

Jeryl L. Mumpower
University at Albany
State University of New York

I would like to draw attention to two sets of ideas regarding extreme event decision making, for possible consideration at the workshop.

1. Tom Stewart, Radhika Nath, and I have used ideas proposed by Hammond (1996), and adapted from Taylor and Russell (1939) to address events that, while more frequent than typical exemplars of extreme events, share many of the properties typically attributed to them -- rareness, high consequence, high uncertainty, disruptiveness, complexity, affecting large numbers of persons, and inequitable. Definitions of extreme events are probably inherently ambiguous and context-dependent, so examples are helpful. We have been concerned about "near-extreme" events such as admission to a psychiatric emergency room, diagnoses of breast cancer, and admission to undergraduate programs in Ivy League schools.

Our paper entitled "Affirmative Action, Duality of Error, and the Consequences of Mispredicting the Academic Performance of African-American College Applicants" has just been accepted by the *Journal of Policy Analysis and Management*. The paper deals with college admissions decisions, as an example of low-probability events in which treatment decisions must be based on diagnostic judgments. Outcomes in such instances can be divided into four exhaustive, mutually exclusive categories: false positives (admit unqualified applicants), false negatives (reject qualified one), true positives (admit qualified ones) and true negatives (reject unqualified ones). The implications of different potential action policies depend on three factors, which need to be considered simultaneously. These are *selection rate* (e.g., percentage admitted); *base rate* (e.g., percentage of those applying who could do the work if admitted); and

predictive accuracy, (i.e., degree of correspondence between predictions of performance and actual performance). Ultimately, the desirability of alternative action policies depends on the *costs* and *benefits* associated with the four types of possible errors and correct classifications.

Our analysis of affirmative action policy leads to three general conclusions that may be generalizable to many low-probability events in which decisions to take action depend on diagnostic judgments. Part of the conclusion of our paper reads as follows:

First, without taking steps to improve predictive capabilities, changes in policies will primarily involve simply trading off one type of error for the other. Uncertainty creates a tradeoff between false positives and false negatives. Reducing one type of error leads inevitably to an increase in the other. The only way to reduce one kind of error without creating more of the other kind is to reduce the uncertainty. Methods for doing this include obtaining better and more reliable information for decision making and making better use of the available information. ... until we develop better tools for predicting the performance of minority group applicants these trade-offs among the two types of errors will plague us.

Second, the appropriate tradeoffs between false positives and false negatives lie at the crux of policy debates in many different substantive areas. Since the uncertainty that necessitates trading one kind of error for another can never be completely eliminated, choosing a course of action means assigning, implicitly or explicitly, relative values to the two kinds of error...

Third...any course of action creates both "winners" and "losers." As much as we would like to eliminate all fairness and injustice, it is unavoidable. This injustice results not from the selfishness or cruelty or incompetence of policy makers or their policies, but simply from uncertainty...

2. Errors of judgment about matters of fact (e.g., predictions of extreme events) may be divided into two classes: systematic errors (bias) and random ones (inconsistency). It may be odd to think of inconsistency in the context of judgments made about singular events or extremely rare ones, so perhaps the concept of “cognitive control” (the original term in the Social Judgment Theory tradition) is more appropriate. Research within the SJT paradigm on multiple-cue and interpersonal learning by Hammond, Brehmer, Stewart and others has repeatedly shown that performance often takes a long time to reach maximal levels because judges find it difficult to make judgments in a perfectly consistent way even after they “have gotten the idea.” Cognitive control or consistency requires practice. In other words, judges acquire knowledge first, but then require practice before they learn to apply it in a sufficiently consistent manner to reach asymptotically maximum performance. This implies that performance in situations that require judges to integrate multiple pieces of information in order to make diagnostic or predictive judgments about rare events is likely to be sub-maximal because of lack of cognitive control in the exercise of knowledge that the judge does possess. This suggests the hypothesis that judgmental performance in situations involving rare events will be improved by finding means to improve cognitive control (i.e., eliminate noise or random error) through practice in simulated conditions or other means.