

Replication of Study 2 by Walsh & Sloman (2004)

Erin Bennett
erindb@stanford.edu

Introduction

This study is about how people revise their beliefs when they are presented with a counterexample to a causal relationship.

There are a few questions:

1. **Are other effects screened off from the effect in the counterexample, given that the cause happened?** That is, do people update their beliefs about that particular causal relationship in isolation, without influencing other parts of their theory? Or do they update their theory in a more complex way such that in this and other counterexamples, other things that usually result from the cause are also less likely? E.g. “My new neighbors tell me they are in a fraternity. I think they will have loud parties late at night and their apartment will be messy. If after a few weekends, the apartment is always really quiet, my expectation might be less strong that it’s messy.”

Walsh & Sloman found that effects are not screened off and that the counterexample results in people expecting other possible consequents less.

Comparison 1: For “A causes B” and “A causes C”, ask participants about two situations:

- 1) A is true. How likely is C?
- 2) A is true and B is false. How likely is C?

If B and C are screened off from one another given A, then both questions should have the same answer. If participants are revising their theories in a more complex way, they might be different. In fact, the answer for 2 tends to be lower than the answer for 1.

2. **Are other effects screened off, given that people’s explanation of the counterexample doesn’t hold?** That is, do people revise their beliefs by hypothesizing one explanation for the counterexample (the cause doesn’t really hold, something happened to prevent the effect, some necessary condition wasn’t met, etc.) and using that explanation to make new predictions? E.g. “Maybe they’re in a really nerdy fraternity. If another fraternity replaces them that isn’t nerdy, the new fraternity will probably have loud parties and messy apartments.”

Walsh & Sloman found that other effects *are* screened off given the cause is true *and* the explanation the participant generated for the counterexample is false.

Comparison 2: Ask participants about two situations:

- 1) A is true and B is false.
 - a. Why do you think that is?
 - b. Also, how likely is C?
- 2) A is true and whatever explanation you gave for the situation in 1 is true. How likely is C?

If B and C are screened off from one another given the explanation for the counterexample, then both questions should have the same answer. In fact, they do.

3. **Do people revise their causal beliefs in the face of a counterexamples?** That is, are these exceptions, conditions, etc. already known and accounted for and participants just have enough evidence for them when they see the counterexample? Or are they adding them in as needed? E.g. “If a new

fraternity replaces my neighbors that may or may not be nerdy (I just don't know), then will my experience with my quiet neighbors influence my expectations about the new fraternity?"

Walsh & Sloman found that people have report different probabilities of effects given causes for (almost) the same question after the counterexample.

Comparison 3: Ask participants about $P(C|A)$ before and after the counterexample.

- 1) A is true. How likely is C?
- 2) A is true and you don't know whether that explanation you gave for the counterexample situation is true. How likely is C?

If beliefs are not revised, you would expect no difference between these responses, because 2 has no more information than 1 (literally, at least). But in fact, responses to these questions tend to be different.

Methods

Power analysis

Planned sample

English-speaking participants on Amazon's Mechanical Turk.

Materials

Walsh & Sloman (2004) report using 6 different stories. Two are used as examples in their 2004 paper. Two more are used as examples in a 2008 report of the same experiment in a book chapter.

Stories:

- Jogging regularly causes a person to lose weight. Jogging regularly causes a person to increase their fitness level.
- Worrying causes difficulty in concentrating. Worrying causes insomnia.
- Following this diet causes you to have a good supply of iron. Following this diet causes you to lose weight.
- Playing loud music in Paul's apartment causes the neighbors on his left to complain. Playing loud music in Paul's apartment causes the neighbors on his right to increase the volume on their TV.

Procedure

Each participant sees each story. Stories are presented one at a time, in blocks.

In each block, first the causal statements in question are presented. These will be visible for rest of the block. Then the following questions are asked (in order):

- 1) A is true. How likely is C?
- 2) A is true and B is false.
 - a. Why do you think that is?
 - b. Also, how likely is C?
- 3) A is true and you don't know whether *[the explanation you gave for 2a]* is true. How likely is C?
- 4) A is true and *[the explanation you gave for 2a]* is false. How likely is C?
- 5) A is true and *[the explanation you gave for 2a]* is true. How likely is C?

Analysis plan

Run t-tests to see whether the following replicate:

- *Comparison 1: Answer2a < Answer1*
- *Comparison 2: Answer2a \approx Answer5*
- *Comparison 3: Answer1 \neq Answer3*

Differences from original study

The sample population of the original study was a group of college students. My sample will be participants on Amazon's Mechanical Turk. Both groups probably reasoning similarly about causes. However, it is possible that workers on Mechanical Turk will be more concerned about giving consistent answers than college students, so for similarly phrased answers we might expect a slight bias for this new population giving the same answer. Therefore, replicating *Comparison 1* and *Comparison 3* might be more informative than replicating *Comparison 2*.

The setting will be online rather than in a lab, which means that participants might be paying less attention, or take breaks in the middle of the study. I don't anticipate that having a strong effect on the results, since these questions don't seem to require a lot of concentration.

I will use some basic natural language processing to repeat back participants' explanations with (hopefully) appropriate syntax. The original study used a human to do this task. Because doing this with a computer is more error-prone, I will have to ask the question in a slightly different way, giving participants a particular syntactic frame to fill in. The resulting program might respond to people with nonsense. It might be good if I could give participants an option to tell me whether the question I'm asking is nonsense so that I can exclude that data.

Method Addendum

Actual sample

Differences from pre-date collection methods plan

Results

Data preparation

Confirmatory analysis

Exploratory analyses

Discussion

Summary of replication attempt

Commentary