® TEACHING STUDENTS TO WRITE ABOUT STATISTICS

FORSTER, Mike, SMITH, David P and WILD, Christopher J
The University of Auckland
New Zealand

To be effective in the real world, our students have to be able to take problems vaguely conceived in natural language terms through the statistical investigation and analysis cycle to arrive at conclusions that they can successfully communicate to others in natural language. Effective communication of statistical findings is one of the critically important legs on which this structure stands. In our second-year applied statistics course we decided some time ago that increasing the ability of our students to communicate their statistical findings was one of the things we most wanted to concentrate on and it now forms a very substantial part of almost all assessments. But assessing a skill you do not systematically teach and develop is not only ineffective but also unfair. In this paper we discuss two highly structured forms of writing which we have developed and call "Executive Summaries" and "Technical Notes", the motivations for concentrating on these particular forms of writing, their strengths and weaknesses, the common scaffolding that stretches across many types of analysis, associated teaching and learning strategies, the challenges we face, and our successes and failures.

INTRODUCTION

To be effective in the real world, our students have to be able to take problems vaguely conceived in natural language terms through the statistical investigation and analysis cycle to arrive at conclusions that they can successfully communicate to others in everyday language. Our second year course in data analysis begins by revisiting basic topics from the first course (one sample, paired and two samples, one-way anova and simple linear regression) with new software (R) and more sophisticated. It then moves on to two-way analysis of variance, the chi-square test, odds ratios, multiple regression, logistic regression and an introduction to time series. The emphasis throughout is on practical data analysis using R. What do we most want our students to take a way from such a course? First, we want our students to be able to recognise which type of analysis is appropriate for a given dataset. Second, we want them to be able to analyse the data using a modern computer analysis package. And finally, we want our students to communicate their findings, not only to the statistically trained, but also to those who have no formal understanding of statistics or its terminology. In this paper, we focus on the third aspect, communication.

There have been many calls for increased emphasis on communication and writing in statistics education, e.g. Gal (2003), Bregar (2002), Samsa and Oddone (1994), Wild (1994), Radke-Sharpe (1991), Hayden (1989). There has been much less written about quite how we should go about developing writing and communication skills in students. We will describe the systems used in our second-year data analysis course. Of all the challenges in this endeavour, perhaps the greatest was the diversity in student language background. As Figure 1 (see over) shows we had students who were born in more than fifty different countries in our second semester class, 2004, and this is the norm rather than the exception. As well as raising equity issues, non-native speakers of English often choose courses from the mathematical sciences in order to avoid courses with high language demands ("language flight.").

In 1997, we began asking students in examinations to write paragraphs describing the main findings from an analysis of data. We asked that they write in non-technical language that would be easily understood by those who have no formal statistical training. We called such paragraphs "Executive Summaries" and the name has stuck even as the conception has developed. There are four main reasons for placing a huge emphasis on writing for a non-technical audience. First, and most obvious, is the real-world need for statisticians to be able to communicate with non-specialists. Second, it is our belief that the discipline of having to write for a general audience forces students to come to a deeper understanding both of statistics and the results of their analyses because they have to wrestle with the real-world meaning of conceptions that are hidden in technical terms. Third, we believed that by formalising our students' thought processes around

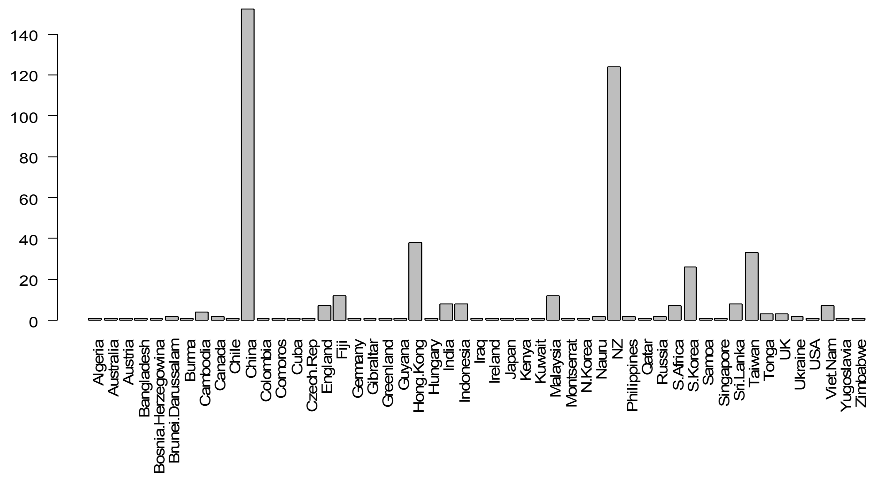

Figure 1. Country of Birth for Students in STATS 20x (Second Semester, 2004).

the reporting of their main findings, we would reinforce some major aspects of data analysis such as the differing interpretations of the results (e.g. additive versus multiplicative models), and also teach them to focus their attention on the main messages in their data. Fourth, we hoped that by focusing more on the communication aspect, our students' ability to formulate and write coherent written reports would be enhanced.

Initially students were just provided with a small number of examples of "executive summaries" written by the teachers intuitively without any clear conception of how they were going about it. The terrible quality of much of the writing we received in response and the gnawing qualms about both the efficacy and ethics of assessing skills we were not actually teaching led, over time, to a much more systematic approach. The highly-systemised forms of writing we are now using in our course, and the teaching processes around them, are the culmination of a seven-year learning process. Gradually the realisation dawned more and more strongly that, at this stage of their development, most students need a great deal more scaffolding and practice to write successfully.

The writing we use involves getting our students to be more active participants in a research cycle. Our initial attempts involved a one-stage process in which recognition, distillation and communication of the main findings from an analysis was all performed at once. Gradually over time we have been thinking more in terms of a two-stage process: an initial "noticing phase" where the students concentrate on their computer output and the main messages they receive from that output, and a second stage where those primary messages are communicated using everyday language. As we developed these forms of writing, we gradually became more and more prescriptive in what we required our students to do for the reasons described above.

COMMON SCAFFOLDING

We wanted students to have structures for writing that were simple enough to internalise so that they could write without needing a checklist. To do this across a wide variety of types of analysis means coming up with a structure that could apply to them all with a minimum of local variations. We have also found that this structure reinforces the main messages from our course. We have been able to tie the communication aspect of data analysis to the general processes that are involved in most of the types of analysis that our students are exposed to in the course and, thus, build and reinforce desirable mental habits.

TECHNICAL NOTES

The idea behind "Technical Notes" was partly separating "noticing" from non-technical communication. These Notes, which do use technical language, also help prepare students to write scientific papers that have statistical analyses in them. The structure we developed for the writing of the Technical Notes was specifically designed to reinforce the processes involved in the Data Analysis Cycle. Usually, in a course on data analysis, the concept of a standard Analysis Cycle is mentioned in the introductory lectures, but rarely discussed again. Repetition, we hoped, would assist in driving the message home and reinforce the idea that a common scaffold is applicable across so many different types of data analysis.

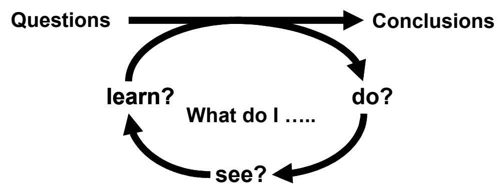

Figure 2. Data Analysis Cycle.

Consequently, the Technical Notes are divided into three sections:

- Exploratory Analysis

We require our students to report the main features they see in an appropriate plot that they have generated of their data. Depending on the data and analysis technique that is required, we expect one or two main features that stand out in their plot to be discussed.

- Checking Assumptions

This requires the students to report, in formal statistical language, the hypotheses associated with and the results of any tests (e.g. Shapiro-Wilk test, Levene test) they have performed, including *P*-values. They are also required to comment on any plots they may have used (e.g. Normal Q-Q plot). They must discuss any transformations of the data they perform and any model building steps that are required (e.g. building a regression model using Backward Elimination).

- Statistical Inference

The students are required to briefly report the main statistical findings from their final model, once again in formal statistical language. We generally restrict this section to discussing the results of the formal test(s) that were performed (*t*-test, *F*-test, Chi-square test), but we do not require our students to report estimates of the size of any effects (confidence intervals) in this section. This could, of course be done here but we emphasise reporting confidence intervals in the Executive Summary.

EXECUTIVE SUMMARY

The major feature of the Executive Summary is that it should not contain any statistical terminology that would only be properly understood by a statistician. When technical language seems unavoidable (e.g. discussing interaction in two-way anova), a non-technical description of any technical terminology must be provided, in the context of their data.

The structure that we developed for the writing of the Executive Summary was designed to reinforce the idea that there is a *set* of main messages that our students should report from any analysis. Our desire was to restrict their report writing to cover the common threads that emerge from most analyses: how the results are to be reported (additive or multiplicative interpretations), the appropriate use of the model they have fitted to the data (e.g. whether a regression model is useful for prediction or whether there is so much noise about the trend as not to be useful), the strength of evidence for any significant effects that are detected as well as reporting any non-significant effects and finally, the quantification of the significant effects using confidence intervals. Thus we developed a global template that has minor variations to cope with certain types of analysis (e.g Categorical Data).

The Executive Summary is divided into five sections:

- Introduction

Here we require a one or two sentence description of the data and the purpose of the analysis. For the students, this usually involves simply taking the main points of the data description and presenting them in their own words.

- A catch-all

The second section is a catch-all for any important information the reader of the report needs to have in order to understand why and how the results are reported. This is where any local variants are discussed. If the data required a transformation before the final analysis was performed, the students need to state this and any implications the transformation has for the interpretation of their findings (e.g. multiplicative, medians). In two-way anova, they are required to discuss whether the factors interact, and give a non-technical description of what *interaction* means in the context of their data. For Regression, we require the students to discuss the goodness of fit of their model and whether it is useful for prediction.

- Strength of Evidence

The students are then required to report the strength of evidence, based on the *P*-values for any effects they have detected, as well as stating which effects, if any, are not significant.

- Quantification

Quantifying the significant effects is then done using the confidence intervals they have generated in the analysis phase.

- Summary

Finally, if the report is quite long or the main message our students see in the analysis gets lost in a mass of figures, a one or two sentence summary of the major findings is required. We have found that a very brief summary of their main findings is often useful for analyses of two-way tables of counts, multiple regression and any data that have come from a designed experiment.

In the assignments and the final examination, we also award marks for general aspects of the students' communication under 4 basic headings: readability, natural language, non-technical language and conciseness.

TEACHING AND LEARNING STRATEGIES

Assessing a skill you do not systematically teach and develop is not only ineffective but also unfair. In light of this, we clearly lay out what we expect from our students in the communication of their findings from an analysis. This is partly answered by providing structure as above. We also give them many worked examples in Case Studies and then provide them with ample practice in writing themselves in the small class tutorials and in their assignment work. An appendix in their printed course notes outlines the basic structure that we require in their reports and details specific aspects that are relevant to particular types of analyses (e.g. dealing with interaction in two-way ANOVA). We also provide a list of technical terminology that *must not* be used in these reports. In some circumstances, technical language is inescapable (e.g. interaction), and when this is the case, we insist that the students give a non-technical explanation of any technical terms they have to use, in the context of the data.

Our course now contains 35 fully worked Case Studies (computer output, Technical Notes and Executive Summary) that are presented in lectures. The lecturer performs the analysis phase live on the computer in lectures and then presents and works through the Technical Notes and Executive Summary pointing out what has been written and why. Our small class tutorials contain 16 examples that are worked through with the students who are then encouraged to write up their findings after the tutorial and we give the students "sample answers" for each of the last three examinations.

In addition our students are required to complete 14 assignment questions where they get a description of the data, the names of the variables and the data file. They are required to analyse the data and write up their findings. In their assignment work, we put our students in exactly the same situation that we face as statisticians. No step-by-step instructions are given. They are required to choose the appropriate type of analysis to perform, analyse the data on the computer and report their findings in their Technical Notes and Executive Summary. We also require them to attach the relevant computer output for each analysis as an appendix to their assignment. The computer output is worth no marks but is included to enable us to determine if they have made any errors in the analysis phase (e.g. not transforming their data when necessary) and give partial credit where it is due. In the final examination we have our students write 2 sets of Technical Notes and 4 Executive Summaries of analyses for which the computer output has been provided.

DISCUSSION

The main strength we see in our approach to teaching the communication of the main findings from an analysis of data is that repeated exposure to report writing in this manner and the detailed instructions they are provided makes it easy for our students to learn how we want them to structure their writing. This applies equally to both the technical and non-technical aspects of their writing. Also, by giving them continual practice in their assignment work throughout the course, our students can learn from their mistakes, before the final examination. Most of our students can now write passages that capture most of the main features of the data and make sense.

The major weakness in our approach is that the reports we get our students to write are very prescriptive in what we get them to comment on and highly structured in their design. While this has made it easier for our second language students, and made grading our students' work relatively easy, it does to some extent stifle creativity in the report writing phase of the analysis. However, writing for a variety of audiences requires complex and difficult skills that cannot be taught in just one course. In our third year applied courses, we are a lot less rigid in what we require from our students as we believe the lessons from our second year course give them enough guidance in communicating their findings in a sensible and comprehensible manner.

Teaching our students to communicate their findings has, on balance, been a very successful aspect of our course. We believe it has assisted our students to appreciate that an analysis of data is not completed when the computer analysis is done and the *statistical interpretation* of the findings are presented. In addition, we have managed to tailor our report writing in such a way that reinforces the data analysis cycle. The Technical Notes have assisted our students in recognising and selecting out only the relevant output from the computer analysis to comment on.

The incorporation of a detailed guide to writing these reports, the inclusion of multiple worked examples in their course materials and redesigning our assignments to incorporate the report writing as a natural last step in their analyses addressed these major short-comings in our early approaches to teaching our students to isolate and communicate their findings in an appropriate and intelligible manner.

ACKNOWLEDGEMENTS

We would like to acknowledge the invaluable assistance of all the staff in our Department for the support, encouragement and assistance we have received while designing this new aspect of our second year course in data analysis. In particular, we wish to acknowledge the specific contributions of Christine Miller, Marti Anderson, Ross Parsonage, Andrew Balemi, Chris Triggs, Russell Millar, Rachel Cunliffe, Alastair Scott and a recent visitor in our Department, Brant Deppa.

REFERENCES

Bregar, L., Ograjensek, I., & Kveder, M.J. (2002). Statistical education and the workplace: Present state of affairs and future challenges. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society, Cape Town, South Africa*. Slovenia: University of Ljubljana.

Crawford, E., & Bowman, A. (2002). Web resources for teaching and learning statistics. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society, Cape Town, South Africa*. Glasgow: University of Glasgow.

Gal, I. (2003). Teaching for statistical literacy and services of statistical agencies. *The American Statistician, 57*(4), 80-84.

Hayden, R. (1989). Using writing to improve student learning of statistics. *Writing Across the Curriculum, 1*(1), 3-9.

Radke-Sharpe, N. (1991). Writing as a component of statistics education. *The American Statistician, 45*(4), 292-293.

Samsa, G., & Oddone, E.Z. (1994). Integrating scientific writing into a statistics curriculum: A course in statistically based scientific writing. *The American Statistician, 48*(2), 117-119.

Wild, C.J. (1994). Embracing the “wider view” of statistics. *The American Statistician, 48*(2), 163-171.