlap in content with the Conclusions (see below) as small as possible.

Introduction

The Introduction, you might have guessed, introduces the report/paper/thesis etc. Its purpose is to provide the reader with the background of the study, which often is a kind of a geological overview. Make sure that you cover only material that is relevant to the content of your study. Many authors believe that the Introduction is a literature review. This is by no means the case. Its purpose is to introduce the motivation for your study and provide the reader with the geological context that is necessary to understand the relevance and the context of the results. Of course you have to refer to the relevant literature wherever you use material from other sources. Typically, in the Introduction and in the Discussion sections you will cite more literature than in those sections that present your data. You need to apply your mind in order to decide what is necessary and what isn't (see below). In many geological studies, particularly those with a field component, a geological overview map will support and illustrate the information presented in the Introduction.

Furthermore, you have to give the reader a rundown on the structure and organisation of the report. What problem do you want to solve, what features are you going to describe, what kind of data will you present, and, in very general terms, what you are going to present as a result (e.g. *"On the basis of new field observation and structural analysis, this study challenges the view proposed by Drinkalot and Neversober (1996), who suggest that shallow-dipping thrusts predate south-dipping reverse faults. Our study also identifies the relative age of flexural slip folds and normal faults with respect to contractional faulting"*).

After reading the Introduction, the reader should have a very good idea what problem you want to solve, in which context it is important, how you are going to tackle it, what kind of data you will present, and, in qualitative terms, what result you will present at the end.

You normally cannot write the final version of the Introduction before you have produced at least a good draft of the descriptive part and the Discussion, because only then you know what you really have found out and what related information you have to present in the Introduction. How should you be able to properly introduce a study does not yet exist in its main parts? However, you obviously need to *know* about the relevant scientific background before you process and interpret your data. Frequently this means that you write a draft version of the Introduction quite early on, but you **have to revise, and perhaps rewrite it** once your study is completed. This is important, because the study needs to correspond in content and form to what you have announced and explained in the Introduction.

Data section/descriptive section

In many studies this section is called "Results". In geosciences this is not always adequate, because first you have to present descriptions of rocks or structures (i.e., observations) before you come to the results of further examinations and investigations, for instance chemical analyses. In the data/results sections the number of citations drops significantly, because here you do not use much of other people's material (who cares what other people have found when you describe your own data?). You focus on your own observation. Here is the right place to use figures (photographs, diagrams, sketches, stereonets, maps, etc.) and data tables.

In this section it is important to find a good balance between description and interpretation. Strictly speaking, the two always need to be separated. However, this is not always practical, and most reviewers/editors/lecturers accept limited interpretation of observations. This applies for instance to the interpretation of the shear sense (i.e. hangingwall transport to the south along a normal fault) right after describing the orientation and nature of the shear plane and the lineation, and the type and geometry of the shear sense indicator. It would *not* apply to the meaning of such tectonics in the regional or orogen scale (such as the possible relationship of the normal fault with the break-up of Gondwana). This comes later in the Discussion.

Another possibility is to add a summary and brief interpretation describing a coherent set of observations, such as observations made in a specific exposure. It is, however, necessary to clearly separate the description of geological features from their interpretation. Start a new paragraph and indicate the beginning of the summary/interpretation (e.g.: *"The observations x, y and z in this exposure suggest contractional tectonics in presence of hydrothermal fluids. This stage was followed by extensional tectonics forming crosscutting normal faults which do not show evidence of syn- or post-tectonic fluid influx. Contractional and extensional faults crosscut fold hinges and therefore post-date folding"*).

As a general rule, observations can be interpreted in the descriptive part as far as such interpretations are immediately possible and obvious from the description. Such interpretations do not extend into their significance or meaning to a larger context. Interpretations in the descriptive sections are usually acceptable as long they do not require regional (or any other) background knowledge that you would have to explain to the reader.

NB: This is the utmost you can do in the descriptive part. You must not burn material that you want to make part of the Discussion section. Therefore, refrain from references to regional processes or comparisons with features seen in other exposures, or the significance of observations in the context of general geological concepts.

Descriptions need to be concise and to the point. Describing field evidence, the reader wants to know where the geological feature is exposed, how it looks like, how often it occurs or how far it can be traced in the field, its size, variability, and what associated features are present. This requires frequently the indication of the location or region of occurrence on a map/aerial photograph/satellite image. Furthermore, you need to illustrate the features adequately. In structural geology this usually means the provision of line drawings and 3D sketches with adequate annotations, such as compass readings for the structures shown, labels for the identified structures, arrows for shear senses, orientation of the outcrop face, scale, etc. Stereonets might illustrate the size of your data set.

The data part is also the section where processed data are presented. Hence, if you have data that need or allow more in-depth assessment, do it in the data section, perhaps in a subsection, if appropriate. In structural geology you would, for instance, plot relevant data on stereonets in order to determine hinge line or axial surface orientations of folds. In metamorphic petrology studies you might include a section on thermobarometry with P-T diagrams in the Results chapter.

Discussion

If you have written a proper descriptive section, the Discussion is usually relatively simple to write. You just put the observations and data presented in the previous section into the context that you have outlined in the Introduction (or in the draft version you already have). In texts on structural features or metamorphic geology this frequently means that you work out the succession of tectonic and metamorphic events that you have described, the conditions under which they have occurred, and their relative or absolute timing. **Most important in this section is that you provide clear scientific reasoning for your interpretations. Just stating your thoughts or opinion is not enough. Reasoning, argument and making meaning of your data and observations is the core purpose of the discussion section. This also sets it apart from any other form of text.** You will point out alternative interpretations that might be suitable to explain your observations and discuss their pros and cons. You do not necessarily have to decide for the one or the other solution if several solutions are equally likely, but you must explain this to the reader. Provide proper reasoning on which your decision is based. The Discussion also will relate your findings to the results and interpretations in the literature, which you have described in the Introduction.

You are not allowed to present new data, evidence or background information in the Discussion. All such material has to be presented either in the data/results sections or in the Introduction. Modify these sections accordingly if necessary. It is highly important that the Discussion and Conclusion deliver what you have promised to deliver in the Introduction, and that you do not discuss anything that requires background information or context that is not provided in the Introduction. Discussion, data sections and the Introduction need to be complementary in content!

Conclusions

What have you worked out in the Discussion? Which problems have been solved and what is the solution to your problem (or the possible options); which new problems have emerged; what other research will have to be carried out? What do the results imply for the larger context? The Conclusions are the results of the Discussion, **with focus on the greater picture**. The Abstract, in contrast, focuses often on the smaller picture of the study, presenting for instance the most important new data that the study has produced. The Conclusions rather focus on the implications of the results in the regional (or topical) context. However, a certain overlap between Abstract and Conclusions is often unavoidable.

2. Style and scientific issues

Don't claim anything as a fact that you have not supported with evidence. Therefore, do not use terms like *indicate, prove, is/are, or demonstrate,* if you only present some observations that just might *be in agreement* with what you want to conclude. If you refer to interpretations or assumptions, rather use *suggest, may suggest, may be related to, is in agreement with, etc*. The latter also applies to interpretations made on circumstantial evidence or common/coeval occurrence of different, but possibly related features. There is nothing wrong with observations that are inconclusive. Several independent observations suggesting the same interpretation can carry enough weight to convince the reader.

Descriptions of rocks and structures are always done in the **present tense**, not in the past tense. Phrases like *"The rock showed garnet and vesuvianite"* makes the reader wonder whether this might have changed since the observation was made (you can safely assume that it has not!): "*The rock **shows** vesuvianite and garnet".*

Quantification: *A lot, many, a large number, dramatic amounts,* are not scientific categories. Avoid using these terms. A lot or many or a large number has different meanings for different readers. Referring to features occurring more or less often in your study area you might use terms like *abundant* (you find it several times in each exposure), *common* (present in more or less all examined exposures), *occasionally* (in a significant minority of exposures), *rare* (you saw it only once or twice, or generally as an exception). If you describe only one location you will have to adapt the meaning of terms accordingly. The best way to avoid hassle is to quantify observations numerically: *"Shear sense criteria indicating hangingwall transport to the north have been seen in six localities (locs.3, 5, 7, 12, 17 and 22)".*

Similar problems arise when authors wish to underpin the **intensity or importance of a feature.** Terms like *extremely, very, highly*, etc. are ambiguous and should be avoided. Just describe the features that are present and try to quantify their abundance. Instead of writing *"Rocks along the fault zone are very (or strongly) fragmented"* you might write: *"Brittle failure produced angular fragments of 0.5 to 2 cm in size throughout the fault zone".* Leave it to the reader to regard this as *very fragmented* or not.

Lack of precision in using correct terminology

A common and **very annoying** example is the practise to describe the orientation of folds in inappropriate terms, such as *" the folds are oriented east-west"*, or similar versions. What is oriented east-west? The hinge line? The dip direction of the axial surface? Or perhaps its strike? Instead, refer to specific fold elements*: "Folds in the area show a moderately to steeply south dipping axial surface and shallowly east or west plunging hinge lines"*. Similar problems occur in the description of surface and lineation orientation, when strike, dip angle, dip direction, trend and plunge are getting mixed up.

Dead wood

Dead wood terms are terms without own meaning or little information value. They often can be found in inexperienced authors' work and make their texts unnecessary long and unpleasant to read. Search your text, find, and eliminate dead wood. It is a prime enemy of the inexperienced author.

Common dead wood are expressions like "*cataclastic S-C fabrics are present*", or "*… are observed*". Rather add some useful information: "*Cataclastic S-C fabrics are developed along reverse faults and form zones of 10-30 cm in thickness"*. Other meaningless terms are: *most notable/it can be noted, namely, are seen, was noticed,* etc. Or: *"An investigation on xyz was conducted"*. Ugly passive voice paired with lack of content. Why not *"An investigation of xyz will evaluate abc"*, saying something useful beyond the fact that is clear anyway, that the investigation that you report on, yes, really, indeed, was conducted.

Or: "*As mentioned earlier, …*" is almost always superfluous. Either the reader still knows what follows and does not need a reminder; in this case (s)he does not need the patronising statement. Or the "earlier mentioned" matter needs to be revisited properly. The best way is usually just to state again what needs to be stated: for instance: "*The hinge line data have shown the dominant E-W trend of regional folds (Section xyz, Fig. xyz), and …*".

Really bad are content-free sentences such as *"In exposure xyz there are many interesting geological features".* Replace such terms or sentences with meaningful content or just delete them. Another common (and annoying) problem is the over usage of the term "*important*", without providing any reason for the feature's or the observation's importance. It often suggests that the author tries to impress the reader and at the same time hide his/her own lack of understanding.

Sentences and paragraphs: the need of coherence and clarity

One **sentence** should always deal with **one sort of material**. Otherwise it will be difficult to follow its content. For instance, *"… the quartzite shows ripples and cross-bedding as well as plumose structures and quartz veins"* is incoherent because it starts off with primary structures (ripples and cross-bedding) which get intermingled with information on secondary plumose structures and veins, which may have formed much later and under very different conditions. Rather carry on with describing primary features (grain size variations, layering, likely primary material deposited, etc.).

Secondary structures, such as joints, veins, or plumose structures that form much later, are described in a **new paragraph or section**. Starting a new paragraph tells the reader that a new type of information is about to be communicated, or, in the Discussion, a new thought is to be developed.

Appropriate subdivision of the text in coherent sentences, paragraphs, sections, and chapters are important tools to guide the reader through the document, and to make it understandable. A dangerous enemy in this regard is the runaway paragraph that extends over more than half a page and deals with a whole lot of different matters. Here clarity is largely lost, and often even repeated reading will not help the reader much to grasp the point the author wishes to make. Typically, a paragraph should not exceed a quarter of a page, preferably less. But also too short paragraphs (two-liners) can be problematic. Where possible they should be merged, extended, or deleted.

Paragraphs have their own internal structure. Beyond coherence it is necessary to set the topic of a paragraph at the beginning in the first sentence. For instance, if you wish to describe the field appearance of a specific rock type or lithological unit in this paragraph, tell the reader: *"The Witteberg quartzite shows different topography along the dominant river valleys compared to areas away from such valleys.* Then the reader knows what to expect and will be able to grasp easier what you want to communicate, and you may carry on with more detailed descriptions. For instance "P*rominent cliffs form along the deeply incised river valley (Fig. xx), but away from the valley the quartzite crops out on the top of rolling hills on the topographic high-ground"*.

At the end of longer paragraphs, you may add a summarising sentence or even a statement with some interpretative tendency (but not more). For instance: *"The prominent morphology of the Witteberg quartzites, with E-W oriented ridges follows the main regional anticlines and synclines and can be used as a tool for mapping and interpreting these structures".*

This shows that longer paragraphs should be structured in a similar way like the whole document, with a kind of introduction that sets the topic, a main part that communicates coherent information on that topic, and, optionally, a summary that tells the reader the essential information in abridged form.

Figures and tables

Figures are vital parts of many geological texts. Readers often browse through the figures in order to get an idea of the nature, quantity and, to a certain extent, quality of the data and observations. It is therefore worthwhile to make figures as informative and appealing to the reader as possible. Diagrams, line drawings, X-Y data plots, stereonets, photographs, **all are called "Figures"**. Nothing is more irritating than having to find out whether stereonet 5 comes before figure 6 or after photograph 3 while trying to follow a text. The use of appendices is discouraged for similar reasons (exception: geological maps that are too large to fit on an A4 page). Usually there is no good reason *not* to give all figures proper numbers and integrate them into the document.

Figures are consecutively numbered and adequately referred to in the text (see below). ALL FIGURES HAVE A FIGURE CAPTION! This figure caption is presented **below** the figure, and this figure caption needs to be **comprehensive enough to inform the reader about the content of the figure**. It also may supply a short *explanation or interpretation*. In professional scientific texts figure captions are often several lines long. Annotation ON the figure is often essential. Scale bars, compass directions, lines and arrows highlighting relevant features, all this contributes to clarity, makes important information accessible to the reader, and are therefore a must.

Tables do not have captions below, but **headings above**. They also are consecutively numbered and, as figures, referred to in the text. Below tables often technical explanations are presented, such as the meaning of symbols used in the tables or different sources of data contained in the table.

In the text you have to make proper use of the features shown in figures. A figure just attached in the appendix is worthless if the text does not refer to it and makes use of its content. This means you have to refer to a figure as soon as possible or appropriate. Do not write long-winding descriptions of a complex geological and tell the reader *at the end* that there is this high-quality 3D illustration in Figure 12, which makes all these explanations clear in a split-second. This would be an excellent way to really annoy the reader/reviewer/supervisor.

Length of the report/ manuscript/thesis

Some authors are under the impression that documents have to be long to be good. This is quite far from truth. There is no strict rule for the length of a scientific text other than: ***As short as possible, as long as necessary.*** Be concise and precise, but do not inflate a small data set in order to make it look like a big one. The readers of your text will spot that instantly and feel that you are wasting their precious time. Do not present material that is irrelevant to the problem you want to solve. Preferably present your data in illustrations and describe these illustrations in the text. On the other hand, do not leave it to the reader to make sense of your data. That is your job. To make this point clear: a good descriptive or data section of a scientific text is largely a concise but comprehensive description and explanation of features shown in illustrations or in tables.

The important and often neglected last step before final proofreading is the revision of your document in order to shorten and simplify the text and to weed out unnecessary material. For early career authors it might be a good idea to re-read writing guides before this last proofreading and editing cycle.

Formalities

The reference list needs to be tidy, complete and consistent in style. Follow strictly the recommended format of the journal or the institution where you want to publish/submit your work. For citations in the text check professional journal articles for appropriate style. Software such as endnote may be used for this task.

Sundries

- Insert **page numbers**
- In longer reports or theses use chapter-specific figure and table numbers (e.g., Fig. 4.1, Tab. 3.2). If you insert an additional figure or table you only have to re-number in one chapter.
- Subdivide the text in **chapters** and appropriate hierarchic **sections** and **subsections**, (1.1, 1.1.1, 1.1.2, etc.), but do not exceed three sub-levels.
- **Careful proofreading before handing in/submission is a must! (I really mean that; a poorly proofread text will irritate the reviewer or examiner and may well trigger harsher assessment of the scientific content. People tend to get picky when they feel that they are used to fix bad writing.)**
- If several conventions are commonly used, spell out which one you use in your text. For instance, when talking about the orientation of planar structures, make clear whether you refer to **strike** or to **dip-directions**.

Common typos and spelling mistakes

- Compass directions (**north, south, east, west**, etc.), **mineral names or rock names** are **NOT capitalised**. Neither are greenschist, amphibolite or granulite facies**.** Oddly, many people use correctly small letters for common minerals, like quartz or muscovite, but then capitalise the rare ones, like xenotime.
- The **Earth** if referred to as a planet *is* capitalised (not the earth as part of the soil or regolith profile). The same applies to the **Moon**, the one that orbits the Earth. If you refer to

unspecified moons of other planets, “moon” is not capitalised; but once you get specific and talk about Io, Ganymede, Europa, Callisto, or Titan, of course they have caps!

- Check out the following websites (or get yourself a book on English grammar): http://owl.english.purdue.edu/owl/ (Purdue Online Writing Lab) http://www.yourdictionary.com/grammar-rules/Grammar-Capitalization-Rules.html
- There is a *z* in quartz. And no *s*.
- Sillimanite is **not** spelled *silliminite* or *silimanite*.
- It is “the right way up”, not the *rite* way. Anyway, better use *normal layering* and *overturned layering*.
- Greenschist is one word.
- There might be closed pubs but no *closed* folds; these would be close folds.
- Nouns used in plural happily live **without an apostrophe before the s**. This is a VERY common mistake. Specifically, many people find that the plural of shale should be *shale’s* instead of shales. “Shale’s” would be correct only if you intend to describe one of its characteristics (e.g. *“ ... the shale’s colour varies from dark grey to …”*). Read up useful sources (e.g. Purdue Online Writing Lab) on the use of the apostrophe in English language. Normally use an apostrophe only:
  ☞ to form possessives of nouns
  ☞ to show the omission of letters

- Fine-grained, coarse-grained, etc., are compounds that refer to a noun (the rock or mineral species) and are **hyphened**.
- The symbol for million years is Ma, not ma (which would make it milli-years)
- It’s the Witteberg Group, not the *Witterberg*. And the Bokkeveld, not the *Bokkerveld* Group.
- Plutons have contact aureoles, not *oriels*.
- Folds may have a vergence, but normally no *vengeance.*
- In order to obtain structural data you normally would use a compass, not a *campus*.
- If you are unfamiliar with a technical term or mineral or rock name, check its meaning before you use it in a text. This applies also to its correct spelling. If you present a wrong spelling of a term readers are usually right when they believe that you actually have no idea what you are talking about.