

Ruscio, J. (1999). Statistical models and strong inference in social judgment research. *Psychology*, 10(27). Available on the World Wide Web: <http://www.cogsci.soton.ac.uk/psyc-bin/newpsy?10.027>

STATISTICAL MODELS AND STRONG INFERENCE IN SOCIAL JUDGMENT RESEARCH

Commentary on Krueger on Social-Bias

John Ruscio
Department of Psychology
Elizabethtown College
Elizabethtown, PA 17022
rusciojp@acad.etc.edu

ABSTRACT: In debating whether an asymmetry in traditional null hypothesis significance testing (NHST) biases empirical results against human rationality, the limited role that NHST can play in scientific reasoning has been overlooked. Platt's (1964) framework of "strong inference" illustrates the proper use of NHST and interpretation of its results, chiefly the ruling out of one potential alternative explanation for observed data (usually chance-level differences or associations). Especially in light of this limited role that NHST can play, it is critical to employ an appropriate statistical model. The inconclusiveness that can result from using an incorrect model is discussed in the context of social projection.

1. Krueger (1998a) proposed that the asymmetry of null hypothesis significance testing (NHST) stacks the deck in favor of rejecting human rationality, which is typically equated with the null hypothesis in social judgment research. Chow (1999) offered a defense of NHST based in part on the distinction between substantive and statistical hypotheses. I would like to elaborate upon this distinction in two ways. First, I situate the (limited) role of NHST within the framework of Platt's (1964) "strong inference." Second, I argue that we must pay closer attention to our choice of statistical model and how it corresponds to the substantive hypothesis of interest. I review the discussion of social projection (as opposed to a false consensus effect) in terms of the underlying statistical model that Stanovich (1998) and Krueger (1998b) have implicitly endorsed, which is inappropriate in light of the normative standard to which they both subscribe.

I. STRONG INFERENCE AND NHST

2. The distinction between substantive hypothesis—the claim that one wishes to test—and statistical hypothesis—ordinarily the "null" hypothesis of chance-level differences between groups or association between variables—is fundamental to statistical methodology in the behavioral sciences. In fact, this nonintuitive arrangement poses a formidable challenge in learning (and teaching) statistics. It is critical to recognize that NHST never directly evaluates the substantive

hypothesis, but can only help to rule out one of its competitors. In well-designed research this competitor may be another substantive hypothesis, although as Krueger (1998a) notes it is often merely chance, or a “fluke” result, that is being rejected. I feel that Krueger (1999) was too quick to dismiss the distinction between substantive and statistical hypotheses. No statistic is self-interpreting. Results contribute to a principled argument that is also based on considerations of design, sampling, measurement, and so forth (Abelson, 1995). To tether substantive conclusions to statistical tests may have the unfortunate consequence of absolving researchers from the responsibility of thoughtfully utilizing empirical results to inform substantive claims.

3. The notion of ruling out a competing explanation is at the heart of what Platt (1964) termed “strong inference.” He conjectured that scientific progress accelerates to the extent that researchers follow a two-part strategy: concocting as many plausible, nonredundant hypotheses as possible and then designing empirical research to systematically weed out incorrect hypotheses. In a Popperian sense, a particular hypothesis gains in verisimilitude as it survives additional risky tests.

4. Platt’s framework contrasts sharply with methods that focus on corroborating a theory simply by obtaining “statistically significant” results in a competition between the theory and chance. I suspect that this weak strategy is what Krueger (1998a) had in mind when he criticized social bias research and its tendency to equate “significance” with “irrationality,” a criticism that is well-deserved. However, the mathematics of NHST are not faulty: interpreted properly, they can help to rule out one potential explanation for observed data, that of chance-level deviations from the null hypothesis¹, whatever that may represent in a given test. Rejecting the null does not confirm any particular alternative, substantive hypothesis. If several plausible alternative explanations remain, an investigator must amass additional evidence to rule them out. Stanovich (1999), for example, reasons from patterns of individual differences to eliminate explanations that compete with “irrationality” in the literature on biases in social judgment, such as performance error or computational limitation.

II. STATISTICAL MODELS

5. To serve a worthwhile purpose within the strong inference framework, NHST must aim to rule out a plausible alternative explanation for observed data. When an inappropriate statistical model is chosen, even this limited role for NHST may be subverted. To illustrate this problem, I will describe the imperfect statistical models implicit in Stanovich’s (1998) and Krueger’s (1998b) discussions of social projection. First, a quick review of the subject is in order. Krueger (1998a) discussed the “false consensus effect” as a social bias that was initially detected through statistical testing against an inappropriate normative standard: using one’s own opinion as a basis for estimating the opinions of others—or “projecting” one’s opinion—was interpreted as irrational. Dawes (1989), among others,

demonstrated that some degree of projection is warranted. This analysis was empirically supported in experiments conducted by Krueger and Clement (1994) and Krueger and Zeigler (1993), each of which demonstrated that some degree of projection increased judgment accuracy².

6. Stanovich (1998) reported that “individuals high in projection tendency were in fact somewhat more accurate in their opinion estimates in our experiments—and they were not more prone to other cognitive biases, nor were they low in cognitive ability or other rational thinking dispositions” (paragraph 4). These findings were used to argue that the normative standard should incorporate some degree of projection, which is a rational strategy. Krueger (1998b) disagreed, arguing instead that a strong positive correlation between projection and cognitive ability would be necessary to draw this conclusion. Each of these treatments of projection fails to take note of the normative standard proposed analytically and confirmed experimentally: although some degree of projection is warranted, too little or too much constitutes a suboptimal judgment strategy³.

7. To test the relationship between measures of projection and other criteria, there are several ways that departures from an ideal degree of projection can be handled in statistical models. For example, one could compute scores representing absolute deviations from normatively appropriate projection, or enter a quadratic term for projection into a regression equation. It is difficult to know how to interpret a linear relationship (or the lack thereof) between a “raw” projection measure and any criterion, and such tests do not seem to rule out any meaningful substantive hypotheses.

FOOTNOTES

[1] Bayesian inference, which has been discussed throughout this debate on social bias, performs essentially the same function: the probabilities of two alternative hypotheses (one of which often corresponds to the null hypothesis, the other to a point prediction associated with a substantive hypothesis) are compared. This helps to rule out at most one competing hypothesis.

[2] Each of these experiments also revealed a “truly false consensus effect” over and above a normatively acceptable degree of projection.

[3] For a thorough discussion of whether and in what way such a “suboptimal” strategy may constitute irrationality, consult Stanovich (1999).

REFERENCES

- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.
- Chow, S. L. (1999). In defense of significance tests. *Psychology* 10(6)
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?10.006>

<ftp://ftp.princeton.edu/pub/harnad/Pycoloquy/1999.volume.10/psyc.99.10.006.social-bias.15.chow>

Dawes, R. M. (1989). Statistical criteria for a truly false consensus effect. *Journal of Experimental Social Psychology* 25: 1-17.

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

Krueger, J. (1998b). What can individual differences in reasoning tell us? *Pycoloquy* 9(77)
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.77>
<ftp://ftp.princeton.edu/pub/harnad/Pycoloquy/1998.volume.9/psyc.98.9.77.social-bias.12.krueger>

Krueger, J. (1999). Significance testing does not solve the problem of induction. *Pycoloquy* 10(15)
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?10.015>
<ftp://ftp.princeton.edu/pub/harnad/Pycoloquy/1999.volume.10/psyc.99.10.015.social-bias.16.krueger>

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology* 67: 596-610.

Krueger, J., & Zeigler, J. S. (1993). Social categorization and the truly false consensus effect. *Journal of Personality and Social Psychology* 65: 670-680.

Platt, J. R. (1964). Strong inference. *Science* 146: 347-353.

Ruscio, J. (1998). Applying what we have learned: Understanding and correcting biased judgment. *Pycoloquy* 9(69)
<http://www.cogsci.soton.ac.uk/psyc-bin/newpsy?9.69> *Pycoloquy* 9(46)
<ftp://ftp.princeton.edu/pub/harnad/Pycoloquy/1998.volume.9/psyc.98.9.69.social-bias.7.ruscio>

Stanovich, K. E. (1998). Individual differences in cognitive biases. *Pycoloquy* 9(75)
<http://www.cogsci.soton.ac.uk/cgi/psyc/newpsy?9.75>
<ftp://ftp.princeton.edu/pub/harnad/Pycoloquy/1998.volume.9/psyc.98.9.75.social-bias.11.stanovich>

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.