Area Validation Exercise

The Challenge of Quantifying Judgment

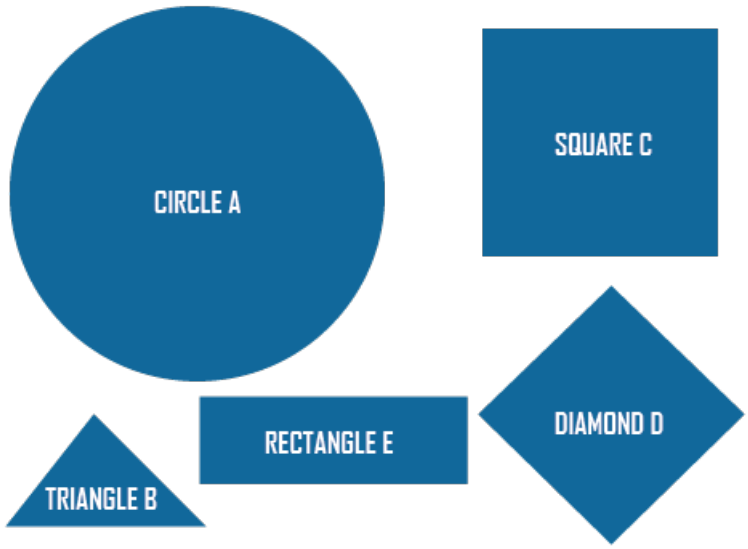
We make hundreds of estimates every day, often without realizing it. We guess how fast a car is approaching before pulling out into traffic, weigh the invisible toll a risky decision might take on our reputation, or try to balance competing priorities when choosing a career path. But there is a massive difference between feeling an answer and expressing it numerically in order to take action. Welcome to the challenge of quantifying judgment. Whether we are dealing with objective physical realities or deeply subjective human values, translating human intuition into data is notoriously difficult. How do we measure the intangible? How do we standardize a gut feeling? This page explores why this gap exists, how it impacts our decision-making, and the frameworks we can use to bridge the divide between human perception and hard metrics.

Welcome to the challenge of quantifying judgment. Whether we are dealing with objective physical realities or deeply subjective human values, translating human intuition into data is notoriously difficult. How do we measure the intangible? How do we standardize a gut feeling?

This page explores why this gap exists, how it impacts our decision-making, and the frameworks we can use to bridge the divide between human perception and hard metrics.

Area Validation Experiment

The relative areas of different geometric shapes (each within an order of magnitude) were derived from pairwise verbal judgments for shapes such as the following:



This experiment has been performed thousands of times, and you can easily do this individually or with a group. The actual relative areas and estimates derived from pairwise verbal judgments with a group of executives at the Ford Motor Company are shown below:

Shape	Rank	Reverse	Proportion	Pairwise Verbal	Actual
Circle	**1**	**5**	33.3%	49.6%	**47.5%**
Triangle	**5**	**1**	6.7%	4.8%	**4.9%**
Square	**2**	**4**	26.7%	23.6%	**23.2%**

© 2019 Expert Choice. All Rights Reserved

Shape	Rank	Reverse	Proportion	Pairwise Verbal	Actual
Diamond	**3**	**3**	13.3%	14.5%	**15.1%**
Rectangle	**4**	**2**	20.0%	7.5%	**9.3%**

Note how close the estimates are in the pairwise verbal column (column 5) to the actual column (column 6). Note also how deficient the estimates are if one were to simply derive the estimates based on the ranking (ordinal measures) of the shapes (column 4 above)

The accuracy of the derived ratio scale priorities is truly amazing considering that the inputs were ordinal measures (words on the fundamental verbal scale). Deriving ratio scale measures from ordinal inputs is somewhat magical since ratio measures have all of the information of ordinal measures, plus interval and ratio meaning as well. In a sense, it gives new meaning to GIGO -- garbage in, genius out! Do not, however, take this for granted. There may be cases where intervals or ratios of the priorities resulting from verbal judgments do not adequately represent the decision maker's feelings. It is incumbent upon the decision maker(s) to examine the resulting priorities and, if they do not adequately represent the decision maker's feelings, to revise the judgments in either the graphical or numerical mode

If graphical or numerical judgments had been made, we would expect the results to be even better, since this is an "objective" problem with known scores. So why use verbal judgments at all?

Verbal judgments are often more appropriate when judging qualitative factors, and all crucial decisions have qualitative factors that must be evaluated. Numerical comparisons can convey the wrong sense of accuracy of the input. If one were to say that clean air is 4.3 times more important than clean water, rather than between moderate and strongly more important, how does one defend the 4.3 rather than 4.2, 4.6, or 5?

See How it Works! | Area Validation Blog

© 2019 Expert Choice. All Rights Reserved