

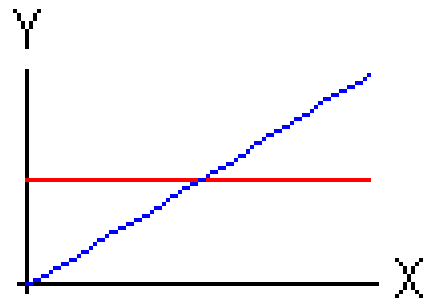
Representative *and* Efficient Designs

Gary McClelland
University of Colorado, Boulder

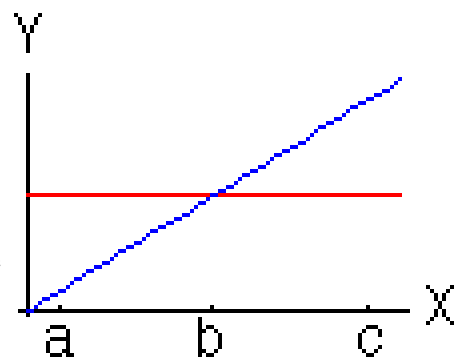
September 1999

When capturing policies or studying judgment strategies, many researchers follow Brunswik's recommendation to use representative designs in which the distribution of stimuli or cases presented to the judge matches the distribution of such cases observed in the environment. Most readers of this website will be familiar with the advantages of representative design and the disadvantages of artificial experimental designs for studying human judgment. However, one major disadvantage of representative designs is that they tend to be inefficient statistically. That is, representative designs often require a large number of cases in order to estimate model coefficients, identify function forms, and assess agreement with environmental models. The inefficiency of representative designs increasingly limits judgment studies with professionals and others who lack the time and patience required to judge many cases. This essay considers how another design principle--efficiency--might be of equal importance to representativeness and how the two design principles might work in cooperation to produce more useful designs for judgment analysis. In this informal web essay, I use graphs rather than equations to make mathematical and statistical arguments.

To highlight the efficiency issues and to illustrate how they can differ from representativeness considerations, let's examine the simple analysis problem in the figure to the right which depicts the possible functional relationships between cue X and judgment Y . We can view the statistical problem in judgment analysis as determining whether the functional relationship between X and Y is more like the sloped blue line or the flat red line. Or if the blue line represents the true relationship in the environment, then the statistical problem is determining the degree to which the judge's slope matches the true slope.

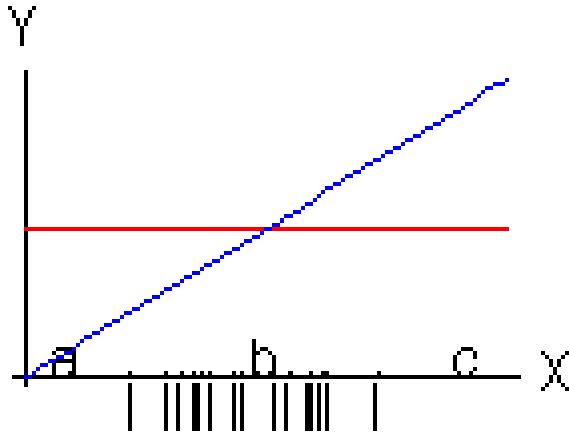


Efficiency Principle. The points **a**, **b**, and **c** in the graph to the right represent three possible cue values we might use for X . These three cue values for X are not equally useful for assessing the slope of the line relating X to Y . At the cue value $X = b$, both the flat red line and the sloped blue line predict the same judgment for Y ; hence, **b** is useless for distinguishing between the possible linear relationships. On the other hand, cue values $X = a$ and $X = c$ are very useful because the red and blue lines make very different predictions for the corresponding judgments of Y . In general, cue values for which the distance between the lines is greatest are most useful. That is, extreme cue values have the greatest statistical efficiency for identifying linear function forms and middle values

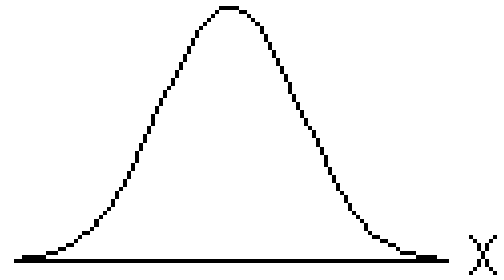


have the least, perhaps even zero, efficiency.

Representative Principle. The actual distribution of cue values for X in the environment is likely to have a normal-like

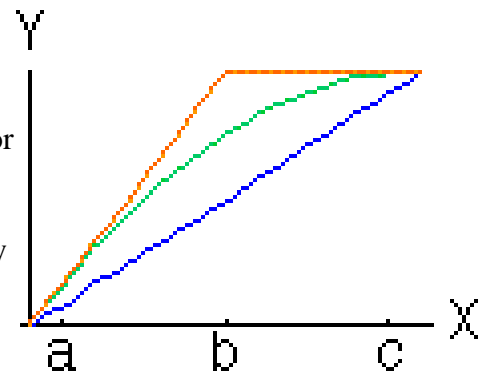


distribution such as the graph to the right. If we sample cue values randomly then the judge will see cases which match the statistical properties (mean and variance) of cases typically encountered in real judgment situations. The graph to the left shows fifteen cue values randomly sampled from a normal distribution. Note that none are as extreme as the highly efficient cue values **a** and **c**. Instead, most of the sampled cue values are between those most efficient cue values and the highly inefficient cue value **b**. Thus, the statistical



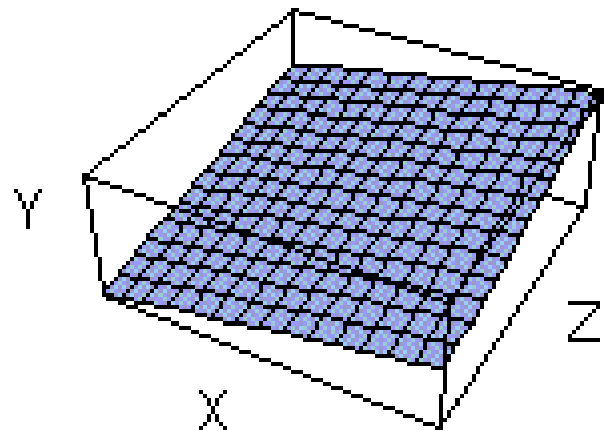
efficiency of most of the cue values in a representative design is intermediate. Increasing the number of cue values judged is the only way to overcome the inefficiency of the design.

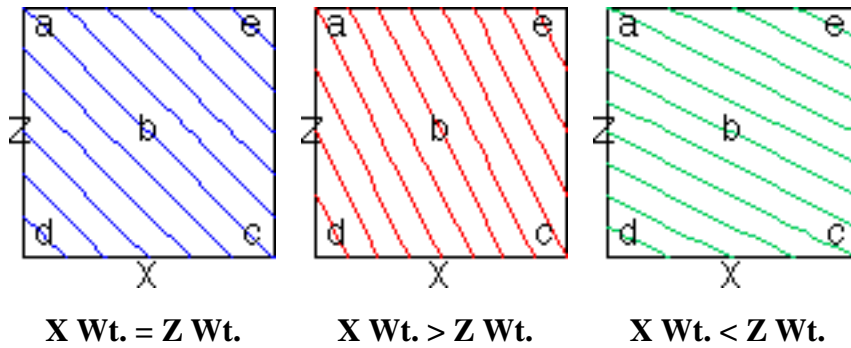
Nonlinear Function Forms. The graph to the right illustrates monotonic, nonlinear function forms that often arise in judgment research: quadratic (green), linear to a maximum (orange), and baseline linear (blue). Although cue values **a** and **c** are very useful for determining the increasing monotonic relationship, they are useless for distinguishing between the alternative function forms. Now it is cue value **X = b** that is most useful because the curves are maximally separated at that point. Again, the randomly sampled values in the representative design would have moderate inefficiency *if the end points of the functions had already been determined efficiently.*



McClelland & Judd (1993) show that for detecting nonlinearity, random samples from normal distributions usually have extremely low efficiencies.

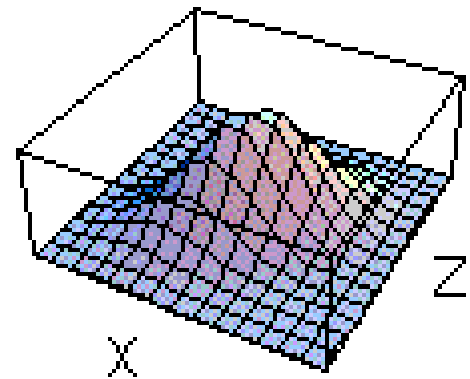
Tradeoff Weights. Many judgment analyses attempt to determine the tradeoff (i.e., partial regression coefficients) between two or more cues. For the case of two cues, this is equivalent to estimating the regression plane, such as the one to the right which depicts equal weighting of cues **X** and **Z**. Instead of considering such planes, it is useful to view the joint relationship as a contour plot or topographic map. In the views below of the **X-Z** plane, the lines depict equi-judgment contours. That is, judgments of cases along a line are equal, and the overall level of the judgments increase from the southwest to the northeast.



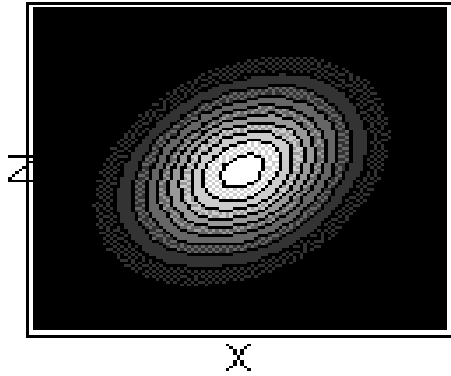


For equal weights (blue), the cues substitute equally for one another. For greater weight on cue **X** (red), smaller changes in **X** than in **Z** are required to move to the next equi-judgment contour. Points **a** through **e** represent possible cue combinations that we might present to judges as cases.

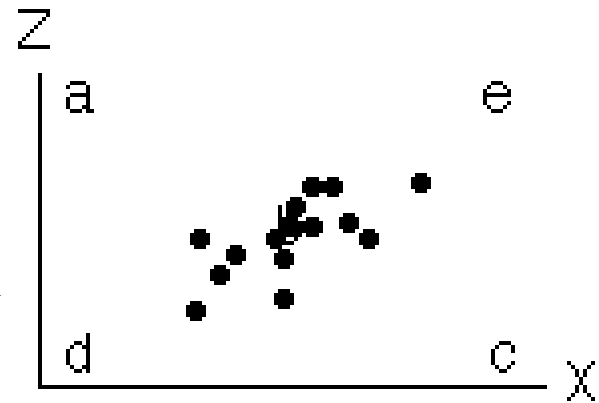
Efficient Design for Tradeoff Weights. Cue combinations along the positive diagonal (such as **d**, **b**, and **e**) provide little or no information for differentiating among the weighting alternatives. In this case, all three relative weightings agree that **b** is five judgment contour lines beyond **d** and **e** is another five judgment contour lines beyond **b**. Hence, judgments from those points provide zero efficiency. In contrast, the alternative relative weightings make very different predictions about the judgments for cue combinations **a** and **c**. For equal weights (blue), **a** and **c** fall on the same judgment contour line so they should be judged the same. For greater weight on **X** (red), **c** is several contour lines beyond **a**; while for greater weight on **Z** (green), **a** is several contour lines beyond **c**. Hence, the judged values of extreme cases on the negative diagonal are very efficient for estimating the relative cue weights.



Representative Design for Tradeoff Weights. For two or more cues, representative design requires that cue combinations represent not only the environmental means and variances but also represent the environmental cue intercorrelations. The graph to the right depicts the bivariate normal distribution for a correlation of about $+0.3$.



The topographic representation (to the left) of this bivariate normal distribution shows the equi-likelihood contours. Sampling from this distribution would yield many cue combinations along the positive diagonal (where they are inefficient) and virtually no cue combinations near the extremes of the negative diagonal (where they would be the most efficient).



The graph to the right illustrates a random sample of fifteen cue combinations from the bivariate normal distribution with $r = +0.3$. As expected, most of the cases fall on the positive diagonal and are therefore very inefficient for determining the relative weights. None of the cue combinations are close to **a** or **c**, which would provide optimal efficiency.

Real-time Efficient Design. With a computer collecting the judgments, it would be feasible to use the previous judgments to construct in real-time the most efficient cue combination to present next. Algorithms for doing so have not been developed, but they should not be difficult to derive.

Efficient Plausible Design. It might appear that efficient designs dominate representative designs. However, efficient designs have a serious disadvantage--they encourage the use of unreasonably extreme cases. For example, to determine relative weights for combined GRE (verbal + quantitative) and GPA when judging graduate student applications, the two most efficient cue combinations would be (GRE = 1600, GPA = 0) and (GRE = 400, GPA = 4), but these are implausible, if not absurd, combinations. Hence, it is necessary to constrain cases in efficient designs to plausible limits established by a representative design.

Augmented Representative Design. An alternative method for balancing efficiency and representativeness goals is to augment a representative design with plausible extreme cases that would otherwise be unlikely to appear in a small random sample. With the added efficiency from the extreme cases, the representative design could be considerably smaller than typically used in judgment studies.

Importance of Extreme Cases. Extreme cases are important not only for statistical reasons but also for psychological and practical reasons. The admissions committee will sail smoothly through their judgments of the student with GRE = 1400, GPA = 3.7 or a student with GRE = 990, GPA = 2.5, but a student with GRE = 1450 and GPA = 2.85 is likely to bring out the differences in each committee member's judgment policy. Pure representative designs in which such cases are rare underestimate latent conflict due to differences in judgment policies. In professional contexts, it is the extreme cases that are

most likely to distinguish the experts. One-hundred year floods are necessarily rare, but we want meteorologists who are able to predict them. A representative design would be unlikely to test a meteorologist on such an extreme event.

"As-If" Representative Design. Even if a non-representative design is used (such as the efficient plausible design or the augmented representative designs described above), it is still possible to estimate the correlations, beta weights, R-sq's, len's model achievement indices, etc. *as if* a representative design had been used. Several strategies exist for doing so: (a) deriving the model from an efficient plausible design and then applying the model to a set of representative cases, (b) using weighted regression where each observations weight is a function of its representativeness, (c) or mathematically adjusting the coefficients based on the variances and covariances of a representative design. As an example of strategy (c), McClelland & Judd (1993) show how to use R-sq estimated from one design to impute the R-sq that would have resulted from another design.

Conclusions and Summary. Representative designs for complex judgment situations require more judgments than busy professionals will tolerate or than most people can do without losing concentration. Efficient designs, relying on extreme cases, use many fewer cases to estimate model parameters with comparable standard errors to those from representative designs. However, unreasonably extreme cases in efficient designs yield arbitrary judgments and cause expert judges to dismiss the judgment task as irrelevant. Combined design strategies such as constraining efficient designs to plausible cases or augmenting smaller representative designs with a few more efficient extreme cases offer significant improvements to judgment studies.

Recommended Reading

Kuhfeld, W.F., Tobias, R.D., & Garratt, M. (1994). Efficient experimental design with marketing applications. *Journal of Marketing Research*, 31, 545-557.

McClelland, G.H. (1995). Asteriod Configus: Why the World Will Always Look Flat to a Brunswikian. Talk presented at the annual meeting of the Brunswik Society, Los Angeles, CA, November 1995. [transperancies available from <http://samiam.colorado.edu/~mcclella/brunswik/flat.pdf> (787 Kb)]

McClelland, G.H. (1997). Optimal design in psychological research. *Psychological Methods*, 2, 3-19.

McClelland, G.H. and Judd, C.M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114, 376-390.

Mead, R. (1988). *The design of experiments: Statistical principles for practical application*. Cambridge, England: Cambridge University Press. [very readable treatment of optimal design principles by a biologist.]