How to report a regression analysis

This document explains what needs to be included where when reporting a regression model(s) in a scientific article. An example of each of the sentences and paragraphs required in the introduction, method, results and discussion is given. On the right (in sans-serif text) is a commentary and explanation of each part.

The example chosen here is deliberately simple to keep things as short as possible: that the variables and measurements don't require special explanations. In a real article you should assume that your reader has completed a psychology degree but is not a specialist in your topic. Any jargon or specialist terms should be briefly explained at the point they are first used.

INTRODUCTION

Academic achievement in higher education is important to students, and university lecturers are concerned to provide the best possible advice to their tutees.

The introduction contains the theoretical rationale for this study design. It usually has an inverted pyramid structure in which the opening is a very broad description of the domain or problem; each paragraph examines the existing literature, narrowing things down so that the final paragraph of the introduction …

Previous surveys of students have found correlations between number of hours worked and student attainment (Bloggs, 2020). However earlier work had provided evidence that students who worked consistently across over the entire academic year performed better than students who massed their practice in the weeks before assessments (Jones, 2010).

Because students who work consistently may also accrue a greater number of hours, this study aims to disambiguate the effects of study duration and consistency.

The final paragraph of the introduction identifies something that we *currently don't know* the answer to and need to collect more data on. This motivates running the study described in the methods.

1

We predict that both total hours studied and the consistency of study will be independent predictors of assessment performance.

The introduction should make at least one specific prediction that is tested in the results section.

METHODS

[…]

Based on the introduction it should be obvious why we want to run *this* study, and the choice of should be an 'obvious' way of answering the question posed.

We used linear regression to predict assessment performance from study duration and study consistency. We compared alternative models using a Bayes Factor. […]

Towards the end of the methods section there will be a paragraph describing the analysis to be used in *general* terms. You don't need to describe each specific. model at this stage — just state the general approach. Optionally, this paragraph has a subheading like "Statistical methods" or "Data analysis".
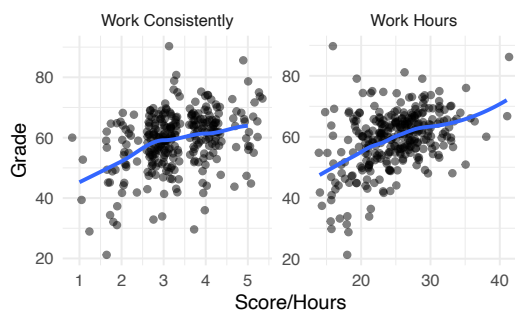
2

RESULTS

Data were obtained for 300 students. Grades ranged from 21 to 90% (mean = 59.2%, SD = 9.57). Students worked between 14 and 41 hours (mean = 24.8, SD = 4.89). Most students (N=140) rated themselves 3 out of 5 on the consistent-working scale (mean = 3.33, SD = 0.851).

We begin by describing the sample. In this example we give figures in the text, but this data could also be provided as a table (e.g. if there were many more variables to describe). 3

Figure 2 suggests…

Always refer to all figures somewhere within the text and…

… that both number of working hours and working consistently were associated with higher grades.

…describe the general pattern of results that each figure shows. Imagine you are telling someone about the plot over the telephone: they don't need to know details about what it looks like, or what colours are used. But they <u>do</u> want to know what you learned from it. 4

Figure 1. Relationship between grades and working hours/working consistently.

Figures should be included roughly where they are mentioned in the text (although some journals ask for them to be at the very end, after the references). Each figure should have a descriptive title with a figure number. All axes should be labelled clearly and include the units of measurement. This example combines two different variables, but you may need to create two separate plots (this was not covered in the lifesavr sessions). 5

To test whether hours and consistency were independent predictors of grades, and to estimate the strength of these relationships…

It is important to give content and explain why a particular analysis is being used. This sentence links our hypothesis, methods, and the insights from Figure 1 with the analysis to come. 6

…we ran two linear regression models. Model A included only study hours as a predictor; Model B included both hours and the consistency.

We need to describe i) what technique(s) were used and ii) how they were applied. We state explicitly that we are using linear regression, that we ran 2 models, and which variables are the outcomes and predictors in each model. 7

Results are presented in Table 1.

Table 1. Results for models A and B

Coefficient	Model A*			Model B**		
	Estimate	95% CI	β	Estimate	95% CI	β
Intercept	38.4	[33.3, 43.5]		34.4	[28.9, 39.8]	
Hours	0.838	[0.636, 1.04]	0.429	0.689	[0.477, 0.901]	0.353
Consistency	—	—	—	2.31	[1.09, 3.53]	0.206
$R^2_{adjusted}$	0.181			0.215		

* Bayes Factor in favour of Model A vs. null model > 10,000
* Bayes Factor in favour of Model B (vs Model A) = 91.2

There was substantial evidence that both number of hours and consistency were a positive influence on grades.

Given our data—model B was 91 times more probable than model A, suggesting that work hours and consistency of work are independent predictors of grades.

DISCUSSION

As expected, we found strong evidence that both car weight and number of cylinders were independent predictors of fuel economy.

[…]

We ALWAYS start the discussion with a restatement of the most important finding: in this case, that our prior prediction was confirmed.
We do not repeat the BF value in the discussion, but the phrase "strong evidence" summarises the BF reported above.

The rest of the discussion would serve to link these new findings to the previous literature and explain implications for theory or practical applications.