DOI: 10.1037//0003-066X.57.1.69

Fisher's Fallacy and NHST's Flawed Logic

Stefan G. Hofmann
Boston University

Krueger (January 2001) recently presented an interesting discussion on the logical validity of null hypothesis significance testing (NHST). He stated, "NHST simply suggests that the results need not be attributed to chance. It suggests that 'there is not nothing'. . . . Such an inference is a probabilistic proof by contradiction (*modus tollens*)" (p. 17). Krueger further explained that the "key concern about this chain of inference is that deductive syllogisms are not valid when applied to induction" (p. 17). Without much further elaboration, he quickly dismissed this argument because the pragmatic value of NHST far outweighs these rather dubious philosophical concerns.

Unfortunately, the underlying philosophical problems with NHST are considerably more complicated and significantly more serious than Krueger (2001) seemed to realize. As I describe in more detail in the following paragraphs, the logic of NHST is based on a misapplication of deductive syllogistic reasoning, because probabilistic statements are incompatible with the rules of deductive reasoning (no matter how practical they are).

Philosophers distinguish between two types of syllogistic arguments: *deductive* and *inductive arguments.* An inductive argument is such that if its premises are true, its conclusion is likely (but not necessarily) true. In contrast, a deductive argument is such that it is impossible for all the premises of the argument to be true while the conclusion of the argument is false. In other words, deductive arguments possess *formal validity.* An argument is formally valid if, and only if, its conclusion (*C*) must be true whenever its premises (*P*1 and *P*2) are true. In addition, an argument is considered *sound* if it has true premises.

Philosophers distinguish among different rules of deductive reasoning to prove that certain arguments are valid. These rules include the *modus ponens* (affirming the antecedent), *modus tollens* (denying the consequence), deductive syllogism, and chain argument. On the surface, NHST appears to be based on the *modus tollens*:

*P*1: If *P* (H_0, or the null hypothesis, is true), then *Q* (datum should not occur).
*P*2: Not *Q* (the datum did in fact occur).
C: Not *P* (H_0 is not true).

The same argument is present in the following example by Cohen (1994):

*P*1: If a person is a Martian, then he is not a member of Congress.

*P*2: The person is a member of Congress.

C: Therefore, he is not a Martian. (p. 998)

This argument is both formally valid and sound (if one believes in the existence of Martians). The problem arises if (a) the premises are not true and/or (b) one makes probabilistic rather than absolute premises. In the former case, an argument might be formally valid but not logically sound. In the latter case, the premises might seem plausible, but the argument is invalid because the concept of probability is introduced. For example, the following argument by Cohen (1994) is sound (i.e., has true premises) but is formally invalid because it contains probability statements that lead to false conclusions:

*P*1: If a person is an American, then he is probably not a member of Congress.

*P*2: The person is a member of Congress.

C: Therefore, he is probably not an American. (p. 998)

This example illustrates that NHST is based on a faulty conceptualization of formal logic. Although it might lead to sensible conclusions in many cases, it can also lead to wrong conclusions in others. However, formal validity and not common sense defines deductive reasoning. An argument can be formally valid only if its conclusion is true whenever its premises are true. By making an argument probabilistic, it becomes possible that its conclusion is false although all of its premises are true. Therefore, a syllogism that is probabilistic by definition does not meet criteria for formal validity because the conclusion is not necessarily true whenever its premises are true. The same argument applies to the *modus ponens* and to the chain argument, which are also implicitly used when interpreting statistical findings.

I agree with Krueger (2001) that Bayes's theorem is one way to reduce the error rate of NHST and to encourage more risky research (see also Hofmann, 1999). However, it does not correct the flawed logic of NHST. NHST is still psychologists' primary method when conducting research probably, for the most part, because of convention (which can be changed). Effect size estimations, power calculations, and confidence intervals are the real solutions, in my opinion. Bayes's theorem might help psychologists to enhance the riskiness of their hypotheses and the verisimilitude of their theories (Popper, 1959). However, to initiate real progress in psychology (e.g., Meehl, 1978; Rosenthal, 1995), researchers will need to move away from the crude and flawed there-is-probably-not-nothing method of NHST toward the measurement and quantification of psychological effects. I doubt that the Human Genome Project could have happened if Fisher had trained Gregor Mendel in empirical methods.

REFERENCES

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49,* 997–1003.

Hofmann, S. G. (1999). Introducing the grandmother test into psychological science. *Journal of Theoretical and Philosophical Psychology, 19,* 167–176.

Krueger, J. (2001). Null hypothesis significance testing: On the survival of a flawed method. *American Psychologist, 56,* 16–26.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Popper, K. R. (1959). *The logic of scientific discovery.* London: Hutchinson.

Rosenthal, R. (1995). Progress in clinical psychology: Is there any? *Clinical Psychology: Science and Practice, 2,* 133–150.

Correspondence concerning this comment should be addressed to Stefan G. Hofmann, Department of Psychology, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215. E-mail: shofmann@bu.edu