The Residual Fallacy Revisited

D. McManaman

The residual fallacy is a statistical fallacy still prevalent today with respect to a number of issues, such as the claim that women are still paid less than men for doing the same work. We might also include here as well the very notion of "white privilege" as an example of this fallacy. The residual fallacy occurs when it is assumed that all unexamined factors that might contribute to an explanation of a particular phenomenon are equal, so that all remaining differences in outcome can be attributed to discrimination, or white privilege, etc.

Allow me to explain using the conditional reasoning we studied in class. We know that the logic of the scientific method requires testing a conclusion, because the method is fundamentally inductive. This means we begin with the evidence and proceed to account for that evidence through conjecture. However, corroborating evidence does not prove a hypothesis; rather, confirming a hypothesis leaves a conclusion underdetermined and thus uncertain. Hence, rigorous testing is required in order to determine which alternative possibilities are the most plausible. You are familiar with the following argument:

p ⊃ q
q
∴p

The "q" in this case stands for "the facts in evidence". And so,

If I come down with the flu, then I will have a fever.
I have a fever (facts in evidence)
Therefore, I have come down with the flu.

As you know, this is invalid (the fallacy of affirming the consequent). The conclusion is possible, but not necessarily the case. There are other possible causes for a fever besides the flu; for example, I may have a kidney infection, or a lung infection, etc.

In the context of scientific experimentation, we speak of *independent* and *dependent* variables. If we were to put this in the context of the conditional syllogism, it would look like the following:

If (independent variable), then (dependent variable)
Dependent variable
Therefore, independent variable.

For example,

I've been experiencing stomach pains recently (facts in evidence). My wife tells me that it is the result of the cookies I've been eating, which are high in fibre. Hence,

If I eat cookies that are high in fibre, then I will experience stomach pains (p ⊃ q) .
I am experiencing stomach pains (q).
Therefore, it is the result of eating cookies that are high in fibre (p)

This argument obviously requires testing. The independent variable in this case is **the high fibre cookies (p)**; the dependent variable is the **symptoms**, that is, **the stomach pains (q)**.

The problem is there are other factors that can account for the dependent variable (the facts in evidence), so to test this particular hypothesis (p), we must control for these other variables, that is, we must try to reduce the effect of **confounding variables (third variables).** Now, a confounding variable is a variable that influences both the dependant variable (stomach pains) and the independent variable (cookies), and if a researcher fails to control or eliminate them, the results of the research are simply unreliable--the cause of the stomach pains may be that third variable.

For example, the doctor questions me and he realizes that when eating these high fibre cookies, I usually do so while drinking Coke Zero. Coke Zero is a possible confounding variable. My drinking Coke Zero might influence my decision to eat cookies, and the Coke Zero might be causing my stomach pains, not the cookies; hence, eating cookies and stomach pains might be correlated, but the one is not necessarily the cause of the other. And so this variable, the Coke Zero, needs to be controlled or eliminated. The reason is that if I stop eating the cookies and the result is that I no longer experience stomach pains, the researcher might conclude that the cookies are the reason for my stomach pains. But it might very well be the Coke Zero, which I only drink when eating those cookies.

The effects of all other **variable predictors** need to be taken into account as much as possible. Other variable predictors might include: grains, sugar, certain dairy products, fish, processed foods, turkey, etc. These might also cause stomach pains.

What we want to do now is test the effect of changing the independent variable on the dependent variable. To test whether the Coke Zero is causing my stomach aches, I am instructed to stop drinking Coke Zero, but continue eating the cookies. If my stomach continues to ache, we can rule out the Coke Zero as a factor, at least for now. I am then instructed to stop eating the cookies. If my stomach ache continues, then we can rule out the cookies, at least for now.

What happened is that when I stopped drinking all diet soda (Coke Zero, etc.) while maintaining every other routine (i.e., I continued to eat the cookies), the stomach pains stopped. Our conclusion is that the high fibre cookies are an **irrelevant variable.**

Another example: I find that I break out in pimples about four times a year. I also notice that my breakouts are preceded by chocolate binges. Hence, I conjecture that the independent variable is the chocolate. I formulate the following major premise:

<If I eat chocolate, then I break out in pimples>

In terms of the conditional syllogism, it would look like this:

If I eat chocolate, then I will break out in pimples
I break out in pimples 4 times a year
Therefore, chocolate is the reason I break out in pimples.

The argument needs to be tested because it is deductively invalid (affirming the consequent). To do so, we want to assess the effect of manipulating the independent variable (chocolate) on the dependent variable (pimples). We are instructed to abstain from chocolate, and we discover that we continue to break out in pimples four times a year. We conclude that chocolate is an **irrelevant variable**. The effect of pimples must be influenced by other factors. We notice that the periods in which I break out in pimples coincide with exam time, and I tend to get stressed at these times. I also tend to eat a lot of chocolate during stressful periods of my life. So, perhaps the stress of exams causes me to crave chocolate and it causes me to break out in pimples (stress, in other words, might have been a **confounding variable**). We know that chocolate is not a factor, so perhaps stress is a factor.

So we manipulate the independent variable (stress) to see how it affects the dependent variable (pimples). I went one full year without the stress of exams, and the result was that I did not break out in pimples. This, of course, does not prove that stress was the reason for my breakouts, but it certainly strengthens the hypothesis.

Control variables must be held constant in order to discern the relationship between independent and dependent variables. If I am comparing the flights of two badminton birdies made by two different companies, I must hold constant certain control variables if my conclusion is to have plausibility. For example, if I were to test one birdie on a bright and sunny day, without the slightest breeze in the air and the other birdie on a cloudy and windy day, my conclusion about their quality would be unwarranted; for the differences in their flight might be due to factors other than the quality of the birdie, such as the weather, or even the person testing the birdie, the kind of racquet he is using, etc. So the weather is a **control variable** that must be held constant in this case. Another control variable is the athlete testing the birdie. We cannot have one athlete testing the birdie on one day, and a different athlete on another. So too, we cannot use on badminton racquet on one day, and different quality racquet on another. We must test the birdies while holding constant the control variables: same weather conditions, same athlete, same racquet, etc. If the flights of the two birdies are significantly different within the same conditions (with the control variables held constant), we know the flight difference was not due to the differences in weather, for there were no differences in weather.

Now, the claim that income disparity between men and women is the result of unjust discrimination must undergo the same kind of testing. The argument looks like this:

There is income disparity between men and women.
Hence, women are victims of income discrimination in the workforce.

The formal argument looks like this:

If there is income discrimination against women in the workforce, there will be income disparity between men and women.
There is income disparity between men and women.
Hence, women are victims of income discrimination in the workforce.

The entire conditional argument is invalid; for there are a number of other possible independent variables that might account for the income disparity. In other words, discrimination might be an **irrelevant variable** (i.e., if income disparity persists after removing all discrimination, the latter is an irrelevant variable). We must test for other **variable predictors**, such as quantity of education, quality of education, type of degree one has, years of experience (often interrupted by maternity leave), skill level, number of hours worked on a weekly basis, etc. If income disparity disappears when we control for certain confounding variables, then discrimination becomes an irrelevant variable (just as when we controlled for stress by removing all exam stress, pimples vanished, rendering chocolate an irrelevant variable).

Control variables must be the same for our conclusion to be warranted. For example, consider the following argument:

If white Americans are genetically superior to Africans, then white people will score higher on IQ tests.
White Americans have scored higher on IQ tests than African kids.
Therefore, white Americans are genetically superior.

Immediately one suspects that we need to control for other variables, such as education, language, and other cultural factors. The **residual** is what is left over after other such factors have been controlled. But at what point do we stop

controlling? Often, the point chosen is arbitrary, or worse, emotionally motivated. A racist, for example, may stop controlling too soon because he wants to believe that whites are genetically superior.

In the discussion of whether discrimination is the reason for income disparity between men and women, we will refer to the two groups under consideration as group M (males) and group F (females). Although two groups may be made up of those with college degrees, there may be other variable predictors discovered through a more precise analysis of the category than the very general category of "college degree".

And it is not enough to hold years of education constant, since there are not only quantitative differences in education (i.e., some programs are 4 years, other 3 years, others 2 years), but qualitative as well (a law degree from Coconut College is qualitatively inferior to a law degree from Princeton). Qualitative differences in education are a variable predictor. Quality can be measured by academic performance (John earned all A's, Betty earned all B's), the ranking of the university (Princeton is top ranked, Coconut College is last), or the difficulty and remuneration factor in the particular field of study (i.e., a math degree will earn you more than a degree in Environmental Studies). In other words, it is unreasonable of me to expect a master's degree in Medieval literature to equal a master's degree in electrical engineering in terms of remuneration capacity. Most industries have no use for a person with a graduate degree in Medieval literature but have use for one with a master's in engineering (that is why English majors are more likely to be waiting on tables after they earn their degree; engineers are usually working as engineers).

The category of "college educated" women and men is problematic from another angle. The reason is that college graduates include people who go on to postgraduate study, and this too influences income. More people in group M may choose postgraduate studies than those in group F. In fact, the ratio of women and men with postgraduate degrees differs from the ratios of those with college degrees. At the bachelors degree level, women outnumber men, but men outnumber women by more than a two to one ratio at the master's level and by 59% at the PhD level. So comparing college educated men and women, which includes those who have gone on to pursue postgraduate work, we are really comparing apples and oranges. In other words, the category "college-educated" lacks precision.

If we wish to compare men and women at the PhD level, we discover once again disparities between men and women, and changing ratios. Women receive 37% of all PhDs; moreover, the areas of study differ significantly from those of the 63% of males who receive a PhD. The PhDs which men receive tend to be more heavily concentrated in math and science and other fields of greater remunerative capacity. Women received almost half of the PhDs in the social sciences, and more than half in the area of education. Men received more than 80% of the PhDs in the sciences and more than 90% in engineering. Even the social sciences are not equally remunerative. For example, two people with a social science degree may show a difference in income by virtue of the fact that one degree is in sociology while the other is in econometrics--an a person with a degree in econometrics will make more than one with a sociology degree, and more men study econometrics than do women.

And so it is simply not the case that the disparity of income between men and women is due to discrimination any more than the disparity of income between me and Oprah Winfrey is due to a pervasive discrimination against philosophy teachers.

The Residual Fallacy and White privilege

The claim is that the very fact that a majority are successful and white, while a minority are less successful and "not white" is simply the result of white privilege. To test this claim, let us simplify society into two groups: East Indian minority and white majority. Furthermore, let us say that the Indian minority makes up 30% of the population, which leaves the white majority at 70%.

Now, let's say that 20% of the Indian population are very successful (2 out of 10). Since the Indian population constitutes 30% of the society that is a white majority, it follows that 6% of the entire population is East Indian and very successful.

Let's say that 50% of the Indian population are middle class (5 out of 10). That's 15% of the entire population is East Indian and middle class.

Let's say that 30% of the East Indian population are poor (3 out of 10). That's 9% of the entire society are poor East Indians.

In the white population, which is the remaining 70%, 10% are white and very poor (1 out of 10), or 7% of the entire population, while 60% of the white population are middle class, or 6 out of 10, which means 42% of the entire population are white middle class (since 70% of the population is white). And finally, we'll say that 30% of the white population are very rich, or 3 out of 10. That means 21% of the entire population is white and very rich.

Looking at the entire population (table 1):

Indian minority (30% of population)	**White majority (70% of population)**
6% = very successful	21% = very successful
15% = middle class	42% = middle class
9% = poor	7% = poor

Out of their own respective population (table 2):

Indian minority (30% of population)	White majority (70% of population)
20% = rich	30% = rich
50% = middle class	60% = middle class
30% = poor	10% = poor

As we can see from table 2, a greater percentage of East Indians are poor (⅔ more), and a greater percentage of whites are rich (⅓ more) and middle class (17% more).

Now, the residual fallacy consists in assuming that all unexamined factors that might contribute to an explanation of a particular phenomenon are equal, such that all remaining differences in outcome can be attributed to discrimination, or white privilege.

But such a conclusion needs to be tested. Other possible variables include: educational backgrounds, educational standards (i.e., medical schools in India are not the same as medical schools here, different degree requirements, which requires extra schooling here in Canada, etc.), different educational opportunities in India, so those who emigrate here have a different distribution of low skilled versus skilled labor; initial language barriers, lack of connections (family or business connections), the time it takes to acquire credit, or differing initial economic conditions (i.e., my parents arrived from India with nothing, so they did not inherit anything from anyone, while your parents inherited money when their parents died, which permitted you to open your own business, because your parents put that inheritance in savings, etc.).

Consider this last condition: my parents arrived from India with nothing, so they did not inherit anything from anyone, while your parents inherited money when their parents died, which permitted you to open your own business, because your parents put that inheritance in savings, etc. Is this white privilege? It is privilege, but it is not essentially white. Whiteness is incidental. It is simply the result of circumstances belonging to one country (Canada) that are different from those of another country (India).

We can compare your parents with the parents of that white student, but we can also compare your parents with the parents of another white student whose parents did not inherit anything. Both are white, but both do not enjoy the same level of privilege. Thus, privilege, like income, is circumstantial, while “white privilege” is categorical. The residual fallacy in this case involves assuming that the distribution of “privilege” is due to the color of one’s skin, or the “category” to which one belongs by virtue of one’s color. However, there are a multitude of factors that prevent a large percentage of white people from “privilege”, namely, a poor work ethic, poor choices, a sense of entitlement, lack of talent, drug use that has had adverse effects on brain development, etc. Such factors also keep some members of a minority group from achieving a level of “privilege” available to them in this country. But it has been factors opposite those previously listed, namely, a good work ethic, hard work, character, talent, persistence and good decisions, etc., that account for the level of privilege that many people currently enjoy. Indeed, some who enjoy privilege may not fit that description, but may have inherited all they have been given. However, such privilege cannot be sustained without those factors.

Now once a generation has passed and the language barrier is no longer a factor--because the children of East Indian parents went to public school and now speak fluently and they have the same educational opportunities as white children, and are achieving comparably with white children and in many cases better than the white students, since the Indian work ethic is significantly better,-- we should see a change in the differences in percentages in Table 2. They should eventually equal out at the very least. If white privilege is nothing more than a fallacy rooted in prejudice, then not only will they equal out, they might even tip in favor of the minorities. And this is just what has happened among the asian population. In 2001, 31% of Chinese in Canada--both those who were born in Canada and those who were foreign born--had a university education, while the national average was 18%. Furthermore, Chinese who immigrated to Canada in the 90s and who were of prime working age eleven years later had an employment rate of 61%; but the national average was 80%. In this case, it was the recognition of foreign qualifications that was the primary issue. But the employment rate for Canadian-born Chinese men who were of prime working age was slightly above the national average (86%), and the employment rate for Canadian born Chinese women of prime working age was 83%, which was higher than the national average, which was 76% (Chinese Canadians: Enriching the cultural mosaic," Canadian Social Trends, Spring 2005, no. 76)