

---

Joachim Krueger (1998) Theoretical Progress Requires Refined Methods and Then Some. *Psychology*: 9 (73) Social Bias (10)

**Volume:** 9 ([next](#), [prev](#)) **Issue:** 73 ([next](#), [prev](#)) **Article:** 10 ([next](#) [prev](#) [first](#)) **Alternate versions:** [ASCII Summary](#)

**Topic:** Social bias

[View Topic](#)

**Article:** 10) Krueger 9(73) Theoretical Progress Requires Refined

[View Article](#)

---

PSYCOLOQUY (ISSN 1055-0143) is sponsored by the American Psychological Association (APA).

---

# THEORETICAL PROGRESS REQUIRES REFINED METHODS AND THEN SOME Reply to Ruscio and McCauley on Krueger on Social-Bias

*Joachim Krueger*  
*Department of Psychology*  
*Brown University, Box 1853*  
*Providence, RI 02912*  
<http://www.brown.edu/Departments/Psychology/faculty/krueger.html>

[Joachim\\_Krueger@Brown.edu](mailto:Joachim_Krueger@Brown.edu)

## Abstract

Can research on social-perceptual biases benefit from improved and diversified statistical methods? Having reached the brink of nihilism, I conclude that (a) any point-hypothesis can be rejected by null hypothesis significance testing (NHST), (b) any such hypothesis can be accepted by Bayesian inference, (c) effect size estimates are meaningful only if that meaning is imported from extra-statistical considerations, and (d) taxonomies of biases and their causes will be messy because most biases are overdetermined.

## Keywords

*hypothesis testing, Bayes' Rule, effect sizes, projection*

---

1. Ruscio (1998) shares many of my concerns about contemporary research on social-perceptual biases. He emphasizes the importance of effect size estimates, the classification of biases by their presumed source, and the study of individual differences in both coherence and correspondence of judgment (after Hammond 1996). In other words, Ruscio calls for further methodological refinements, hoping that these

will ultimately yield theoretical progress and a technology for bias reduction. McCauley (1998) believes that social psychologists' concerns with judgmental biases and their statistical methods of choice are largely ideological. According to his view, the world is a terrible place which social psychologists try to understand by demonstrating, time and again, the limitations of human perception and decision making. McCauley suggests that the information-processing paradigm cannot address the most pressing issues at hand (e.g., ethnic strife, multiculturalism, etc.). To believe that it could would be naively optimistic.

2. Although both Ruscio and McCauley express dissatisfaction with the current research program in social cognition, their concerns fall on opposite ends of the spectrum of what one can be concerned about. Should we refine and press on (Ruscio) or start over (McCauley)? I sympathize with McCauley's view that the information-processing paradigm has created blind spots. Any dominant paradigm makes progress (by definition!) to the detriment of its alternatives. The emphasis on humans as information processors has pushed other metaphors to the background (e.g., humans as lawyers or moralists). It is not clear to me, however, how we can "get beyond error and bias to create a social psychology that has something to say in a world of cultural diversity and ethnic conflict, to groups who have real differences that they are not wrong to care about." To study these important topics, we need [again] a paradigm that tells us what questions to ask and how to analyze the data we collect. It is not clear yet how the alternative metaphors can accomplish that.

3. Ruscio's call for further methodological refinements is important if one chooses to press on in the existing paradigm. Two of the three refinements (reporting effect sizes, doing debiasing studies) were among the recommendations I offered at the outset (Krueger 1998a). The third, to classify biases according to their presumed psychological causes, presents an additional task. Although I recognize the appeal of this approach, I doubt that it is sufficient to move research to a new qualitative level. As an illustration, I turn again to research on social projection, because I know this area better than I know others. After rediscovering social projection as a demonstrable phenomenon, Ross, Greene and House (1977) set the research agenda by suggesting a variety of possible sources of this bias. Hundreds of studies followed to examine the role of these sources (e.g., the availability heuristic, stimulus attributions, ego-protective motives, etc.). All of these studies used null hypothesis significance testing (NHST) to ask whether the presumed causes (e.g., availability) would strengthen the phenomenon. They did. Many contributing causes were identified as being sufficient, but none as being both necessary and sufficient (Krueger 1998b). Inasmuch as judgmental biases are overdetermined, and inasmuch as research focuses on one potential cause at a time, no compelling classification of biases by source can be obtained. NHST does not just favor the detection of bias, it also favors the detection of presumed causes of bias.

4. Much of my critique of NHST focused on the consequences of statistical power. The more data one collects, the less likely they will be under the null hypothesis. Indeed, the more data one collects, the less likely they will be under ANY point hypothesis UNLESS the point is the mean of the sampled data. Consider a sample of 10 observations (0, 0, .1, .1, .2, .2, .3, .3, .4, .4) with a mean of .2 and a standard deviation of .15. Suppose these observations express some social perceptual bias (e.g., projection). When tested against the null hypothesis,  $H_0$  (zero projection), the effect is highly significant,  $t(9) = 4.24$ ,  $p = .002$ . When tested against an alternative hypothesis,  $H_1$  (e.g., a rational inductive degree of bias of, say, .3), the effect merely approaches significance,  $t(9) = 2.12$ ,  $p = .06$ . NHST suggests that  $H_0$  should be rejected and  $H_1$  retained. Using Bayes' Rule, one can ask how likely the hypotheses are given the data. Assuming that  $H_0$  and  $H_1$  were equally likely a priori, the posterior probability for  $H_0$  is  $p(H_0|D) = (p(H_0)p(D|H_0))/(p(H_0)p(D|H_0)+p(H_1)p(D|H_1)) = (.5*.002)/(.5*.002+.5*.06) = .03$ . The posterior probability for  $H_1$  is  $1 - p(H_0|D) = .97$ . In short, the Bayesian inferences are consistent with the decisions based on NHST. Now suppose sample size is doubled ( $N = 20$ ) without any change in other sample parameters. The  $p$  values are .000006 for the test against zero, and .006 for the test against .3. That is, NHST suggests that both  $H_0$  and  $H_1$  should be rejected. The Bayesian posterior probabilities

are .001 and .999 for  $H_0$  and  $H_1$ , respectively. Given the data, the alternative hypothesis (here: rational induction) has become far more likely than the null hypothesis (here: no projection).

5. This example illustrates that Bayesian inferences remain consistent when sample size increases, whereas categorical decisions derived from NHST do not. Nevertheless, the Bayesian posteriors become more polarized (e.g., the posterior probability of the alternative hypothesis rose from .97 to .999). Bayesian posteriors (e.g.,  $p(H_0|D)$ ) are as sensitive to sheer increases in sample size as the traditional  $p$  values derived from NHST (e.g.,  $p(D|H_0)$ ). For any point hypothesis, other than the one that is identical to the empirical mean,  $p$  values can be driven toward zero. Therefore, when two alternative hypotheses are used, as in the foregoing example, the  $p$  values for both approach zero with increasing sample size. This does not mean, however, that the posterior probabilities for both hypotheses approach zero. Instead, the probability of whichever hypothesis is closer to the empirical mean (by whatever small degree) will approach 1. If, for example, the empirical mean were .16, with  $H_0$  and  $H_1$  being 0 and .3, respectively, a sufficiently large sample would inspire near certainty in the truth of  $H_1$ .

6. It may be these kinds of sample size effects that explain the attractiveness of effect size indices (which are independent of  $N$ ). Ruscio argues that effect size estimates "provide richer and... more pertinent information" than Bayesian analyses. I hesitate to endorse this conclusion because it does not help us clarify the MEANING of an obtained effect size. Effect sizes come into play either within the context of NHST or independent of it. The first case, which has been recommended by the APA Task Force on Statistical Inference (1996), calls for reports of effect sizes that are significantly different from zero. If there is only a directional hypothesis (i.e., no a priori expectation about the size of the effect), any significant effect is reported, and every reported effect is considered equally relevant. That is, effect sizes are considered relevant if they can be distinguished from the null hypothesis. Whether they can be distinguished depends again on sample size (and on the precision of the measures). Ironically, then, the fact that the researchers cared enough to collect the necessary data to demonstrate the significance of a small effect reflects the perceived relevance of the effect. This also means that small effects tend to have smaller standard errors than large effects.

7. In the second case, distinguishing the effect size from zero is considered irrelevant. Because effect sizes are independent of sample size, the goal of statistical analysis is not to make inferences about hypotheses, but to obtain reliable point estimates. Here, the reliability of an estimate need not increase with its proximity to a point hypothesis that is to be rejected. Abandoning all hypotheses as benchmarks for comparison leaves open the question of when estimates are sufficiently reliable. Survey researchers and pollsters have apparently agreed that an error margin of  $\pm 3\%$  is acceptably low. No equivalent consensus in research on social judgment is in sight. To complicate matters, response scales and other behavior measures vary widely (rather than being the ubiquitous percentage estimates of the polling industry).

8. The question remains of what our measures tell us about the rationality of human judgment. The refinement of statistical methods ultimately does not satisfy. The appeal of rationality as a psychological attribute hinges in part on the objectivity of its defining standards. But most research is conducted to DENY respondents' rationality by using methods whose justification also rests on their presumed objectivity. In other words, we use rational methods to prove that our subjects think irrationally. Throughout this debate, my argument has been that the rationality (i.e., objectivity) of our methods is a negative one. In the areas of social judgment that I reviewed (projection, self-perception, attribution), these methods are guaranteed to work. They demonstrate irrationality because there is no positive view of what rational judgment is.

## REFERENCES

American Psychological Association (1996) Task force on statistical inference initial report.  
[www.apa.org/science/tfsi.html](http://www.apa.org/science/tfsi.html)

Hammond, K. R. (1996) Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice. Oxford University Press.

Krueger, J. (1998a). The bet on bias: A forgone conclusion? *PSYCOLOQUY* 9(46)  
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.46>  
<ftp://ftp.princeton.edu/pub/harnad/Psycology/1998.volume.9/psyc.98.9.46.social-bias.1.krueger>

Krueger, J. (1998b) On the perception of social consensus. *Advances in Experimental Social Psychology* 30: 163-240.

McCauley, C. (1998) The bet on bias is cockeyed optimism. *PSYCOLOQUY* 9(71)  
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.71>  
<ftp://ftp.princeton.edu/pub/harnad/Psycology/1998.volume.9/psyc.98.9.71.social-bias.9.krueger>

Ross, L., Greene, D. & House, P. (1977) The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology* 13: 279-301.

Ruscio, J. (1998) Applying what we have learned: Understanding and correcting biased judgment. *PSYCOLOQUY* 9(69) <http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.69>  
<ftp://ftp.princeton.edu/pub/harnad/Psycology/1998.volume.9/psyc.98.9.69.social-bias.7.krueger>

---

**Volume:** 9 ([next](#), [prev](#)) **Issue:** 73 ([next](#), [prev](#)) **Article:** 10 ([next](#) [prev](#) [first](#)) **Alternate versions:** [ASCII Summary](#)

**Topic:** Social bias

[View Topic](#)

**Article:** 10) Krueger 9(73) Theoretical Progress Requires Refined

[View Article](#)