

A seven-point gain that was the school year

Literacy rose 7.0 points between baseline and endline, and a proposal is calling it the effect of the school feeding programme. The schools without the programme rose 7.6 points. All of the gain and slightly more belongs to the year, not the meals.

Pierre Robentz Cassion

Lesson 1 of 8

Impact Evaluation Methods — The counterfactual

What this lesson covers

- One question, three answers
- The counterfactual, stated in one sentence
- Why the before-after number is so persuasive
- Three things a before-after change contains
- When a before-after comparison is legitimate
- Report it whole
- What comes next

One question, three answers — In Python

```
import pandas as pd
import numpy as np

assessment = pd.read_csv("learning-assessment-2024.v1.csv")
assessment["pct"] = assessment["raw_score"] / assessment["items"]
literacy = (assessment[assessment["domain"] == "literacy"]
            .pivot_table(index="student_id", columns="assessment_round",
                          values="pct").dropna())

roster = pd.read_csv("school-roster-2024.v1.csv").drop_duplicates(
    "student_id", keep="last")
d = literacy.join(roster.set_index("student_id"), how="inner")

print(d.groupby("feeding_programme")[["baseline", "endline"]].mean().round(4))
print(f"overall gain: {(d['endline'] - d['baseline']).mean():+.4f}")
```

One question, three answers — In R

```
library(dplyr)
```

```
d |> summarise(baseline = mean(baseline), endpoint = mean(endpoint),  
              .by = feeding_programme)
```

One question, three answers

	Baseline	Endline	Gain	n
Feeding programme	57.75%	64.41%	+ 6.66 pts	398
No programme	54.67%	62.29%	+ 7.62 pts	187
All students	56.76%	63.73%	+ 6.97 pts	585

One question, three answers

- "The programme delivered a 7.0-point gain." — Before-after, paired, 95% CI +6.1 to +7.9, $t = 15.2$
- "Schools with the programme are 2.1 points ahead." — Endline cross-section, 95% CI -1.2 to +5.4
- "The gain was 0.96 points lower in schools with the programme." — Both groups against their own baselines, 95% CI -3.5...

The counterfactual, stated in one sentence

Literacy in the feeding schools rose 6.66 points compared with what would have happened without the programme.

The counterfactual, stated in one sentence

- The second clause is the counterfactual, and it is never observed — The same children in the same year without meals is. . .

The counterfactual, stated in one sentence

Claim	Counterfactual it assumes	Is that plausible?
Before-after	The same children would not have changed at all	No — a school year happened
Endline cross-section	The two groups would have been identical	No — they differed by 3.1 points at baseline
Difference-in-differences	The two groups would have changed by the same amount	Arguable, and lesson 3 is that argument

The counterfactual, stated in one sentence

- Naming the counterfactual is the whole method — Once it is written down, the before-after claim collapses without any. . .

Why the before-after number is so persuasive — In Python

```
gain = d["endline"] - d["baseline"]
se = gain.std(ddof=1) / np.sqrt(len(gain))
print(f"{gain.mean():+.4f} 95% CI [{gain.mean() - 1.96*se:+.4f}, "
      f" {gain.mean() + 1.96*se:+.4f}] t = {gain.mean()/se:.1f}")
```

Why the before-after number is so persuasive — In R

```
t.test(d$endline - d$baseline)
```

Why the before-after number is so persuasive

- $t = 15.2$ — Every gate in the statistics course passes
- None of those checks can detect a missing comparison group — and that is the thing worth carrying out of this lesson

Three things a before-after change contains

- The passage of time — Children get older and are taught
- Regression to the mean — Whoever was selected for being worst off will look better next round even if nothing was done, . . .
- Everything else that happened — A drought, a currency movement, another agency's programme in the same district, a . . .

Three things a before-after change contains — In Python

```
by_school = d.groupby(["school_id", "feeding_programme"])[  
    ["baseline", "endline"]].mean()  
by_school["gain"] = by_school["endline"] - by_school["baseline"]  
print(by_school.groupby("feeding_programme")["gain"].describe()[  
    ["mean", "min", "max"]].round(4))
```

Three things a before-after change contains — In R

```
by_school |> summarise(mean = mean(gain), min = min(gain), max = max(gain),  
                        .by = feeding_programme)
```

Three things a before-after change contains

- Every one of the 24 schools gained — from +1.2 points to +12.2

When a before-after comparison is legitimate

- When the counterfactual is genuinely flat and you can show it — Three or more pre-programme rounds with no trend is an...
- When the outcome cannot change on its own — A latrine does not build itself
- When you are monitoring, not evaluating — "Attendance rose from 84% to 88%" is a true statement about the world and a...
- The sentence to watch for is "as a result of" — It is where a monitoring number becomes an impact claim, and it is...

Report it whole — Example

Literacy outcomes, baseline to endline 2024

All students	56.8% to 63.7%	+6.97 points	95% CI +6.1 to +7.9
With feeding	57.8% to 64.4%	+6.66 points	
Without feeding	54.7% to 62.3%	+7.62 points	

The overall change is reported as a monitoring result and is not attributed to the feeding programme. Schools without the programme gained slightly more, so the difference-in-differences estimate of the programme effect is -0.96 points (95% CI -3.5 to +1.6, clustered on 24 schools).

The before-after change assumes children would have learned nothing over a school year. That counterfactual is not defensible here and the comparison group is what replaces it.

Report it whole

- The last paragraph is three lines and it is the difference between a monitoring report and an impact claim — Writing. . .

What comes next

- The difference-in-differences estimate above rests on the two groups of schools being comparable in the ways that matter.

Read the full lesson, with runnable code

```
https://data-analysis.cassion.dev/courses/impact-evaluation-methods/  
ie-01-the-gain-that-was-the-school-year
```