out respect to what others said. Ceraso et al. (1990) have reported that social psychologists of the time were shocked by "early returns" from Asch's studies showing that people "stick to their guns" (p. 8). Even Asch's (1956) final results provide more compelling evidence for truth-telling than for conformity. If Milgram's (1974) basic study had been described like Asch's has been, the focus would be entirely on the 35% of participants who refused to continue. Why do we emphasize the minority responses in Asch and the majority responses in Milgram? K&F's complaint of pervasive negativity is an excellent candidate.

A crucial reason for the story having unfolded the way it has is that Asch adopted a zero-tolerance norm (sect. 2.2.1). Asch (1952) thought he had created a simple moral dilemma between *truth* and *consensus,* in which truth was good and consensus was not. To agree with the unanimous majority on even a single trial was to have erred. But perhaps Asch's imputation of error was rash (sect. 3.1). Campbell (1990) argued that consensus is a good (he calls it *trust*), and that it should be integrated with other values (e.g., truth) to guide behavior in the Asch situation. He proposed that it would be most rational to believe that the majority is correct, but that the most moral action would be to offer one's dissenting view.

Hodges and Geyer (submitted) have suggested that Asch's analysis is simplistic and Campbell's is unrealistic. Interpreting the Asch situation in terms of values and conversational pragmatics, they ask: How does one speak the truth in a complex, tense, and frustrating situation? How does one answer the experimenter's questions in a way that simultaneously honors one's peers, the experimenter, one's own perception, and the situation in which all are embedded? Although any one trial prevents a resolution, over 12 critical trials, the actors in the Asch situation can try to balance as best they can their differing obligations. This means that they might occasionally give an incorrect answer, not because they are gullible, cowardly, or incompetent, but as an implicit signal that they have "heard" the majority and that they are open to further conversation despite the sharpness of the disagreement. By engaging in a local error, actors may be communicating a larger truth about the tension of their multiple obligations and their frustration in realizing all of them.

If this analysis is correct, then there is an irony in Asch's work, which, like the paradox of the fundamental attribution error (FAE; sect. 3.1.3.1), deserves to be "savored like a fine Merlot." It is this: Asch could see that his own deception – the design of the experiment – was part of a larger quest for truth, yet he would not accord his participants the same latitude.

Whatever the merits of the foregoing hypothesis, it illustrates K&F's call for more positive, balanced approaches (sect. 4) to social cognition and action, and for considering whether the norms by which behavior are judged are "incomplete, wrong, or misapplied" (sect. 3.1.2). Furthermore, it attempts to attend to the whole range of behavior (sect. 4). Most explanations of Asch's experiments are so fixated on explaining "conformity" that they overlook Asch's two main results: the preponderance of dissenting responses and the enormous range of responses. Hodges and Geyer (submitted) hypothesized that there might be three different strategies for integrating truth, consensus, and other values, suggesting that together these differing strategies would provide for group survival better than any one strategy alone. Their hypothesis illustrates K&F's suggestion that "multiple norms may need to be considered" (sect. 3.1.2). As Funder puts it elsewhere: Situations are complex, generating multiple motivations, such that "life is a continuous struggle to balance them all and find some kind of workable compromise" (Funder 2001b, p. 23).

Moving toward a more balanced social psychology that understands behavior as guided by multiple values will be difficult. Asch (1990) noted that a "central" theme of his research had been "that there is an inescapable moral dimension to human existence. . . . Yet psychologists have been among the most determined opponents of this claim" (p. 53). Thus, the open discussion of values K&F call for (sect. 5) will not come easily.

K&F briefly acknowledge the difficulty in their reference to the debates that have emerged about decision-making norms (sect. 3.1.2). Finding the balance they call for (sect. 4.3.2) will require negotiating some middle way between the enlightenment rationalism that tempted Asch (Leyens & Corneille 1999), and the subjective relativism that tempts them (i.e., participants' own goals define what is right; sect. 3.1.2). If values are simple and obvious, no discussion is needed; if they are merely what individual psychologists "consider desirable" (sect. 5), no discussion is possible. Discussions, as Asch realized, require real constraints and real obligations. In fact, his purpose in doing the experiments was to demonstrate that clear physical constraints and real moral obligations make rational behavior possible.

What obligations – which K&F refer to in Lewinian terms as "force fields" (sect. 4.3.2) – frame social relations and provide the basis for judging our actions and decisions (Hodges & Baron 1992; Sampson 2003)? Asch thought truth was our primary obligation. K&F emphasize accuracy. Truth and accuracy are crucial to human survival, but there is more that needs to be included if we are to flourish. For a start, there is compassion (sect. 5).

The goodness of judgment index

Lee Jussim

Department of Psychology, Rutgers University, Piscataway, NJ 08544.
jussim@rci.rutgers.edu **http://www.rci.rutgers.edu/~jussim/**

Abstract: Evidence is presented indicating that mainstream social psychology material leads undergraduates to conclude that people are irrational. To help address the problems identified by Krueger & Funder (K&F), a new statistic, the Goodness of Judgment Index (GJI), is presented. A concrete example based on a recent study is used to show how the GJI can be used to bring some balance back to research emphasizing error and bias.

Krueger & Funder (K&F) are right in almost every one of their particulars. People raise decent families, create wonderful works of art, invent computers and cell phones, hold warm and enjoyable gatherings among friends, figure out how to regularly travel at over 600 miles per hour, teach their children how to walk, talk, and ride bikes, create vast universities for educating young adults, and so forth. How could all this possibly occur if people did little more than engage in one immoral behavior after another, and, when not subordinating, torturing, or murdering one another, went about making the dumbest decisions imaginable?

I realize that no social psychologist has ever written anything quite as starkly damning as the prior sentence, but much of the body of work of social psychology leads to an impression consistent with that stark, dark, prior sentence. I recently taught an honors introductory social psychology class – these are among the most intelligent and thoughtful students in psychology. Their readings prior to the midterm included: Myers's (1996) introductory social psychology text's chapter on the self, Aronson's (1999b) *Social Animal* chapter on self-justification, Merton's (1948) classic article on self-fulfilling prophecies, La Piere's (1934) "attitudes do not predict behavior" study, and two chapters each from Cialdini's (1993) book on social influence, and Ross and Nisbett's (1991) book on the person and the situation. These are well-respected and mainstream social psychological writings.

One midterm essay question was, "According to social psychological research, are people mostly rational or mostly irrational?" Three quarters of my students concluded that social psychology demonstrated that people were mostly irrational. See the following examples.

First student, introductory sentence: "Through taking this class, I have come to the conclusion that people are, and have always been, primarily irrational."

Second student, introductory sentence: "People are not rational beings; rather they are rationalizing beings."

https://doi.org/10.1017/S0140525X04390084 Published online by Cambridge University Press

Third student, concluding sentence: "I guess that we are probably irrational and spend our lives trying to convince ourselves that we are rational."

This testifies to just how little insight social psychology has provided into how people do things well.

K&F have made many profound points, and I would like to augment one. Specifically, they highlighted how, in most studies, accuracy or rationality is the null hypothesis. It is, therefore, impossible to obtain evidence of accuracy or rationality, because it is impossible to "confirm" the null hypothesis. They are right about this, but there is another problem with null hypothesis testing in such research. The statistical comparison is usually deviation from perfection (whether that perfection means zero difference between experimental groups or deviation from predictions of some sort of normative model).

Unfortunately, however, perfection is so high a standard that researchers rarely apply it when testing their own models, theories, and hypotheses! If the results confirm the hypotheses on two out of three dependent variables in Study One; and then reappear on three out of four dependent variables in Study Two, the paper will most likely get published and the tone will emphasize confirmation of the author's theory.

This situation is explicit in use of goodness-of-fit indices when testing structural equation models. One of the earliest tests of model quality was the chi-square test, which evaluated whether the correlations or covariances predicted by the hypothesized model significantly deviated from the observed correlations. With large samples, the chi-square is almost always significant, however, indicating that almost all hypothesized models are imperfect. It is not often very useful, however, to know that one's proposed model is not perfect. Typically, we want the model to be good, not perfect.

This inspired some researchers (e.g., Bentler & Bonnett 1980) to develop measures of fit that indicate how much the model does explain. One common fit measure is the Normed Fit Index (NFI). For the NFI, one assesses the improvement in fit obtained in one's hypothesized model, as compared to that of the null model – one which assumes no variable is related to any other variable. The NFI is computed by:

$$\frac{(\text{chi-square[null model]}) - (\text{chi-square[hypothesized model]})}{(\text{chi-square[null model]})}$$

If the hypothesized model has a chi-square of 80, and the null model has a chi-square of 1,000, the NFI equals .92. The model may be significantly worse than perfect, but this result can be metaphorically interpreted as meaning that it accounts for about 92% of the covariances among the observed variables.

In the same spirit, I propose a Goodness of Judgment Index (GJI) for studies of error and bias. Most results in studies of judgment, decision-making, and prediction can be readily translated into a 0–1 scale, because such studies use frequencies, percentages, or correlations as their basis for identifying error and bias. The GJI then becomes:

$$\frac{(\text{Maximum possible imperfection} - \text{actual degree of imperfection})}{\text{Maximum possible imperfection}}$$

"Maximum possible imperfection" is the most anyone could possibly be wrong under the circumstances. Imperfection can be operationally defined as errors, discrepancies from predicted values, discrepancies from perfect correlations, and so forth.

The GJI is simple to use and indicates the proportion of improvement of social judgment compared to complete error. Scores above .5 mean that the judgment is closer to complete rationality or accuracy than to complete error or irrationality; scores below .5 mean that the judgment is closer to complete error or irrationality than to complete rationality or accuracy.

Consider the following data from a recent study on bias in social perception (Monin & Norton 2003). During a community water crisis, Princeton students were asked to stop showering for about two days. The researchers surveyed Princeton students, and asked them: (1) How many showers they took during the ban; and (2) what percentage of other Princeton students showered during the ban. I only focus on Day 3 of their results, although the basic pattern holds throughout. During Day 3 of the ban, 47% of their respondents admitted to taking showers. Bathers predicted that 66% of Princeton students took showers; nonbathers predicted that 47% of Princeton students took showers. Interpretation of the results focused almost entirely on error and bias.

Their results, however, provided more evidence of accuracy than of bias. One does not need the GJI for the nonbathers – their prediction of the proportion of bathers was dead-on accurate (47%), and their GJI would be 1.0. Even the predictions of the bathers were more rational than irrational. Because 47% of the students bathed, the greatest possible imperfection would be 53% (which would occur if 100% of the students were predicted to bathe). The observed imperfection among bathers is 66% – 47% = 19%. GJI = (.53 – .19)/.53 = .64. So, the bathers' judgments, which clearly showed a false consensus effect, were closer to complete accuracy than complete inaccuracy.

I hope that K&F's article inspires researchers studying bias to develop tools to estimate, not just the degree to which people make biased and irrational judgments, but also the degree to which they make unbiased and rational judgments. When data are consistently presented in a balanced manner, we will have taken a major step toward a balanced social psychology.

Building an even better conceptual foundation

Tatsuya Kameda[a] and Reid Hastie[b]

[a]*Department of Behavioral Science, Hokkaido University, Sapporo, Japan 060-0810;* [b]*Center for Decision Research, Graduate School of Business, The University of Chicago, Chicago, IL 60637.* **tkameda@let.hokudai.ac.jp**
reid.hastie@gsb.uchicago.edu
http://lynx.let.hokudai.ac.jp/members/kameda

Abstract: Krueger & Funder (K&F) spend too much time on their critique of some classic studies in social psychology. They should have spent more time developing their constructive ideas about better methodologies and, especially, better conceptual foundations for the field. We endorse their exhortation to consider social behavior in its ecologically adaptive context, and we present a few ideas of our own about how to develop a more comprehensive conceptual framework.

Krueger & Funder (K&F) are unhappy with two traditions of research in social psychology. They provide a biting and occasionally witty critique of classic demonstrations of misbehavior (over-conformity, over-obedience, and failures to help) and the heuristics and biases approach to social cognition. Many of their insights into weaknesses of these approaches have merit, although their relentlessly negative perspective vastly undervalues the enormous positive contributions of research from these traditions. Any school of behavioral research can be subjected to scathing criticism; indeed, the authors' own *Realistic Accuracy Model* is also limited in many ways.

Acknowledging that there is something to learn from a critical evaluation of these major pillars of the social psychological canon, we are more interested in K&F's affirmative suggestions for improvements in methods and theory. Most of their suggestions are methodological, and we have already expressed our enthusiasm for the correspondence framework for the analysis of degrees of accuracy (see Hastie & Rasinski 1987, for an introduction to the correspondence–coherence distinction, a discussion of "methodological logics" for judgment research, and a comment on the manner in which the null hypothesis test has obstructed the study of social judgment). And, like K&F, we find the Bayesian framework to be conceptually superior to the Neumann-Pearson null

https://doi.org/10.1017/S0140525X04390084 Published online by Cambridge University Press