the WINNOWER

SOCIAL SCIENCES

Modus Tollens Bitches

SIMINE VAZIRE [1]
1. UC Davis

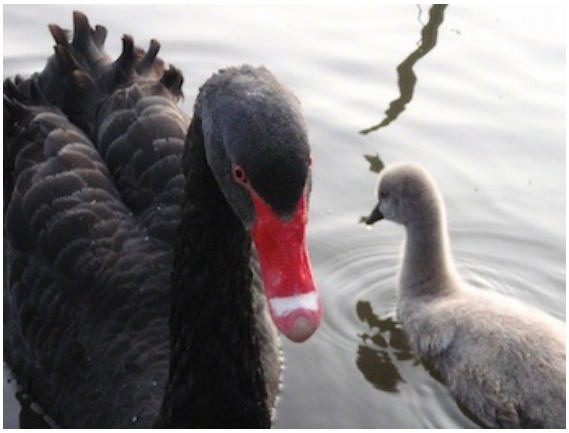
READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:
simine@gmail.com

DATE RECEIVED:
June 10, 2015

KEYWORDS:
reproducibility, scientific integrity, research methods, philosophy of science



i am not going to say anything original in this post.

my philosophy of science friends tell me that i should abandon popper. before i do, i'd like to have one last fling with him.

i am teaching two research methods classes this quarter, so i'm teaching, you know, how scientists are objective and our theories are falsifiable and we seek to disconfirm them. (hello psc41 students!)

today's topic: modus tollens and falsifiability.

for those of you who don't remember day 2 of research methods, here is modus tollens (aka, denying by denying, a valid form of logical argument):

why is modus tollens important for empirical research? i am cribbing from meehl 1978 here (who himself was stealing from others, including of course popper), but basically, it's how we make our theories falsifiable. we make risky predictions of the sort:

if my theory is correct, then i should observe a certain pattern of data.

if the data show what we predicted, then our theory is left standing. of course it doesn't prove that our theory is correct, that would be affirming the consequent, and we all know that's a logical fallacy.* but we do have a tendency to say that the data are consistent with our theory (true), and even that the data support our theory (maybe less warranted, depending on how risky the prediction was - if a lot of

other theories are also consistent with the data, then i'm not so sure the data really support our theory). basically, if the data come out as predicted, we celebrate and claim to have found evidence for our theory.

the interesting part comes when the data don't come out as predicted. according to modus tollens, if we find ~q (not the predicted results), we should conclude ~p (our theory was wrong). let's put it this way:

if my theory is right (T) then i should observe these data (D)
i observe data that are not those i predicted (~D)

therefore, my theory is wrong (~T)

that's how it should work, and that's what makes our theories falsifiable.
but, as meehl pointed out, what we are really testing is something like the following:

if my theory (T) is right, my hypothesis (H) is appropriately derived from my theory, and my methods (M) are sound, then i should observe these data (D)

or:
if T.H.M ---> D

so now what happens when we observe ~D? well, we no longer have to conclude ~T, because we have two other options: ~H and ~M. that is, instead of concluding that our theory is wrong, we can throw the hypothesis or the methods under the bus.

~M is especially tempting, because all it requires is to say that there was something wrong with the methods. it's a very easy escape route. blame the subject pool. blame the weather. blame the research assistants.**

as etienne lebel and kurt peters pointed out, also echoing meehl, in psychology we are often quite happy to throw our methods under the bus. that's because we are measuring things that are very difficult to measure, and also because we probably don't pay enough attention to the validity of our methods. in a way, there is a perverse incentive to gloss over the methods - the less well-defined and rigorous they are, the easier it is to escape the ~T conclusion by escaping through the ~M door. if we bolster our methods, and think of all the possible pitfalls ahead of time, we block that escape route and are left in the uncomfortable position of risking falsifying our hypothesis or our theory. god forbid.***

this is a problem for (at least) two reasons.

1. when a researcher is conducting original research, it is too easy to ignore null results by attributing them to a quirk in the methods. there is always some rationale for sticking a null result in the file drawer, so long as you can point to some potential flaw in the method.

2. when a researcher is conducting a replication of someone else's work, it is too easy to ignore null results by attributing them to a quirk in the method. it's even easier here, because you don't even need to point to a specific potential flaw in the method - you can just say that the researcher lacked expertise, or did not have the tacit knowledge necessary.

what can we do?

1. improve our methods. if we really want our theories to be falsifiable, we need to close the ~M loophole. think ahead of potential flaws in your design, and only run the study once you are fairly certain that you will believe the results, even if they are null. of course we can't anticipate every flaw, and some results legitimately belong in the file drawer. but we can shrink the loophole by using well-validated measures, a large sample, thinking ahead of potential moderators (and pre-registering them

when possible).

2. don't throw the methods under the bus. consider the possibility that the theory or hypothesis is wrong. make it your new year's resolution: i will not write off a null finding or a failed replication unless there is a glaring error -- one that my enemy's mother would agree is an error.

3. when publishing your original results, provide enough detail about all important aspects of the method so that if someone attempts a replication following the procedure you describe, you cannot blame their result on poor methods. the onus is on us as original researchers to specify the things that matter. of course some are too obvious to specify (e.g., make sure your experimenter is wearing clothes), but these are covered by the 'glaring error that your enemy's mother would agree with' clause in #2. if you don't think another researcher in your field could competently carry out a replication with the information you publish, that's on you. if your procedure requires special expertise to carry out, specify what expertise. if you can't, your theory is not falsifiable.

still 3. of course, sometimes a failed replication makes you think of a potential moderator you hadn't thought of before. good for you. i don't really care, unless you are willing to go out and test that moderator. show me a confirmatory study in which the moderator does its moderation thing, and i will believe you. otherwise, your modified theory is as good as any untested theory.

but never mind, popper is dead. of course i know**** that naive falsification can't be right because of the problem of compensatory adjustments and the probabilistic nature of science. but still, aiming to make our theories more falsifiable***** seems like a good idea to me.

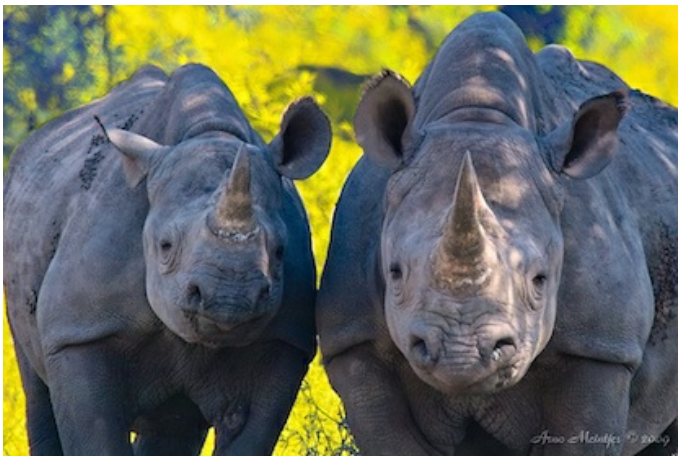
* you're a logical fallacy.

** don't blame the research assistants.

*** this week i managed to convince 94% of my class that they should switch doors in the monty hall problem. this is not actually related to my blog post, but i wanted to tell you about it anyway. to be fair, 27% thought it was the right strategy at the beginning of class, so i only get credit for 67%. still, i am pretty pleased with myself.

**** thank you wikipedia.

***** i can feel the philosophers cringing.



why do the black swans get all the attention?