

Chapter 4 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 4 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 4 Data Collection Methods

Solution to Exercise 4.E1. *Laboratory.* Recruit volunteers to a research session, measure baseline party preference, randomly assign them to view a real campaign advertisement or a neutral control video, then measure preference again. Data: preference before and after, on a rating scale (ordinal, often treated as interval), plus a forced choice (nominal). Strength: complete control over exposure, so any change is attributable to the advertisement. Limitation: participants watch attentively in a room with nothing else happening, which is not how campaign advertising is encountered.

Survey experiment. In a national online survey, randomly assign half of respondents to see the advertisement embedded in the questionnaire before the preference question. Data: one preference measure per respondent plus assignment, with demographics. Strength: a large, representative sample at low cost, and the comparison is between randomised groups. Limitation: it measures a preference stated moments after exposure, which may not persist.

Field experiment. Randomly assign postal areas or households to receive campaign mail, and measure preference in a follow-up survey weeks later — or, better, measure actual turnout and vote choice where records permit. Strength: real advertising, real voters, ordinary conditions. Limitation: no way to know who read the material, and contamination between neighbouring areas.

Natural experiment. Exploit a boundary where media markets do not align with electoral districts, so some districts receive heavy advertising aimed at a neighbouring race for reasons unrelated to their own politics. Compare preference change across the boundary using existing survey data. Strength: scale, realism, and an exposure nobody assigned. Limitation: the whole argument rests on the two sides of the boundary being otherwise comparable.

Recommendation. For an initial finding, the survey experiment: it is affordable, representative, and cleanly randomised. For a claim about real campaigns, the field experiment, because the laboratory result may not survive contact with the way people actually encounter advertising. Decide which question you are asking before you recommend a design; the recommendation follows from it and not the other way round.

Solution to Exercise 4.E2. (a) Digital trace data, and secondary. It was most likely collected by searching Twitter's interface or archive for housing-related keywords over some period, classifying each post's sentiment with an automated tool, and computing the share expressing concern.

(b) Two reasons, from several.

The population is wrong. Twitter users are not Canadians — they are younger, more urban, more educated, and more politically engaged. Users who post about

housing are narrower still. Nothing about the platform's data can correct for this.

Posts are not people. A small number of highly active accounts generate a disproportionate share of posts, and organised campaigns amplify deliberately. Counting posts weights people by how much they post; a population estimate requires weighting them equally.

A third reason: people tweet about housing when they are angry about it, so the platform selects for concern before any measurement happens.

(c) A probability-sample survey with a known sampling frame — Statistics Canada's Canadian Housing Survey, or a national omnibus poll with published methodology. It supports inference to the population because every member has a known chance of selection, the question wording is fixed and documented, and results can be weighted to population benchmarks.

Solution to Exercise 4.E3. (a) IV: whether the school received the anti-bullying programme. DV: student wellbeing, measured with a validated scale. Control variables: school size, neighbourhood socioeconomic composition, and baseline wellbeing before the programme began. Prior bullying incidence is also defensible.

(b) A reasonable choice is a survey administered to students before and after, in programme and comparison schools.

What does it measure? Self-reported wellbeing — what students are willing to report, not their internal states.

What can it answer? Whether wellbeing changed more in programme schools than comparison schools. With random assignment of schools it supports a causal claim; without it, an association.

What are its risks? Social desirability, since students may not want to report distress; and the school context, where students may doubt their answers are private.

Structured observation of playground interactions is a strong alternative, and combining the two would be stronger still — each fails where the other holds up.

(c) Two ethical issues. *Minors* cannot consent independently: parental consent plus student assent is required, and students must be able to decline without consequence. *Distress and disclosure*: asking about bullying may surface active harm, which creates a duty to have a response protocol and support available, and may trigger mandatory reporting obligations.

A third: withholding the programme from comparison schools. A waitlist design addresses it.

(d) If schools were randomly assigned, this is a field experiment supporting a causal claim about the programme's effect on reported wellbeing. If schools chose to adopt it, the design is quasi-experimental and supports only an association — schools that adopt an anti-bullying programme likely differ from those that do not.

Solution to Exercise 4.E4. (a) A repeated cross-sectional survey shows how the

population's attitudes changed: the national average in 2015 against 2025. It cannot say whether individuals changed their minds, because different people answered each time. A stable national average is consistent with enormous individual churn.

A panel follows the same respondents, so it can establish that individuals changed, in which direction, and in what order relative to other events in their lives.

(b) The cross-section is cheaper, faster, and does not suffer attrition; each wave can also use a fresh representative sample, so it stays representative of the current population. The panel is expensive and slow, loses respondents non-randomly, and its sample ages with it — after ten years it no longer represents the adult population, because nobody young has entered.

(c) For the stated question — *changes in attitudes over ten years* — the answer depends on what “changes” means. If the interest is in whether the country’s views shifted, repeated cross-sections are sufficient and much cheaper. If the interest is in whether people change their minds, only a panel will do. The move here is to notice the ambiguity rather than pick one and hope.

(d) Cross-sectional: the Canadian Election Study, run at every federal election since 1965, or the World Values Survey waves. Panel: the Canadian Election Study carries panel components across recent cycles, and Statistics Canada’s Longitudinal and International Study of Adults follows the same individuals.

Solution to Exercise 4.E5. (a) Variables. *Neighbourhood income level* — ratio if measured as median household income, ordinal if grouped into bands. *Number of extracurricular activities* — numerical, discrete, ratio. *Academic achievement* — ratio if a test score or GPA, ordinal if a grade band. The unit of analysis is the child, with neighbourhood as a context variable, which matters for what the findings can claim.

(b) *Primary*: a survey of parents or students in schools sampled from high- and low-income neighbourhoods, asking about participation, with school records supplying achievement. *Secondary*: census data on neighbourhood income linked to school board administrative records on enrolment in school-run activities and standardised test results.

(c) *Primary approach*. Internal validity threat: selection — families who respond to a school survey are likelier to be engaged, which correlates with both activity participation and achievement. External validity threat: the sampled schools may not represent Quebec schools generally.

Secondary approach. Internal validity threat: the records capture only school-run activities, so children in community or religious programmes appear to participate in nothing — and that misclassification is likely to differ by neighbourhood income. External validity threat: coverage is limited to one board’s jurisdiction and its particular programme offerings.

(d) For an initial exploratory study, the secondary approach: it is fast, cheap, covers many children, and would establish whether the pattern exists at all. For a definitive policy evaluation, neither is sufficient — the question is causal, and both designs are observational. A policy evaluation would need random assignment of an activity-access intervention, which is feasible in a way that assigning neighbourhood income is not.

Solution to Exercise 4.E6. 1. *Double-barrelled*. Teaching quality and academic advising are separate. Split into two questions.

2. *Loaded*, and also *leading*. “Excellent” asserts a premise the respondent has not agreed to, and “do you agree” invites agreement. Rewrite: “How well do the university’s library services meet your needs?” with a labelled scale from “not at all well” to “very well.”

3. *Vague quantifier*. “Frequently” means different things to different students, and “stress” is undefined. Rewrite: “In the past seven days, on how many days did you feel so stressed that it interfered with your studying?”

4. *Recall error*, and also *double-barrelled* in a subtle way — it asks for an average across three years, in which study patterns almost certainly changed. Rewrite: “In the past seven days, about how many hours did you spend studying outside class?”

5. *Double negative*. “Not be unwilling” requires unpicking two negations. Rewrite: “How likely would you be to recommend this programme to a friend?”

The response bias: acquiescence. A long questionnaire made entirely of agree–disagree items invites respondents to agree with everything, both because agreeing is the low-effort response and because a tired respondent satisfices. The two compound: satisficing supplies the motive, and the uniform agree–disagree format supplies the opportunity. The standard countermeasure is to reverse-word a portion of the items so that uniform agreement becomes visible as a contradiction.

Solution to Exercise 4.E7. Your study is your own, so this solution sets out what each of the four parts has to do, then carries the worked specimen from Exercise 3.E6 forward.

1. *Answer the three questions specifically*. “A survey measures attitudes” does not answer what the method measures. What was recorded, from whom, and by what procedure does.

2. *Search for the secondary data, do not assume*. Name what you searched — specific programmes, specific variable lists — and what you found, including “nothing suitable, because the item was not asked after 2018.” A negative result, honestly reported, answers this part fully, and it is the commonest real outcome.

3. *Find a genuine flaw in each question*. It is easy to draft three questions and declare them fine. The flaws to look for are the ones in §4.2.3 — a vague quantifier, an implied premise, a recall demand.

4. *Name the missing group and its consequence.* Not “some people might not respond” but which people, and what their absence does to your conclusion.

The specimen, continued. The Chapter 3 study asked whether paid work during the semester lowers CEGEP students’ term grades.

(1) *Method:* a self-administered online survey of students, linked with their consent to institutional grade records. *What does it measure?* Reported weekly paid hours, and grades as the registrar recorded them — one self-report and one administrative measure, which is the right split, because the grade is the variable a student has most reason to misreport. *What can it answer?* Whether students who work more hours have lower averages in the same term. Not whether working caused it. *What are its risks?* Recall error on hours, which vary week to week; and selection, since students under academic pressure are the least likely to fill in a survey.

(2) *Secondary search:* Statistics Canada’s Labour Force Survey carries hours worked by age and student status, but not linked to grades. The Canadian Undergraduate Survey Consortium surveys work and study time, but covers universities rather than CEGEPs and reports grades by self-report band. Result: nothing suitable exists at the level needed, which justifies collecting primary data — and saying so is part of the answer.

(3) *Three draft questions, each with its flaw.*

1. “How many hours per week do you work?” — *vague reference period.* A student with an irregular schedule cannot answer. Repair: “In the past seven days, how many hours did you work for pay?”
2. “Does working interfere with your studies?” — *implied premise and a leading frame.* It presumes interference and invites agreement. Repair: ask for hours studied, and let the association answer the question instead of the respondent.
3. “How many hours did you study each week last semester?” — *recall demand* across four months. Repair: the same seven-day window as the work item, so the two are comparable.

(4) *The missing group:* students who dropped out during the semester, who are disproportionately those working the most hours and doing worst. They are absent from both the survey and the grade file. Their absence removes the cases where the effect should be largest, so the study will *understate* the association — and the direction of that bias has to be stated, because a reader who assumes attrition just adds noise will read the result as weaker evidence rather than as a floor.

Solution to Exercise 4.E8. (a) A survey experiment. Respondents were randomly assigned to different versions of the same question inside a questionnaire, which is the defining feature.

(b) Social desirability bias, and non-response driven by it. Income is prone to this for two reasons that compound: reporting income invites judgement in both directions, and higher earners have a specific additional motive — privacy about wealth, and awareness that the figure could be used to place them. The result is that income non-response is concentrated at the top, which biases estimated mean income downward. Note that this is systematic, so a larger sample makes it worse rather than better.

(c) The answer depends on the survey you chose. The Canadian Election Study and the World Values Survey both carry trust items. Expect the real wording to be more specific than you predicted — naming a particular institution, or specifying a four-point rather than an eleven-point scale — and expect the object of trust to be narrower than “the government.”

That gap is the point. The real item measures something slightly different from what the concept name suggests, which is the operationalisation problem from Chapter 2 appearing in a published instrument rather than a textbook example.