Incorporating Tables, Figures, Statistics and Equations Effectively Into Your Writing

1. The essential basic elements for tables and figures

The main text should:

(a) tell the reader when to look at a table or figure (and so this reference should appear *before* the table of figure in single column text);

(b) introduce the contents of the table or figure;

(c) point out any key features or trends which the reader should take a note of;

(d) draw a conclusion from the table or figure which answers the "So what?" question. (Schematics of experimental apparatus are one exception to this guideline.)

Example 1

main text

One measure of social disadvantage is relative access to primary health care. As Table 1/Fig. 1 shows, in comparison to capital city residents, Woop Woop's residents **have significantly lower** per capita access to both GPs and pharmacists. This suggests that governments should look at ways of encouraging more doctors and pharmacists to move to rural areas.

Topic sentence to introduce the topic of the paragraph. "As Table 1 shows, …" would rarely be a useful topic sentence.

(a)

(b) and (c). Note that the important feature is only described *qualitatively*, as to given any actual numbers here would be redundant and potentially confusing to the reader.

(d)

Table 1. Access to primary health care providers in rural Woop Woop is significantly lower than in capital cities.*

	Woop Woop	Capital City Average
GPs per 100,000	75.6	103.4
Pharmacists per 100,000	52.0	62.5

* Source: AIHW (1996).

- Table has a descriptive title which is numbered for easy referencing within the text and to also help the reader easily find the table referred to.
- Note that titles are generally found on *top* of the table.
- Note also that if a table or figure is *not* your original creation, you *must cite its source.*

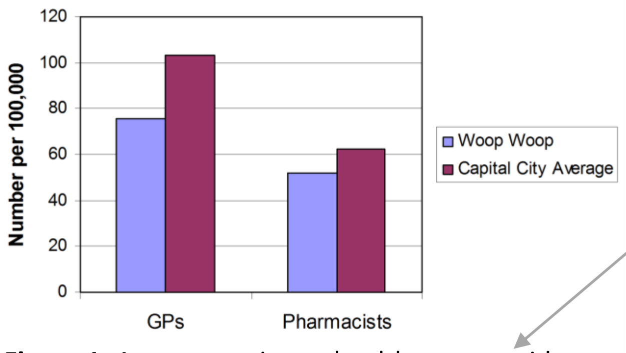


Figure 1. Access to primary health care providers in rural Woop Woop is significantly lower than in capital cities (AIHW, 1996).

- Figure has a descriptive caption which is numbered for easy referencing within the text and to also help the reader easily find the figure referred to.
- Note that captions are generally found *below* the figure.
- Again, if a figure is *not* your original creation, you *must cite its source.*
- Note that when referencing the source, "adapted from" would only be used if you have changed the original in some way.
- Note also that the caption does *not* include "Bar chart of …" as that would be redundant. It can take a conscious effort to write captions which don't include redundant "Graph of / Bar chart of" etc.

Tables versus Figures:

- tables are more precise but can take more work from the reader to interpret;
- graphs and charts are less precise but are generally much easier to interpret qualitatively.

Which is better to use in any given circumstance depends on the writer's objectives and the reader's perceived needs.

2. Figure captions and labels should be used to help the reader understand what they are looking at

Example 2

main text

The rotary lawn mower (Fig. 1) is an ancient piece of technology used by the Suburbites to ... (ref. 1). The design of the cutting blade assembly, illustrated in Fig. 2, shows a number of design features which demonstrate the engineering sophistication of this civilisation. For example, the cutting blade was not held fixed in place, but was allowed to swivel on a recessed bolt, thus allowing it to swivel out of the way if something hard is hit, protecting the blade from damage. Note also ...

Topic sentence to introduce the topic of the paragraph. In this case a reference to a figure (not provided) appears parenthetically in the topic sentence.

(a) reference to figure shown below

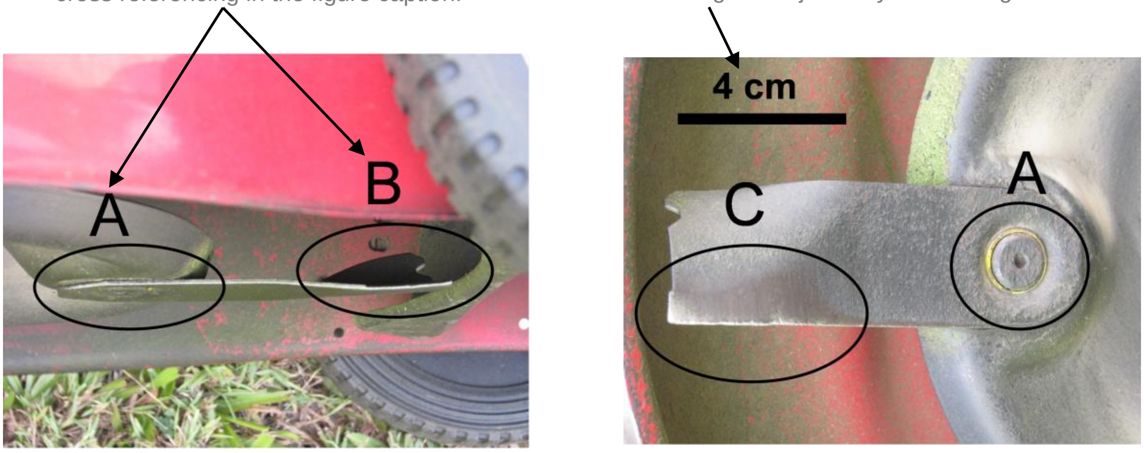
(b) introduction to contents of figure

(c) important features to pay attention to

(d) conclusion to be drawn from the figure

Important features labelled for easy cross referencing in the figure caption.

Note the use of a scale bar to help the reader gauge how big the object they are looking at is.



(a) (b)

Figure 2. Underneath views of a lawnmower showing the cutting blade design.[1] (a)[2] Side view[3] illustrating how the bolt holding the cutting blade is recessed (labeled A)[4] and how the rear edge of the blade is bent upwards to lift cut grass and create a fan to blow the cut grass into the catcher (labeled B)[4]. (b) View from beneath[3] again illustrating the recessed bolt (labeled A) together with a polished part indicating the location of wear[5] (labeled C).

[Photos and design analysis provided by W. L. Rowland.]

Caption notes:

1. General description of what is shown in the figure.
2. Each component part needs to be explained to the reader.
3. To help the reader interpret what they are looking at, they need to know from what direction they are viewing the object.
4. To make it easy for the reader to identify an important design feature it has been labelled.
5. The *significance* of what the reader has been directed to look at needs to be *explained*.

Four basic ways to refer to your tables and figures in the main text:

1. (a) As shown in Table 3, ... / As can be seen from Fig. 2, ... / As indicated in Fig. 6-4, ...
   (b) The result of X was Y, as shown in Fig. 3.
2. Regarding the question of ..., the data in Table 5 shows that ...
3. The factor loadings for the four factors are given in Table 9.
4. As hypothesised, it was found that Group A performed significantly better on the test than Group B (Table 7). OR ... (see Table 7).

Example 3: Additional points for figures

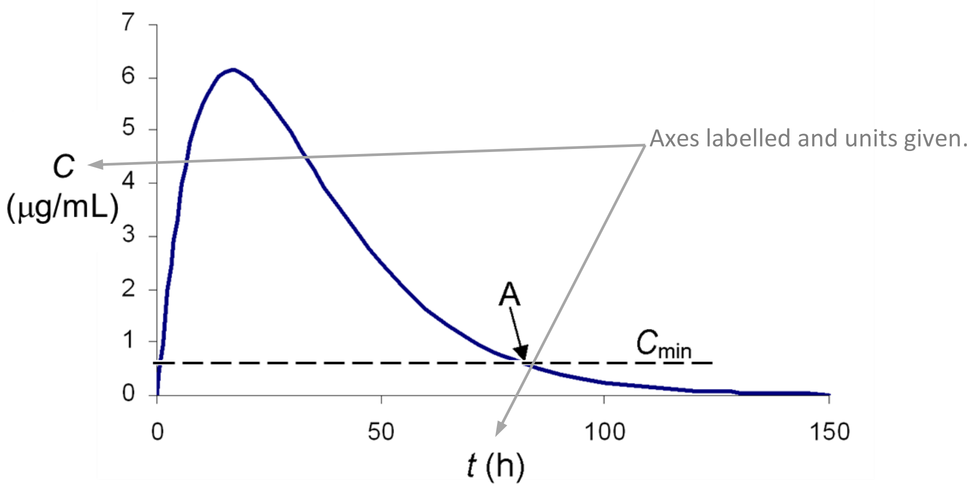


Figure 1. Time dependence of drug concentration in the blood (solid curve). The horizontal dashed line indicates the minimum efficacious concentration and hence the point *A* indicates the latest time at which more drug should be administered if continuous effectiveness is to be achieved.

Significance of different line types explained. Often, different sets of data are indicated by different symbols such as open and closed diamonds and triangles. A legend should explain these.

3. Make your tables and figures as independent of the main text as much as you can without making them excessively wordy

Do this because readers often try to read tables and figures independently of reading the main text, either because they are taking a "short cut" or because on a later reading they are just looking for a specific piece of information.

Example 4[1]

Table 2. The number of 1st year engineering students (N = 108) providing the indicated answer to the question of the units of each term in the differential equation $dD/dt = 100 - 0.01D^2$ shows that very few realise that the units of the terms in a differential equation need to be homogeneous.

	mg/hr*	mg	mg^2	no units	other	no answer
dD/dt	84	6	–	2	3	13
100	10	74	–	6	6	12
-0.01D^2	14	35	26	–	13	20

* The correct units for each term.

Footnote: Help for the reader who may not know what the correct answer is.

Simply writing – "Student answers to the first diagnostic quiz question (N = 108)." – does give the reader a very general idea of what the contents of the table are, but, to interpret the table, the reader would then have to search back through the text to find out what the diagnostic question asked.

M. Cargill & P. O'Connor (2009), *Writing Scientific Research Articles: Strategy and Steps*, pp. 25-26, argue that "story telling" captions and titles are more communicative than simply descriptive titles. Compare the title / caption in Example 1 with: "Rural-city comparison of access to primary health care providers."

[1] Adapted from D.R. Rowland (2006), "Student difficulties with units in differential equations in modelling contexts," *Int. J. Math. Educ. Sc. Technol.* **37** (5), 553-558.

4. Error bars are critical for interpretation

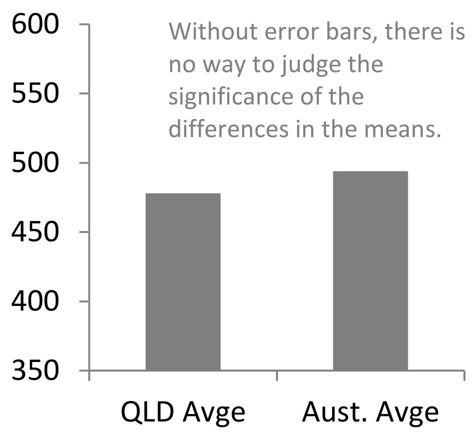
Because of population variability, any sample only provides *estimates* of the actual population characteristics a researcher is trying to measure. Consequently, there is always a certain amount of "statistical uncertainty" in experimental results. When reporting experimental results graphically therefore, "error bars" should be used to depict:

(a) how accurately you have measured something
(b) the amount of statistical uncertainty in a result (sample means only provide an estimate of population means)
(c) the amount of natural variation there is in the population being measured.

Standard errors of the mean (SEM) and confidence intervals (CI) are generally used to depict (a) and (b), while standard deviations (SD) are generally used to depict (c). Which you should use in any given situation depends on what you want to show, but leaving them out suggests your data is much more certain than it really is.

Since error bars could be SEMs, SDs or CIs, it is essential that in your caption you tell the reader what type of error bar you have used.

Example 5



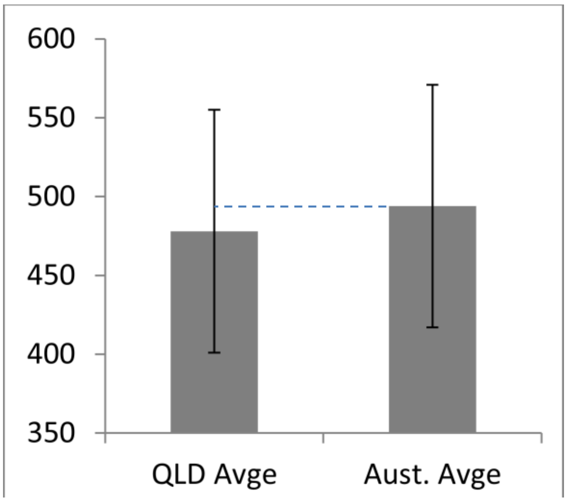


Figure 2. Queensland versus Australian averages on the reading component of the Grade 5 NAPLAN test for 2009 (error bars are standard deviations) revealing a small effect size difference (Cohen's *d* = –0.21).

When comparing group means, authors generally provide SEMs or CIs as they can give some idea as to whether any differences in means are statistically significant or not. In this case though, the sample sizes are so large that the issue is not whether we can be sure there is a difference, but how big is the difference compared to natural variation, the effect size, as this gives some indication as to how difficult it might be to reduce the size of the gap. As natural variation is measured by the SD, that is what has been given.

- Error bars with Excel: http://www.ncsu.edu/labwrite/res/gt/gt-stat-home.html

5. Ensure the meaning of the numbers reported in tables are clear

Some things to note:

- Ensure column headings are adequately descriptive.
- Provide sample sizes.
- Indicate whether errors are standard deviations (SD), standard errors of the mean (SEM), or confidence intervals of what size (CI).
- The statistical significance of group differences should be noted.
- The statistical test used to calculate a *p*-value or equivalent should be indicated.

Example 6

Sample sizes are important information to provide to the reader as they help the reader judge the likely significance of a result.

TABLE 4. Mean Motion Scores

Imaging phase	Overall population ($n = 90$)	Fixed-timing delay group ($n = 45$)	Fluoroscopic triggering group ($n = 45$)	P-value[a]
Precontrast phase	1.50 ± 0.45 (1.0–3.0)	1.37 ± 0.40 (1.0–2.3)	1.63 ± 0.46 (1.0–3.0)	0.005
First arterial phase	2.05 ± 0.69 (1.0–4.0)	1.94 ± 0.60 (1.0–3.7)	2.16 ± 0.76 (1.0–4.0)	0.028
Second arterial phase	2.10 ± 0.71 (1.0–4.3)	1.93 ± 0.59 (1.0–3.3)	2.27 ± 0.78 (1.0–4.3)	0.030
Third arterial phase	2.19 ± 0.77 (1.0–4.0)	2.00 ± 0.70 (1.0–3.7)	2.37 ± 0.80 (1.0–4.0)	0.022
Portal venous phase	1.53 ± 0.45 (1.0–3.0)	1.36 ± 0.40 (1.0–2.7)	1.69 ± 0.45 (1.0–3.0)	0.001
Late dynamic phase	1.45 ± 0.49 (1.0–3.3)	1.34 ± 0.48 (1.0–3.0)	1.56 ± 0.48 (1.0–3.3)	0.007

Data are pooled for all three readers and are summarized as mean ± standard deviation (range).
[a]Comparison of fixed-timing delay group vs. fluoroscopic triggering group. P-values were calculated with the Mann-Whitney test.

Source: J. MAGN. RESON. IMAGING 2016; 43: 1073–1081.

Note the importance of precisely defining the nature of the numbers provided in the table.

The statistical test used to calculate *p*-values is important methodological information and so should be provided to the reader.

6. Reporting statistical results in the body of your text

Example 7

The following example[2] illustrates many of the key features.

> "Table 2 contains descriptive data on the achievement test for the groups ... separated by instructional method and level of mathematics anxiety. Note the relatively homogeneous effect of the discovery method for the low- and medium-anxiety groups that is suggested by the much smaller variances in their scores. As expected, a factorial analysis of variance showed a significant anxiety effect, F (2, 69) = 10.11, $p < 0.01$, with the students with high mathematics anxiety showing lower achievement than the students with low mathematics anxiety (see Table 3). The difference between methods was not significant, but there was a significant interaction between method and anxiety level, F (2, 69) = 4.96, $p < 0.01$. The groups with low and medium levels of mathematics anxiety scored higher, on the average, with the discovery method, whereas the groups with a high level of mathematics anxiety scored higher, on the average, with the expository method."

The type of statistical test used should be indicated so expert readers can assess its appropriateness, and non-expert readers can learn how these analyses are done.

Provide the test statistic, degrees of freedom, and the *p*-value (though note that an exact *p*-value is more informative than a "less than" or "greater than" statement.

To simply say, "there was a significant interaction between method and anxiety level" doesn't tell the reader the nature of the interaction. *Aim to help the reader interpret, where necessary, what your statics are saying.*

The type of statistical test used should be provided. Remember, you should provide enough details about your methods that another researcher in the field could replicate them.

Readers want to know not just, *Was the difference statistically different?*, but also, *How well did each group actually do?*

- An independent samples *t*-test showed that the difference on the spatial reasoning test between the group who had received training ($n = 49$, $M = 54.2$, $SD = 14.7$) and the control group ($n = 52$, $M = 47.0$, $SD = 16.2$) was statistically significant, $t(97) = 2.30$, $p = 0.012$, Cohen's $d = 0.46$.

Exact p-values are more informative than statements like $p < 0.05$, except when $p < 0.001$.

A *p*-value only answers the question, *How confident are we that the difference is not just sampling error?*, it doesn't directly answer questions about the practical significance of an effect, so some measure of *effect size* also provides valuable information.

[2] Source: P. S. Clute, "Mathematics anxiety, Instructional method, and achievement in a survey course in college mathematics," *Journal for Research in Mathematics Education* **15**: 50-58, 1984.

- However, readers might find all that data easier to digest if it is put in a table. And certainly they will find the key results easier to locate. E.g.[3]

"The grade averages were generally higher for the intervention group than for the control group, but only the problem solving exam grades showed a statistically significant difference between the groups with an effect size of Cohen's *d* = 0.65 (Table 2)."

Table 2. *Results of the exams in the cardiovascular system module*

	Problem-Solving Exam				Multiple-Choice Exam			
	n	Mean	SD	*P* Value	*n*	Mean	SD	*P* Value
Control group	39	2.79	1.39	0.0013	39	3.52	3.52	0.098
Intervention group	83	3.60	1.18		83	3.80	3.80	

n, Number of students per group. *P* values were determined by a nonpaired Student's *t*-test for homogenous variance. Results were also determined using the Mann-Whitney test.

- When results aren't statistically significant.
  - The hypothesis that the treatment group (n = 49, M = 49.5, SD = 15.5) would score higher on average than the control group (n = 52, M = 45.3, SD = 20.3) on the XYZ test <u>was not supported</u>, $t(99) = 1.17$, $p = 0.12$.
  - Or: The difference between the treatment group (n = 49, M = 49.5, SD = 15.5) and the control group (n = 52, M = 45.3, SD = 20.3) on the XYZ test <u>was not statistically significant</u>, $t(99) = 1.17$, $p = 0.12$.

[3] Adapted from, *Advances in Physiology Education* 32: 312–316, 2008; doi:10.1152/advan.00021.2007

7. Incorporating equations in your writing

Example 8

Superscripts refer to checklist items.

"The intravoxel incoherent motion (IVIM) theory is an advanced method to separate diffusion and perfusion effects using diffusion-weighted imaging (DWI) [1]. The IVIM theory states that the incoherent blood flow in the capillaries causes a dephasing of the blood magnetization when diffusion gradients are applied. Le Bihan [1] described the resulting signal decay[3] by

$$\frac{S(b)}{S(0)} = (1 - f) \cdot \exp(-bD) + f \cdot \exp\big(-b \cdot (D^{*} + D)\big),^{4} \qquad (1)^{2}$$

where[4] f is the perfusion fraction[5] and D the diffusion coefficient.[5] $D^{*}$is the pseudodiffusion coefficient[5] and describes the additional signal void due to incoherent blood flow in the capillaries. To extract these parameters, the signal must be measured with at least four different diffusion weightings[5] b (including $b = 0$ s/mm^2)."

Source: Magnetic Resonance Imaging 29 (2011) 766–776.

Example 9

"For a left-handed system, the phase[3] in the image domain for a gradient echo sequence with an echo time TE can be written as:

$$\phi(\vec{r}) = \phi_0(\vec{r}) + \gamma \Delta B_z(\vec{r}) TE, \qquad (1)$$

assuming that the main field is in the z-direction. Here,[5] $\phi_0(\vec{r})$ is a coil-sensitivity dependent phase offset and $\Delta B_z(\vec{r})$ is the z-component of the field variation …

Equation (2)[2] can be written as a convolution between $M_z(\vec{r})$ and the point-dipole response $G(\vec{r})$ [25]:

$$\Delta B_z(\vec{r}) = \mu_0 M_z(\vec{r}) * G(\vec{r}), \qquad (3)$$

where $G(\vec{r})$ (the Green's function) is given by:

…

and θ is the angle between $\vec{r}$ and $\vec{z}$. By using the convolution theorem on Eq. (3),[2] $\Delta B_z(\vec{r})$ can be found easily."

Source: Magnetic Resonance Imaging 33 (2015) 1–25.

<u>Incorporating equations checklist</u>

1. <u>Principle</u>: Enough detail should be provided on your methods that a competent researcher in the field could replicate what you have done. But aim for the minimum amount of description that would allow this.
   - This means that standard equations for statistical tests for example, do NOT need to be provided. Just state the name of the test used and the outcome. (i.e. do NOT provide the formula to calculate a SD or *F* or *t*.)
2. Large equations and equations that you want to be able to refer to later need to be put on a separate line and numbered. Such equations are referenced by their number.
3. What an equation is for should be explained.
4. Equations should be incorporated into sentences so that the sentence reads like a proper, grammatically correct and correctly punctuated sentence.
5. All terms in an equation should be defined.

Learn more:

- Some good guides to the reporting of the results of statistical tests both in the body text and with tables and charts are:
  https://www4.uwsp.edu/psych/cw/statistics/Wendorf-ReportingStatistics.pdf
  http://abacus.bates.edu/~ganderso/biology/resources/writing/HTWstats.html and
  http://discoveringstatistics.com/docs/writinglabreports.pdf
- A *Graphic Presentation Handbook* developed by the former *School of Geographical Planning and Environmental Management* at *The University of Queensland* can be downloaded from the bottom of this page: http://www.uq.edu.au/student-services/learning/types-of-assignments