

Chapter 18 Solutions

Making Sense of Social Data

Mark Paradis, Ph.D.

September 2, 2026

These are the worked solutions to the *end-of-chapter* exercises for Chapter 18 of *Making Sense of Social Data*. Solutions to the exercises inside each section are in the book itself, printed immediately after the exercises they answer.

Do the exercise before you read the solution. A solution read first is just more text to read; a solution read second is a check on your reasoning, and only the second one teaches you anything.

The complete set for every chapter is at <https://www.markparadispolitics.com/books/making-sense-of-social-data/online/solutions/>.

Chapter 18 Assessment and Interpretation

Solution to Exercise 18.E1. (a) Two of the four. *What was compared* — treatment against control on a wellbeing scale — and *how uncertain*, through t and p together. Missing: *how big* (no means, no difference, no effect size, and no statement of the scale) and *who it is about* (no n in the sentence, though df implies it).

(b) The degrees of freedom give the sample: $n_1 + n_2 = df + 2 = 60$. Then

$$d \approx \frac{2t}{\sqrt{df}} = \frac{2 \times 2.41}{\sqrt{58}} = 0.63$$

A medium-sized effect, which the original sentence gave the reader no way to know.

(c) Something close to: “Participants in the treatment condition ($n = 30$, $M = \dots$, $SD = \dots$) scored higher on the _____ scale than controls ($n = 30$, $M = \dots$, $SD = \dots$), $t(58) = 2.41$, $p = .019$, $d = 0.63$, 95% CI [...].”

What had to be invented: the two group sizes (thirty each, which the degrees of freedom permit but do not require), the means and standard deviations, the name and range of the scale, and the interval. What did *not* have to be invented is the effect size — and noticing which of the missing pieces was recoverable and which was not is the exercise.

Solution to Exercise 18.E2. (i) $p < 0.01$ has a leading zero on a quantity that cannot exceed 1, and gives a threshold where the exact value belongs. $V = 0.31$ has the same leading-zero problem: Cramér’s V runs from 0 to 1. Corrected: $\chi^2(df, N = \dots) = 8.4$, $p = \dots$, Cramér’s $V = .31$ — and the degrees of freedom and N have to be supplied, since a chi-square without them cannot be checked.

(ii) The reverse error. A mean and a standard deviation are in the units of the variable and can be any size, so both keep the leading zero: $M = 0.68$, $SD = 0.12$, $N = 40$. This one reads as careful and is wrong.

(iii) Two leading zeros to remove — $p = .004$ and $\eta^2 = .087$ — since neither can exceed 1. The F and its degrees of freedom are correct as they stand.

(iv) The formatting is right throughout: r and p lose the zero, n is lower case for a single group. What is wrong is the reporting. At $n = 12$ a correlation of .55 has an interval so wide that it is compatible with almost anything, and quoting $p < .05$ rather than the exact value hides how narrowly it cleared. Nothing here breaks a typographic rule and a reader is still misled, which is the distinction §18.3.2 and §18.1 draw between formatting a result and reporting one.

Solution to Exercise 18.E3. Answers depend on the three rows you pick. Three worked, one from each part of the book:

Mean and spread. *What was compared* — nothing; it is a description of one variable, which is why the row is the simplest in the table. *How big* — M . *How*

uncertain — *SD* describes the spread of the cases rather than the precision of the mean, so strictly this row discharges the third obligation only partly. *Who* — *n*. The quantity not reported and available: a standard error, or an interval on the mean, which is exactly what Chapter 12 adds.

Chi-square test of independence. *What* — the two variables, named in the surrounding prose rather than in the statistic. *How big* — Cramér’s *V*. *How uncertain* — χ^2 with its *df*, and *p*. *Who* — *N*, which this row carries inside the parentheses. Not reported and available: an interval on *V*.

Simple linear regression. *What* — the two variables and the direction. *How big* — *b*, in dollars per year, which is the row’s great advantage: the estimate is already in units a reader thinks in. *How uncertain* — *SE*, *t*, *p* and the interval, all four. *Who* — through *df*. Not reported and available: the standard error of the estimate, which §16.5.4 calls the number a reader can actually use and which is almost never printed.

Solution to Exercise 18.E4. (i) Kind 4 — the arithmetic outrunning the design. The association may be perfectly measured; the causal direction and the confounds are untouched by it. Wealthier neighbourhoods have both more parks and lower obesity, which produces this pattern with no causal link between the two. Defensible as a statement of association, or with a design that fixes the direction.

(ii) Kind 2 — dropping the uncertainty, and kind 1 alongside it. “Worked” is a claim about size and *p* is not one. Defensible with the estimate, its interval and an effect size, and with “worked” replaced by something the numbers support.

(iii) Kind 3 — reading a pattern that is not there. Four rises in a row is exactly what §7.5.5 showed can appear in one sample and vanish in the next. Defensible with an interval on the change, or with the resampling argument that section makes.

(iv) Kind 5 — rounding along the way. Each step is defensible on its own and the sequence is not: .0847 to three decimals is .085, and rounding that again to .09 discards the precision the first rounding kept. Round once, at the end: $\eta^2 = .085$.

Solution to Exercise 18.E5. The Results paragraph reports the table, the statistic with its degrees of freedom and *N*, the *p*-value and Cramér’s *V*, plus the direction of the association in words. The Discussion says what the association might mean, what it cannot establish about direction, and how it sits against what is known about education and turnout.

How many sentences you have to move is the exercise, and the usual answer is between one and three — and they almost always move in the same direction, from Results into Discussion. The pull is one-way because the interesting sentences are all interpretive and they arrive while you are writing up the numbers.

The sentence most often misplaced is the one naming the association “weak”. That one is a Results sentence, not a Discussion sentence, because “weak” here is a description of *V* against a stated convention (§14.5) rather than a judgement about

what the finding means.

Solution to Exercise 18.E6. (a) The strongest version: at a given α the two are mathematically linked. A 95 per cent interval for a difference excludes zero exactly when the two-sided test rejects at $\alpha = .05$, so printing both is printing one fact twice, and §18.3.4 says not to report the same number twice.

(b) The reply: they are linked at one value and carry different information everywhere else. The test answers about the single value the researcher nominated; the interval lets a reader check any value they care about — whether the effect could be twice as large, whether it could be too small to matter, whether it is compatible with a rival study's estimate. And the interval reports in the units of the outcome, which the p -value never does. The rule against repeating a number is about the same quantity in three places, not about two quantities that happen to agree at one point.

(c) They say the same thing when the only question is whether zero is plausible — a null of no difference, tested two-sided, at the interval's own confidence level. They part company whenever the question is about a value other than zero: §16.3's slope is significant, and its interval spans a factor of seven, so the test says “not zero” while the interval says “and we have very little idea how large”. Those are not the same finding.

Solution to Exercise 18.E7. Your answers depend on the article, but what you should be able to say about it does not. What to expect:

(a)–(b) The estimate and n are almost always answerable from the paper. Uncertainty usually is. The effect size is the one most often recoverable rather than reported — §18.2.3 gives three routes, and one of them will usually work. What is checked and what was decided after the data were seen is the question most often unanswerable at all, and noticing that is the point of asking it.

(c) Most published articles do not say. What it costs a reader is precise: without it you cannot know how many analyses were run, so the p -value cannot be read as Chapter 13 defines it. That does not make the paper wrong. It makes the single number you would most like to rely on unreadable in the strict sense, which is why §18.2.4 treats pre-registration as a statement about what the reader may conclude rather than about the authors' honesty.

(d) Abstracts almost always claim more than the Results support, and usually by omission rather than by overstatement — the interval and the effect size drop out, and the hedging goes with them. If you find an abstract whose claims are fully supported by its own Results section, it is worth noticing which journal it was in.

Solution to Exercise 18.E8. (a) Answers depend on which comparison you pick. Whichever it was, what it established was the same thing: the simulation resembles the real world in shape and not in level. Chapter 10's comparison found `worldsim`

giving Canada a life expectancy no country has reached; Chapter 11's found a rental share that matches Statistics Canada exactly, and matches it because the simulation was written to. Both are evidence about the code.

(b) Format is what is being marked here, not the number. A correct answer looks like the rows of Table 18.1: the quantity, its units, the year, the source, and a citation — for example, a median after-tax income reported as a dollar figure for a named year from the Canadian Income Survey, not as “about \$75,000 recently”.

(c) Something close to: “The data being simulated means no finding in the book is a fact about Canada, and the book says so every time. What it buys is the thing real data cannot: the true answer is known, so a method can be checked against it — which is how §18.1's interval could be shown to miss the population mean it was estimating. You would not publish from these datasets, and you are not meant to; you are meant to learn on them and then do the same arithmetic on data where nobody knows the answer.”