

# With No Cause at All, the Bottom Rises **1.76** Standard Deviations and the Top Falls **1.76**

— Everyone equally able, both measurements independent, and the bottom decile moves **+1.756597** while the top decile moves **−1.759617** — One treatment looks opposite depending on which group was chosen —

[Paper 371]

---

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

September 2026

— *From thinking to syncing.* —

## Abstract

It is said that drilling the weak improves them and praising the strong spoils them. This paper asks whether that is the effect of the treatment—the answer is that it happens at the same size with no treatment at all. No new mathematical theorem and no new law is claimed. First, everyone is given equal ability and the two measurements are independent noise (Section 2). Second, this is the core of the paper. The bottom decile's second measurement rises by **+1.756597** (Section 2). Third, at the same time the top decile falls by **−1.759617** (Section 2). Fourth, both second means are near zero (0.008800 and  $-0.000170$ ) (Section 2). Fifth, the middle half barely moves ( $+0.004677$ ) (Section 3). Sixth, the separator is whether the group was chosen on the first measurement, not what the treatment was (Section 5).

Scope of this paper (scope note): No new mathematical theorem and no new law is claimed—regression to the mean, Galton’s regression, and regression discontinuity are all standard. We do not build statistics—corrections for regression (control-group design, difference-in-differences) are not entered. They are named and no more. We use no real data—actual records from education, medicine, or sport are not used. The computation is on standard normal draws. We do not build causal inference—identification conditions for treatment effects are not the subject. Only what is seen when the treatment is zero. Correlation is not treated—the two measurements are taken as fully independent. In reality they correlate, and the regression is smaller by that much. We do not generalise—not every before-and-after comparison is false. Only those in which the group was selected on the first measurement. Relation to earlier papers: Paper 364 showed selection alone raising the mean by 1.52 standard deviations—there the issue was which group is visible; here it is the “change” when one group is measured twice. Paper 353 showed a claim turning false when its condition is dropped—here the dropped condition is how the group was chosen. Paper 300 showed whether two things share a root is decidable—by that test, a treatment effect and a regression effect have distinct roots. Paper 361 showed a classifier failing when tried on a known instance—here the known instance is a treatment of exactly zero. What is added is giving first and second means for five groups, showing that bottom and top move in opposite directions by nearly equal amounts, showing both second means returning to zero, placing a control in which making the second measurement identical to the first removes the effect exactly, and putting the separator on selection by the first measurement.

## 1. Notation and the Question

Setting: Give 100000 people an ability of exactly 0 and make the two measurements independent standard normal noise.

$$X_1 \sim \mathcal{N}(0, 1), \quad X_2 \sim \mathcal{N}(0, 1), \quad X_1 \perp X_2$$

Choose groups by  $X_1$  and look at their  $X_2$ .

### Observation 1.1

There is no treatment. Nothing was done between the measurements—so any change is entirely an appearance.

### Observation 1.2

There is no difference in ability either. All drawn from one distribution—the first ranking is settled by luck alone.

### Observation 1.3

And still a “change” appears. Because the operation of selecting on the first measurement enters—this paper gives its size as a number.

Question: How large an “effect” appears with no treatment?

## 2. +1.76 at the Bottom, −1.76 at the Top

Calculation 2.1: First and second means of groups chosen on the first measurement (computed in this paper, 100000 people).

| Group by first     | Count | First mean | Second mean | “Effect”  |
|--------------------|-------|------------|-------------|-----------|
| Bottom 10 per cent | 10054 | −1.747797  | 0.008800    | +1.756597 |
| Bottom 25 per cent | 25013 | −1.269273  | 0.004057    | +1.273330 |
| Middle 50 per cent | 49885 | −0.003051  | 0.001626    | +0.004677 |
| Top 25 per cent    | 25102 | 1.274214   | 0.008623    | −1.265591 |
| Top 10 per cent    | 10060 | 1.759448   | −0.000170   | −1.759617 |

### Observation 2.2

This is the core of the paper. +1.756597 at the bottom decile and −1.759617 at the top—opposite in sign, nearly equal in size, and not one treatment was applied.

### Observation 2.3

Both second means are near zero. 0.008800 and  $-0.000170$ —both groups have returned to the population, as they must when ability is equal.

### Observation 2.4

So the “change” lives on the first side. That the first mean was  $-1.747797$  makes the whole of it—not anything that happened at the second.

### Observation 2.5

One treatment looks opposite by group. Applied to the bottom it looks effective; applied to the top it looks harmful—and the treatment is “do nothing”.

### Observation 2.6

This is the most dangerous shape in practice. The observation “scolding improves and praising spoils” is fully explained by it—and being able to explain something is not evidence that the explanation is right.

## 3. Away from the Extremes, the Regression Is Small

---

### Observation 3.1

The middle half barely moves.  $+0.004677$ —one three-hundred-and-seventy-sixth of the bottom decile’s 1.756597.

### Observation 3.2

So the size is set by how extreme the selection is.  $+1.273330$  for the bottom quarter,  $+1.756597$  for the bottom tenth—the harder the selection, the larger.

### Observation 3.3

There is practical guidance in that. A design that picks an extreme group and compares before with after produces the strongest false effect—and picking the extreme group is usually the natural decision to start where the trouble is.

### Observation 3.4

A control removes it. Make the second measurement identical to the first and the effect is exactly 0—naming the noise as the cause.

### Observation 3.5

To be stated honestly—real pairs of measurements are correlated. This paper took them as fully independent—with correlation the regression is smaller. The 1.756597 here is an upper bound.

## 4. The Difference from Paper 364

---

### Observation 4.1

Paper 364 treated the mean of the visible group. When dropouts are not recorded, the survivors' mean looks high—there only one measurement was taken.

### Observation 4.2

This paper treats the change in one group. Measured twice, and the difference examined—nobody drops out; everyone is recorded twice.

### Observation 4.3

So the remedies differ. Paper 364's bias vanished once dropouts were recorded—this one does not vanish even with complete records. A control group is needed.

#### Observation 4.4

Distinct roots by the criterion of Paper 300. One is a cropping of observation, the other a combination of selection and re-measurement——both produce “it looks better” from different roots.

#### Observation 4.5

And this one is harder to detect. The data have no gaps, so looking at the records shows nothing wrong——only looking at the design reveals it.

### 5. The Separator: Was the Group Chosen on the First Measurement?

#### Observation 5.1

The separator——whether the group was chosen on the first measurement; not what the treatment was.

| How the group was chosen    | Treatment | Apparent effect  | True effect |
|-----------------------------|-----------|------------------|-------------|
| Bottom 10 per cent of first | none      | <b>+1.756597</b> | <b>0</b>    |
| Top 10 per cent of first    | none      | <b>−1.759617</b> | <b>0</b>    |
| At random                   | none      | <b>about 0</b>   | <b>0</b>    |

#### Observation 5.2

The treatment is the same in all three rows. Nothing was done——only the selection differs, and the apparent effect runs from +1.76 to −1.76.

#### Observation 5.3

Which is why randomisation works. Do not select on the first value and no regression occurs——this is one of the things randomisation is for.

#### Observation 5.4

The separator can be moved, though not easily. In practice one is often forced to choose the group in trouble——then choose the control group by the same criterion. Matched selection makes the regression equal in both groups, so it cancels in the difference.

#### Observation 5.5

To be stated honestly——1.756597 and  $-1.759617$  are values under stated conditions: standard normal, independent, bottom and top tenths. Correlation shrinks them——what does not move is the structure: bottom and top move in opposite directions.

#### Observation 5.6

One last thing——one treatment of exactly zero can produce both “it worked” and “it backfired”. Choose the group according to the conclusion you want——a design whose conclusion can be chosen does not support that conclusion. Look at results without looking at the design and the two cannot be told apart.

## 6. The Census and the Close

---

1. Everyone equally able (Observation 1.2): **the first ranking is luck.**
2. No treatment (Observation 1.1): **any change is appearance.**
3. **+1.756597** for the bottom tenth (Calculation 2.1): **looks like improvement.**
4.  **$-1.759617$**  for the top tenth (Calculation 2.1): **looks like harm.**
5. Both second means near zero (Observation 2.3): **0.008800 and  $-0.000170$ .**
6. Middle half **+0.004677** (Observation 3.1): **one three-hundred-and-seventy-sixth.**
7. Harder selection, larger effect (Observation 3.2): **+1.273330** for a quarter.
8. Identical second measurement gives exactly **0** (Observation 3.4): **noise is the cause.**
9. A different root from Paper 364 (Observation 4.4): **cropping, or select-and-remeasure.**
10. Correlation shrinks it (Observation 3.5): **these values are an upper bound.**

Closing: It is said that drilling the weak improves them and praising the strong spoils them——but the same thing happens, at the same size, with no treatment at all. Give 100000 people equal ability and make the two measurements

independent noise——nothing is done between them, so any change is entirely an appearance. And still the bottom rises and the top falls——the bottom decile moves **+1.756597** and the top decile **−1.759617**: opposite in sign and nearly equal in size. Both second means are near zero (0.008800 and −0.000170) — both groups have simply returned to the population. So the “change” lives on the first side: that the first mean was **−1.747797** makes the whole of it. One treatment looks opposite by group——applied to the bottom it looks effective, applied to the top harmful, and the treatment is “do nothing”. The observation “scolding improves and praising spoils” is fully explained by this — and being able to explain something is not evidence that the explanation is right. The size is set by how extreme the selection is——the middle half moves only **+0.004677**, one three-hundred-and-seventy-sixth of the bottom decile’s. A design that picks an extreme group and compares before with after produces the strongest false effect — and picking the extreme group is the natural decision to start where the trouble is. Make the second measurement identical to the first and the effect is exactly 0, naming the noise as the cause. One thing separates them ——whether the group was chosen on the first measurement; not what the treatment was. The treatment is identical in all three rows and only the selection differs, yet the apparent effect runs from **+1.76** to **−1.76**. Which is why randomisation works; and where one is forced to choose the group in trouble, choosing the control by the same criterion makes the regression cancel in the difference. To be said honestly——real pairs of measurements are correlated and the regression is smaller by that much; the **1.756597** here assumes independence and is an upper bound. And this differs in root from Paper 364: the data here have no gaps, so the records show nothing wrong and only the design reveals it. One last thing——one treatment of exactly zero can produce both “it worked” and “it backfired”. Choose the group according to the conclusion you want. A design whose conclusion can be chosen does not support that conclusion, and looking at results without looking at the design cannot tell the two apart.

## References

---

1. Galton, F. (1886). Regression Towards Mediocrity in Hereditary Stature. *Journal of the Anthropological Institute*.
2. Campbell, D. T., Kenny, D. A. (1999). *A Primer on Regression Artifacts*. New York: Guilford Press.
3. Barnett, A. G., van der Pols, J. C., Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*.
4. Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
5. Bland, J. M., Altman, D. G. (1994). Regression towards the mean. *BMJ*.
6. Yamagishi, Y. (2026). *The Three Hundredth Census: Whether Two Things Share a Root Is Decidable* [Paper 300]. Zenodo. 10.5281/zenodo.22171914
7. Yamagishi, Y. (2026). *Adding Noise Can Improve Detection — But “More Is Better” Is Not the Rule* [Paper 353]. Zenodo. 10.5281/zenodo.22696835
8. Yamagishi, Y. (2026). *A Habitable Zone Drawn from Equilibrium Temperature Puts Earth Outside It* [Paper 361]. Zenodo. 10.5281/zenodo.22697436

9. Yamagishi, Y. (2026). *Selection Alone Raises the Mean by 1.52 Standard Deviations and Shrinks the Spread to 0.45* [Paper 364].  
Zenodo. 10.5281/zenodo.22697530

---

— *From thinking to syncing.* —

**Yuuki YAMAGISHI** (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)