Data Commentary

Robert Blake
Revised March 2013
Outline

- Common mistakes when presenting figures & tables
- Structuring data commentaries
- Varying the strength of claims
- Examples of data commentaries from journals
- Language for dealing with imperfect data
- Choosing tenses in results & other sections
- Further reading
Common mistakes with Graphical Presentation

- Figure/Table is too small for quick access to information
- Figure/table is located too far from text describing it
- Poor layout: figures in series of tests are not arranged to allow comparison
- Information in figure/table is not accessible to any reader other than immediate supervisor: lacks sufficient context & explanation
- Axes are unlabelled or poorly labelled
- Values for axes are not specified
- Values- obscured or difficult to read in relation to trend line
- Figure is neither numbered nor labelled
- Label in wrong position: i.e. should be over a table √, below a figure √
- Plagiarised data: dataset is not your own but source not acknowledged
Common mistakes with Data Commentary

- Figure/table is not referred to in text
- Figure/table is referred to in text but the number is omitted
- Table or figure is referred to but content & main purpose are not explained in text
- Key patterns/trends are not summarised/highlighted for reader
- Commentary does not accurately describe the data
- Unrealistic statements made in the text about the figure/table:
  - claim is too strong & does not reflect the data produced
    [add hedging- see later slides]
  - claim is too weak to describe the data produced
    [too much cautious hedging-see later slides]
- Commentary is not succinct, therefore purpose & main trends are obscured
Providing a commentary for data in graphical format

Writing the data commentary requires as much careful preparation as the presentation of graphical information.

Commentaries have 2 main functions:
1. direct the reader’s attention to the figures
2. enable the reader to focus on the most important aspects of the results.

Data should not be presented without a commentary to guide the reader through them.

Keep the commentary near to the graphical presentation of the data.

[material in slides 5-19 & 23 adapted from Weissburg & Buker (1990) & Swales & Feak (1994)]
Providing a commentary for data in graphical format

Commentaries can consist of up to 3 elements:

1. A location/ summary statement
2. Statements highlighting the most important findings
3. Statements that comment on the results (in combined results & discussions sections)
Data commentary: location & summary statements

Location/summary statement:
This directs the reader’s attention to
a) the location of the figures containing the results
b) expresses the main idea of the graphical information

e.g. Table 3 illustrates the arithmetic mean...
e.g. The results in table 4 show that most of the compounds
Data commentary: location & summary statements

Alternatively the sentence summarises the main idea of the graphical information with the location statement in parentheses e.g.

Caffeine was somewhat more prevalent than theophylline in preventing leaf-eating (figure 1)
Data commentary: location & summary statements

Typical location statements, using the PASSIVE FORM:

- The most common forms of transmission are shown in table 5
- The most common forms of transmission are given in table 5
- The most common forms of transmission are provided in table 5
- As can be seen from table 5, the size of particles is reduced when ...
- As shown in table 5, the size of particles is reduced when ...
Data commentary: location & summary statements

Typical location statements, using the ACTIVE FORM

- Table 5 shows ...
- Table 5 provides ...
- Figure 4.2 gives ...
- Figure 4.2 suggests ...
- As revealed by the graph, ...

SOME VERBS FOR REFERRING TO DATA

<table>
<thead>
<tr>
<th>show</th>
<th>provide</th>
<th>display</th>
</tr>
</thead>
<tbody>
<tr>
<td>summarise</td>
<td>demonstrate</td>
<td>illustrate</td>
</tr>
<tr>
<td>reveal</td>
<td>suggest</td>
<td>indicate</td>
</tr>
<tr>
<td>display</td>
<td>give</td>
<td>present</td>
</tr>
</tbody>
</table>
Other elements in a data commentary

Statements that present the most important findings e.g.

As can be seen in the majority of cases, the source of transmission can be detected by ...

Statements that comment on the results which can be categorised into:

- generalisations from the results
- explanations for the possible results
- comparing the results with those of other researchers

Note the specific to general pattern
A basic data commentary in a results & discussion section

FOREIGN LANGUAGE IN THE ELEMENTARY SCHOOL A COMPARISON OF ACHIEVEMENT

Figure 7.2 displays the mean percentile scores on the four subtests for non-immersion and immersion French students. Students in the French immersion programs performed significantly better than their non-immersion peers on all four Modern language Association tests by more than two to one in terms of scores attained on each of the subtests. For example in the listening subtest, immersion students scored at the 80th percentile, while non-immersion students scored at the 14th percentile. Clearly, the findings indicate that the amount of exposure to a foreign language has a positive effect on student performance. It appears that the intensity of immersion program (an average of 75% of total instruction per week in French compared to approximately 10% for non-immersion) and the use of the foreign language to study basic subjects results in substantial differences in performances in all four skill area of the MLA test. [Table not included. Source Weissburg and Buker (1990:)]
A basic data commentary in a combined results & discussion section

[LOCATION OF RESULTS] Figure 7.2 displays the mean percentile scores on the four subtests for non-immersion and immersion French students.

[MOST IMPORTANT FINDINGS] Students in the French immersion programs performed significantly better than their non-immersion peers on all four Modern language Association tests by more than two to one in terms of scores attained on each of the subtests. For example in the listening subtest, immersion students scored at the 80th percentile, while non-immersion students scored at the 14th percentile.

[COMMENTS] Clearly, the findings indicate that the amount of exposure to a foreign language has a positive effect on student performance. It appears that the intensity of immersion program (an average of 75% of total instruction per week in French compared to approximately 10% for non-immersion) and the use of the foreign language to study basic subjects results in substantial differences in performances in all four skill area of the MLA test.

[NB Divided into 3 paragraphs & headings added for purposes of analysis only; normally a single paragraph or 2 at the most]
Qualifications and Strength Of Claim

Highlighting statements need to be well judged to reflect the quality of your data & the strength of the claim you can make about the data. You may need to

- be cautious-& sometimes critical about the data.
- to choose appropriately cautious language (hedging) to express this caution. Hedging = language that indicates certainty or commitment in varying degrees.

Slides 14 to 19 examine ways of qualifying or moderating a claim in data commentaries.

*Material in slides 16- 20 & 24-adapted from Swales & Feak (1994:86-89)*
Probability (1)

There are many ways to indicate probability in scientific writing. One straightforward way is the modal auxiliary, which is used to ‘show how certain or committed we feel about something’ (Leech at al, 2001, p.287) e.g. will, may, might

Notice how the claim progressively weakens in the 3 sentences.

• A reduced speed limit *will result* in fewer highway injuries.
• A reduced speed limit *may result* in fewer highway injuries.
• A reduced speed limit *might/could* result in fewer highway injuries
**Task**: Put the following sentence variations in approximate order from 1 (strongest claim) to 6 (weakest claim).

Furthermore, several case studies have indicated that climate change ______________ the detrimental effects of human exploitation and mismanagement.

a. has contributed to  
b. has caused  
c. may have contributed to  
d. has probably been a major cause of  
e. has been one of the causes of  
f. might have been a small factor in

[Adapted from Swales & Feak (1994:)]
Task: You can also reduce the strength of claims by choosing a weaker verb e.g.

The widespread use of cfcs has caused global warming.  \textbf{(stronger)}

The widespread use of cfcs has contributed to global warming. \textbf{(weaker)}

Task: Underline the verb making the \textit{weaker} claim e.g. The results indicate/establish that there is a link between smoking and lung cancer.

1. Table 9 suggests/shows that Chinese scientists may need help with writing English.
2. The latest series of experiments question/undermine much previous research.

[From Swales & Feak 1994]
Choosing Weaker Or Stronger Verbs (2)

3. The results given in figure 4 validate/support the second hypothesis.

4. The quantities displayed in the table have been assumed/shown to be about 98% accurate.

5. The test results create/suggest a basis for modification of the research design.

6. Changes in ambient temperature may have influenced/distorted the test results.

7. In their earlier work, they failed/neglected to take ambient temperature into account.

8. As can be seen from table 3, the new pollution controls have encouraged/stimulated research into particulate filters.

9. Figure 12 depicts/clarifies the genetic relationship...

[Slightly adapted From Swales & Feak 1994:88-9]
Notice how these sentences weaken in strength with weaker modals, adjectives etc

It is certain that ...
It is almost certain that ...
It is very probable/highly likely that ...
It is probable/likely that ...
It is possible that ...
It is unlikely that ...
It is very unlikely/highly improbable that
There is a strong possibility that
There is a good possibility that
There is a definite possibility that
There is a slight possibility that
There is a remote possibility that a reduced speed limit will result in fewer injuries.

How is this reduction in strength of claim achieved?

Swales & Feak 1994:87
Examples of data commentaries from journals (1)

Look at the examples of data commentary from a range of journals:

1) Underline the hedging

Several properties of a plant community might influence its resistance to climate change. [i] Previous exposure to climatic extremes [7,8]. For example, plant communities that frequently experience dry conditions may be more resistant to the effects of extreme droughts. [ii] Species richness. Some experimental evidence suggest that diverse plant communities are more resistant and resilient.
Examples of data commentaries from journals (2)

2) Which format is used here for referring to the figure?
Following leaf injury or application of OGAs, systemin, or JA through the transpiration stream, pin 2 transcripts become detectable after 2 to 4 hours (Fig 1). .... (O’Donnell et al. 1996.)

3) Identify 2 verbs used for hedging here [O’Donnell et al. 1996.]
These data suggest that at least one site of ethylene action in the wound response is the regulation of JA levels in the plant. Whereas ASA pre-treatment abolishes any rise in JA, presumably through its multiple inhibitory effects, the specific inhibition of the ethylene synthesis or action only reduces the overall level to some 20 or 30% of that found in wild type. This suggests that two processes contribute to the wound-induced increase in JA, only one of which is ethylene-dependent. This data suggest working levels in the plant .... (O’Donnell et al. 1996.)

Language For Dealing With Imperfect Data (1)

Hedging can also be used where you are having problems with your data & need to reduce certainty when discussing possible causes:

The difference between expected and obtained results *may be due to* the incorrect calibration of instruments

[‘may’= possibility i.e. less certainty than ‘is’]

The discrepancy *can be attributed to* the small sample size

[‘can’ reduces certainty of ‘is’ i.e. there may be other explanations; *is*= only explanation.]
Language For Dealing With Imperfect Data (2)

The lack of statistical significance is probably a consequence of weaknesses in experimental design.

- The anomaly in the observation can probably be accounted for by a defect in the camera.

- The difficulty in dating this archaeological site would seem to stem from the limited amount of organic material available.

[Source Swales & Feak 1994:97]
Tenses For Results Sections

Reporting what you found i.e. results = past tense
Both species of ants exhibited a behavioural sequence.
Referring to figures & tables = present [simple]
Figure 1 shows time series of anthropogenic emissions of carbon dioxide to...
Discussing your results: implications/conclusions= present & present perfect
Thus, the recumbent protein is stable in plants grown at greenhouse temperatures.
Tenses For Other Sections

Reporting what you did i.e. method = past tense
...we isolated a 36-kD glycoprotein from...

Introduction = Present simple e.g.
Recent research *shows*
& Present perfect = e.g.
Research *has established* that....
One strategy that *has evolved*...
Further reading


*Almost Everything You Wanted to Know About Making Tables and Figures* Bates University (2003) http://abacus.bates.edu/~ganderso/biology/resources/writing/HTWtablefigs.html This provides a very useful introduction, with many clearly explained examples.

Weissburg and Buker's *Writing Up Research: Experimental Research Report Writing for Students of English* (1990): geared at international students studying science, it provides excellent analysis of the structure of data commentary & gives lots of language advice on structure and tenses.

Swales and Feak's (1994) *Academic Writing for Graduate Students*: geared at international students. It has a lot of useful advice on structuring various types of postgraduate writing e.g. critical reviews, writing summaries [for literature reviews]. The 3rd edition (2012) is in the University Library.