How to... Produce a bad results section

You worked hard for your data – why share them all? **DANIEL B. WRIGHT** *and* **SIÂN WILLIAMS** *come to the rescue.*

COMPLICATED equations, confusing figures, arcane technical expressions; all commonly found in psychology results sections. In this article we'll show you how to achieve these lofty heights of mind-numbingly boring techno-babble. Inappropriate use of statistical procedures, bad graphs, poor writing style... we'll cover the lot. Your findings will be so obscure that even you won't understand them.

Our approach follows Howard Wainer (1984), who described how to make graphs as uninformative as possible. His approach was to 'concentrate on methods of data display that leave the viewers as uninformed as they were before seeing the display or, worse, those that induce confusion' (p.137). But perhaps Wainer didn't go far enough: we show how entire results sections can be made to 'induce confusion'. Many authors of results sections published in the most respected journals already recognise the value of obscurity. Indeed, the American Psychological Association created a task force to discover why so many follow our approach (Wilkinson *et al.*, 1999).

Statistical tests: Failing the four Rs

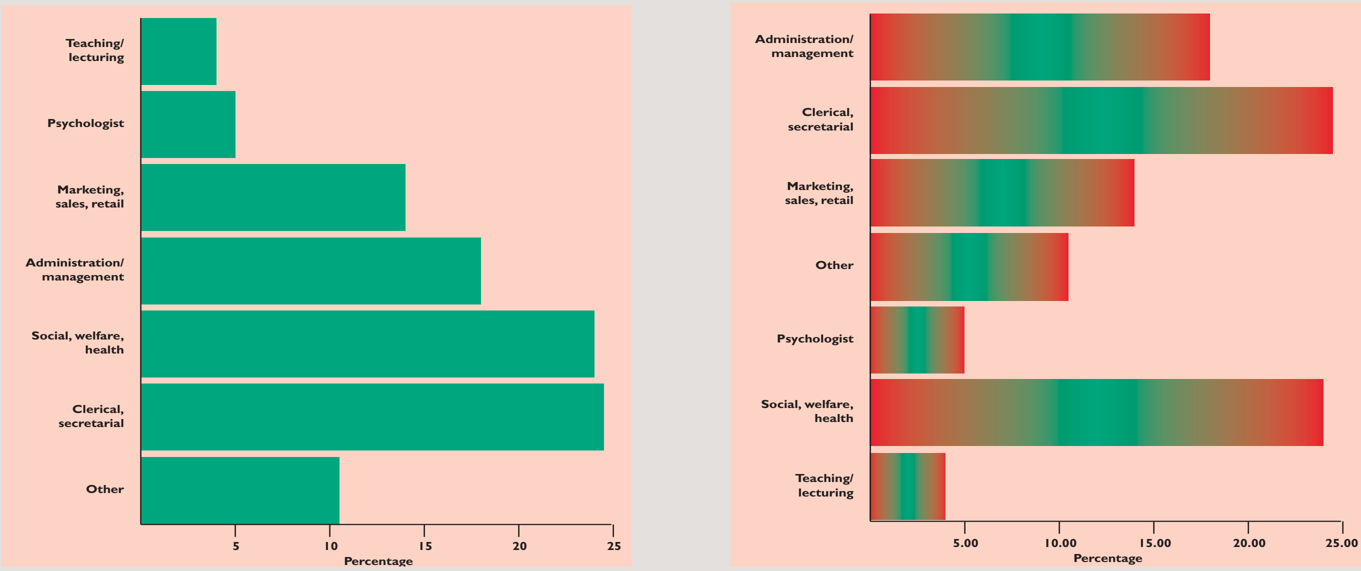
There are three basic ways to miscommunicate findings: numerical, graphical and verbal. The first, numerical, relates to the statistical tests that people conduct.

To produce a results section that is completely misleading, the author could conduct statistical tests that are clearly wrong. Easily achieved when you consider that when questioned, many researchers do not even understand concepts fundamental to much of the statistics that they use, like what *p* means or what a confidence interval is (Oakes, 1986). But, unfortunately, reviewers have a tendency to notice when wrong tests are used. A subtler tactic that can often get past reviewers is only conducting and reporting the final hypothesis-testing statistics, and not exploring the data. Failing to explore the data adequately can mean that interesting facets of the data will not be discovered by the researcher and thus will be hidden from the reader.

Hoaglin *et al.* (1983) discuss the four Rs of understanding data:

- *Resistance* Some statistics are not 'resistant' – they are heavily influenced by a small fraction of the data. (This concept is closely related to a statistic being robust. Resistance is a characteristic of robust statistics.) The mean, ANOVAs, ordinary regressions, and so on, are not resistant, so you should use them (or you could always see Wilcox, 2001, for an introduction to some alternatives).
- *Residuals* This refers to how different

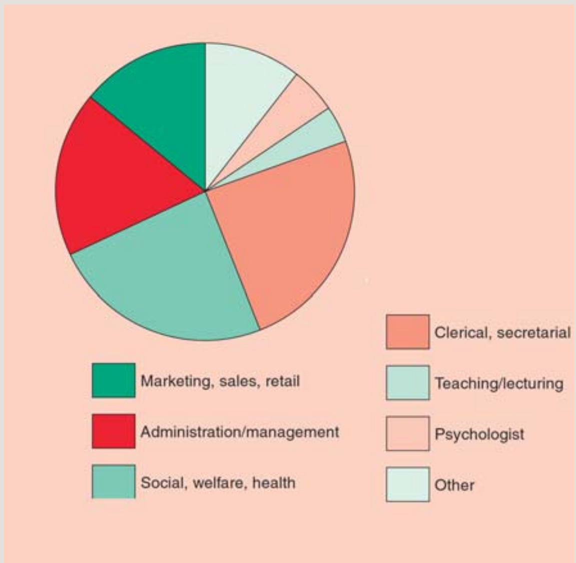
FIGURE 1 Destination of psychology graduates in the UK: same data, different presentation styles



points do not fit with the model. Much as Piaget showed how focusing on children's errors could shed light on cognitive development, it is necessary to examine the residuals to judge the worth of any model.

- *Re-expression* Should the raw data be rescaled to make the analyses and interpretation simpler? Often this is to make the scale of the data more appropriate for the theories being investigated or to make the data more resistant (usually more like the normal distribution). For example, reaction time data are often transformed so that the distribution is not as positively skewed.
- *Revelation* Your methods of analysis can often reveal interesting and unexpected aspects of the data and help inform theories. Following our advice should limit this possibility.

As our goal is to help people create misleading and uninformative results sections, we recommend ignoring the four Rs. Instead, find an introductory textbook that has a flowchart that presents simple questions like 'Are you interested in an association, or group differences?' and 'What is the level of measurement of the data?' and directs the reader to one particular test. This approach, used on its own and with little consideration of the questions, should leave you with just an r, F or χ^2 value as the only means to decide whether the statistical model being considered is appropriate. Don't bother with graphing and examining the descriptive statistics before performing any inferential statistics.

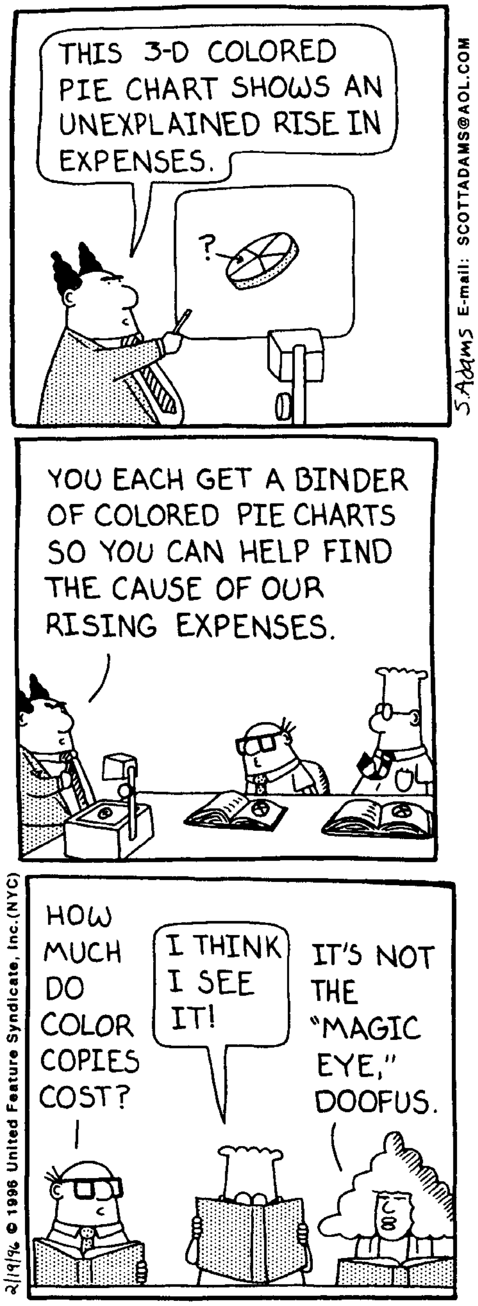


Making bad graphs

For decades the science of graphical display developed so that politicians could construct misleading graphs, with the assumption that the audience was not interested in the numbers unless they were made artistically appealing. Tufte (2001) and Wainer (1984) show some graphs, from highly respected sources, that hide the data from readers, display data inaccurately and present them in a cluttered and confused manner. Many of these methods have arisen because computers can add extra frills to graphs – what is called 'chartjunk'. While computers have made making graphs that clearly communicate the findings easier, they have also created a fertile environment for you to smother your data with technology: 'like weeds, many varieties of chartjunk flourish' (Tufte, 2001, p.107).

As an example, Figure 1 shows three graphs giving the destinations of psychology graduates in the UK in 1999 (data from the BPS document *Studying Psychology*, 2001). In the first graph the reader can see the frequency increasing from 'teaching/lecturing' to 'clerical, secretarial', but the reader won't know that with most graphics software you can tick lots of fancy options. In the middle graph we've alphabetised the items, added a really neat visual illusion to the bars and altered the axis labels. Looks pretty and makes the information more opaque. Many of the patterns that can be used to fill bars create visual illusions about the size, the shape and even the apparent motion of the bars. The third graph is a pie chart, which generally makes it more difficult for the reader to extract information (Hollands & Spence, 2001). Then the boss from Dilbert shows (or he'd say 'demonstrates') the pinnacle of bad graphs. According to the SYSTAT manual, these false 3–D pie charts 'incorporate nearly every visual illusion discussed in this chapter' (Wilkinson, 2000, p.13).

Being creative with colour can further demonstrate how technology can triumph over communication. Consider using red and green, to make the graph particularly bad for the approximately 10 per cent of males who have difficulty distinguishing these colours.



Miscommunicating with the written word

Even if you have written a brilliant literature review, have meticulously and clearly described your study, and have written a fluent and informative discussion, there are still several ways to make your paper 'induce confusion'.

Assume all rules of writing style are not applicable for results sections. Many readers expect results sections to be confusing. This is because the standards of good writing do not apply to results sections. In their book *The Elements of*

Style Strunk and White (1979) argue the case for 'cleanliness, accuracy, and brevity' (p.xiii) in the use of language. Messiness, inaccuracy and extravagance will ensure that people avoid even reading the results section.

Sometimes it is worth doing more than just inducing confusion. Sometimes it is worth producing annoyance. Most scientists are driven by a sense of curiosity (Grigorenko, 2000). We question things. You can feed this curiosity by reporting only some of the information. For instance, omitting the degrees of freedom for an analysis may leave your audience curious about your sample size. It is also possible, and very easy, to neglect to describe any transformations you have made to the data, or how you dealt with any outliers. However, our favourite method is conducting a statistical test, say a *t* test, and not reporting the means. Imagine telling someone in a pub about a study and instead of saying 'the group who were given the drug answered, on average, 10 per cent more questions correctly, and this much of a difference would be unlikely if the drug had no effect', you said '$t(18) = 2.30$' and nothing else. The American Psychological Associations task force on statistical inference (Wilkinson *et al*., 1999) stressed the importance of reporting descriptive statistics. If you wish to produce a bad results section, ignore everything in their report.

Try to make the statistical techniques sound as complicated as possible. Some statistical techniques are very complex. Assume that the reader has a PhD in statistics and knows every statistical technique. Use lots of jargon, particularly if you are using some esoteric technique that you have just learnt. Explain every minute mathematical aspect of the technique. This is easily done by paraphrasing statistics books and manuals – you don't need to understand them yourself. Use a thesaurus to slightly change the meanings of words so that it is not copying word-for-word from the manual. Given that some words have precise technical meanings, this should also confuse readers who felt that they understood the techniques. For example, the words *components* and *factors* have different meanings, and result from different statistical procedures, but are sometimes interchanged.

Use the word *significant* as if it meant the effect was large and important. Some words have a different meanings in English and Statisticalese. Usually the words have the same basic meaning, but have a more precise definition in scientific jargon than in English. But sometimes the meanings are very different. *Significant* in Statisticalese means that assuming the null hypothesis is correct, data as extreme as observed should occur less than 5 per cent of the time. The word *significant* means something very different in English: important. One way to confuse readers is to assume being statistically significant means that the effect is significant in the English sense of the word.

Be careless about using causal language when describing correlational studies. Both causal and associative hypotheses are important in psychology. However, they are different with respect to the theories that are being investigated, and they require different research designs. We recommend casually using words like *cause* and *influence* when conducting studies where there has been no manipulation.

But if you insist on doing it properly…

Irony aside, many of the techniques that we have shown can be easily avoided. There were several themes running through this article.

For analysis, use exploratory techniques to understand the data before you leap into statistical tests. Become friends with your data (Wright, 2002, 2003). Don't just check if a result is statistically significant: look at the size of the effect, ask if it is robust, check to make sure that it is consistent with your graphs, and ask if your finding makes sense.

For graphs, showing technical wizardry and making them 'pretty' can be at odds with the aim of clearly and accurately communicating results. See Wainer and Velleman (2001), and Tufte's marvellous trilogy *The Visual Display of Quantitative Information* (2001, first edition published 1983), *Envisioning Information* (1990), which is about picturing nouns, and *Visual Explanations* (1997), which is about picturing verbs.

For writing styles, think about your audience: second-year undergraduates should be able to understand what you write. Strunk and White (1979) and Sternberg (2000) highlight numerous ways in which writing styles can be used to improve the presentation of research findings.

The art of conducting and communicating statistics is difficult. Abelson (1995) describes how people should consider what they are trying to persuade the reader about. He gives five MAGIC criteria (Magnitude, Articulation, Generality, Interestingness and Credibility) that you should bear in mind whenever you are reporting a statistical result. You should be excited by your results and convey this to the reader. If you are bored by your results, your readers will be too. If scientific papers were murder mysteries, the results section would reveal the killer.

■ *Daniel B. Wright is in the Psychology Department at the University of Sussex. E-mail: danw@sussex.ac.uk.*
■ *Siân Williams is at the Royal College of Paediatrics and Child Health. E-mail: sian.williams@rcpch.ac.uk.*

References

Abelson, R.P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum.
Grigorenko, E.L. (2000). Doing data analyses and writing up their results: Selected tricks and artifices. In R.J. Sternberg (Ed.) *Guide to publishing in psychology journals* (pp. 8–120). Cambridge: Cambridge University Press.
Hoaglin, D.C., Mosteller, F. & Tukey, J.W. (1983). *Understanding robust exploratory data analysis*. New York: Wiley.
Hollands, J.G. & Spence, I. (2001). The discrimination of graphical elements. *Applied Cognitive Psychology, 15*, 413–431.
Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioural sciences*. Chichester: Wiley.
Sternberg, R.J. (Ed.) (2000). *Guide to publishing in psychology journals*. Cambridge: Cambridge University Press.
Strunk, W. Jr & White, E.B. (1979). *The elements of style* (4th edn). Needham Heights, MA: Allyn & Bacon.
Tufte, E.R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.
Tufte, E.R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.
Tufte, E.R. (2001). *The visual display of quantitative information* (2nd edn). Cheshire, CT: Graphics Press.
Wainer, H. (1984). How to display data badly. *American Statistician, 38*, 137–147.
Wainer, H. & Velleman, P.F. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology, 52*, 305–335.
Wilcox, R.R. (2001). *Fundamentals of modern statistical methods: Substantially improving power and accuracy*. New York: Springer-Verlag.
Wilkinson, L. (2000). Cognitive science and graphic design. In SPSS Inc., *SYSTAT 10: Graphics* (pp.1–18). Chicago, IL: SPSS Inc.
Wilkinson, L. and the Task Force on Statistical Inference, APA Board of Scientific Affairs (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594–604.
Wright, D.B. (2002). *First steps in statistics*. London: Sage.
Wright, D.B. (2003). Making friends with your data: Improving how statistical results are reported. *British Journal of Educational Psychology, 73*, 123–126.