University of Guelph
Numeracy Project

About Simpson's Paradox: Examples

TABLE OF CONTENTS

About Simpson's Paradox

Example 1

► The University of Guelph is notorious for its first year introductory chemistry class, also known as "killer chem." This mandatory class is often used to weed out students in science majors. It is thought that those students who take university-level chemistry in high school will be better at succeeding in "killer chem" than those who only took college-level, or none at all.

The following table illustrates that those "killer chem" science students who took university-level chemistry in high school had a 10 point difference in grades when compared to students who did not take chemistry in high school:

	High School Chemistry	No Chemistry Background	Difference
Number of Students	60	10	---
Average "Killer Chem" Grade	75	65	10

►   Now, let us consider a different group of people. Students enrolled in a general arts program at Guelph are offered the choice to take an arts version of "killer chem" as a first year elective. This is to allow arts students a chance to take a course they normally would not to enrich their university experience. As expected, most arts students did not take high school chemistry. However, for the few that did, the table below shows a 10 point difference in first year "killer chem" over those who did not.

	High School Chemistry	No Chemistry Background	Difference
Number of Students	10	60	---
Average "Killer Chem" Grade	95	85	10

It seems in both cases, whether you are an arts or science major, taking high school chemistry makes it more likely that you will have a higher grade in first year "killer chem." Does this hold when the groups are combined?

►   The following chart is the combined data of both the arts and science students who took university-level chemistry in high school:

	Number of Students	Average "Killer Chem" Grades	Grade Points
Science	60	75	4500
Arts	10	95	950
Total	70	---	5450
Average	---	---	77.86

The average university grade for chemistry ("killer chem") was calculated by adding all the grade points (5450) and then dividing by the total number of students (70).

► The following chart is the combined data of both the arts and science students who did not take university-level chemistry in high school:

	Number of Students	Average "Killer Chem" Grades	Grade Points
Science	10	65	650
Arts	60	85	5100
Total	70	---	5750
Average	---	---	82.14

Again, the average university grade "killer chem" was calculated by adding all the grade points (5750) and then dividing by the total number of students (70).

As we can see, the difference between the combined group averages (taken high school chemistry vs. did not take high school chemistry) indicates a reversal of association. In fact, $82.14 - 77.86 = 4.28$ average grade points more for those with no background in chemistry. What caused this discrepancy?

In fact, there are a number of factors at work here, including the combination of groups of unequal size (in both cases, 60 people + 10 people). In addition, you might consider the combinations to represent the joining of unequivalent groups, as chemistry for the arts is generally considered less rigorous than chemistry for science majors.

Example 2

► This next example is taken from a real-life longitudinal study done in South Africa. Many thanks to Christopher H. Morrell for granting use of this study as an example.

The study in question was the Birth to Ten study (or BTT). The focus of this study was to pinpoint factors related to the emergence of cardiovascular disease risk factors in a cohort of children living in an urban environment in South Africa.

The initial birth cohort was formed in 1990 and consisted of 4029 births. Five years later, these children were followed up.

► The follow-up, five years after baseline, consisted of a detailed questionnaire focusing on a variety of health factors. The follow-up comprised only 964 children (compared to 4029 at baseline). This is a glaring problem, exacerbated by the fact that a lot of data was missing from the baseline group.

As a result, findings were inapplicable, unless the characteristics of the 964 children followed up were similar to those of the non-responders. In other words, as necessity dictated, the five year follow-up group was compared to those who did not participate in the follow-up.

A factor considered by the researchers was whether the mother had medical aid (akin to health insurance) at the time of the child's birth.

► The following table outlines the number (and percentage) of mothers with and without medical aid at the time of their child's birth:

	Children Not Traced	Five-Year Group
Had Medical Aid	195 (16.6%)	46 (11.1%)
No Medical Aid	979 (83.4%)	370 (88.9%)
Total	1174 (100%)	416 (100%)

(Morrell, 1999)

As the table indicates, 11.1% of those in the five-year cohort had medical aid, while 16.6% of those who were not followed up had medical aid. Statistical significance was established at the $p = .007$ level.

► The table below goes a step further, breaking the previous data down by race:

	White		Black	
	Children Not Traced	Five-Year Group	Children Not Traced	Five-Year Group
Had Medical Aid	104 (82.5%)	10 (83.3%)	91 (8.7%)	36 (8.9%)
No Medical Aid	22 (17.5%)	2 (16.7%)	957 (91.3%)	368 (91.1%)
Total	126 (100%)	12 (100%)	1048(100%)	404 (100%)

(Morrell, 1999)

► As the previous table shows, 83.3% of white participants in the five-year follow-up group had medical aid, while 82.5% of those who were not followed up had medical aid. Conversely, amongst black participants, the corresponding percentages were 8.9% and 8.7%.

This data indicates a reversal of the association found before the groups were broken down by race. That said, the difference between the percentages over race is not statistically significant ($p = .945$ for whites and $p = .891$ for blacks, respectively).

5

►   Given the social context of this example, the reversal of association might easily be explained. Given the dynamic of the country, whites have greater access to medical aid than blacks, and a majority of the follow-up for this study was done on a black population. So, when the racial groups were combined, a relatively small percentage of the subjects had access to medical aid.

►   As discussed, at the five-year follow-up, very few whites agreed to attend the screening exams.

Referring back to the social dynamic of the country, the lack of representation from the white population at follow-up could be due to a feeling of having little to gain from participating in the study. On the same token, black participants may have valued the medical check-up and screening provided to the children in the study.

As this example excellently illustrates, a number of "lurking" variables can contribute to instances of Simpson's Paradox.