

Statistical Deception

- Lying with statistics
- Putting a “positive spin” on the facts

Statistical deception is not necessarily a bad thing, but you need to be aware of it rather than just accepting statistics at face value.

Many problems with statistics involve problems with gathering the data:

- 1. Non-representative samples**
 - too small
 - too large
 - not randomly chosen
 - convenience sample
 - purposely chosen wrong
- 2. Generalizing to the wrong population**
 - sample is drawn from a different population than the results imply
- 3. Comparing apples and oranges**
 - groups being compared were different to begin with
 - difference is due to something other than the results imply
- 4. Survey bias**
 - leading questions
 - **“Good boy” effect**
 - People will give the answer they think you want to hear.
 - **“NOYB” effect**
 - The more personal a question is (the more it is “none of your business”), the more likely people are to lie.
- 5. Placebo effect**
 - In medicine a placebo is a fake treatment actually helps because people **think** it will work.
 - Doesn't have to deal with medicine.
 - In general, when people think they are being watched or treated, they often act differently than they would otherwise.

Other issues can come up in interpreting and publicizing statistical results:

- 1. Convenient averages**
 - Choosing the average that makes you look the best (or your opponent look the worst), even if it's not really a “typical” average
 - ...Is the number they give really “average”?
- 2. Assuming everybody is average (or close to it)**
 - ...How spread out are things?
- 3. Not adjusting for different sample sizes**
 - Comparing raw numbers instead of percentages
- 4. Not adjusting money amounts for inflation**
- 5. Screwing up the math**
- 6. Extrapolating from a partial result.**

7. Assuming cause and effect
 - Remember: correlation just means “relationship”
 - A confounding variable may be skewing the results.
8. **Ignoring Occam’s Razor**
 - complex or unbelievable explanations
 - Occam’s Razor says the simplest explanation is generally the best.
 - Always consider the simplest explanation first.
9. Changing the subject
 - a.k.a. “Moving the bullseye to fit the arrows”
 - saying a result means something different than it really does
 - putting a “good spin” on the data
 - finding one small thing about the results that supports what you want to find
10. Reporting information from biased sources that have a vested interest.
 - Always ask “Who says so”?
 - Try to get information from neutral parties who don’t have a stake in the outcome.
11. Using a non-standard significance level
 - deciding after the fact on a level that guarantees significance
12. Misuse of the word “significant”
 - implying significant means big, important, or dramatic
 - REMEMBER: it just means “unlikely to have happened by chance”
13. Discounting significance because something is “just statistics”