Magnitude Scaling & Perceived Hostility Towards the United States

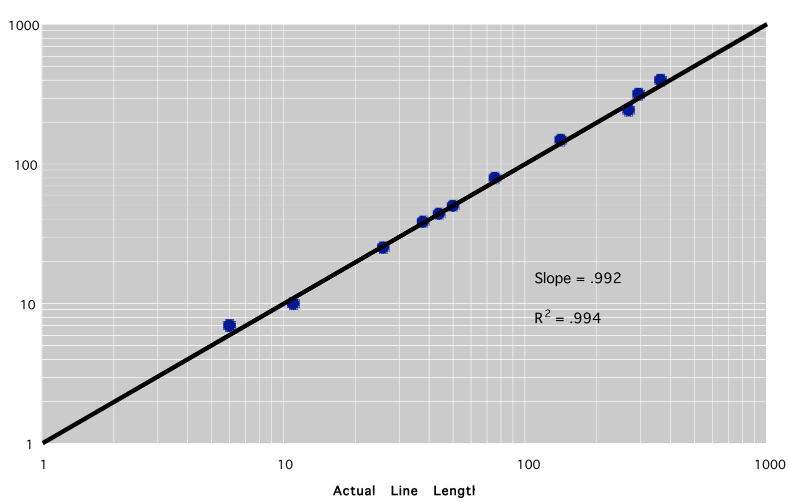
Many scales in the social sciences are categorical – respondents rate stimuli by selecting one of a fixed number of responses. These categorical scales, while simple to administer and analyze, have limitations. First, they provide only an ordinal level of measurement. This could limit the types of statistical analyses one could conduct on the data. Second, the labels a researcher uses to describe the categories have been shown influence subjects' responses (Lodge, 1981). Finally, by having a limited number of categories, we lose a good deal of information. Subjects are forced to place a wide range of stimuli into a limited number of categories.

If the limitations of categorical scaling are due to the limited resolution of the categories, perhaps we should allow respondents to select from an infinite number of categories. This would allow us to measure the actual magnitude of our respondents' perceptions and provide us with an interval level of measurement.

Consider the line length exercise found in Appendix A. 83 subjects were told the length of a *reference line* was 50-units. The subjects were then presented ten other lines of varying lengths, one at a time. Subjects were asked to provide a numeric estimate of the lengths of those ten lines relative to the length of the reference line. For example, if a subject perceived a line to be ten times as long as the reference line, the subject would provide a numeric estimate of 500 (10 x 50 = 500). Subjects were allowed to use any positive number for their estimates, including decimals or fractions.

Figure 1 shows the outcome of this exercise. After recording the perceived line length estimates from all 83 subjects, the geometric mean of those estimates was calculated for each line (take the log of each estimate, find the arithmetic mean of those logs, and then raise 10 to the power of the mean of the logs). These geometric means were then plotted against the actual lengths of each line on a logarithmically ruled graph. On such a graph (called a log-log graph), equal distances on each axis mark off equal ratios. If the magnitude of people's judgments (perceived line length) is proportional to the magnitude of the stimuli (actual line length), the resulting graph would be a straight line. The correlation between perceived and actual line lengths was found to be r = .994, showing a strong linear relationship. A linear regression analysis finds the slope of the line in Figure 1 to be 0.992. These results indicate that the perception of line lengths for the subjects in this study is proportional to the actual lengths of the lines.

Figure #1: Perceived Line Lengths vs. Actual Line Leng

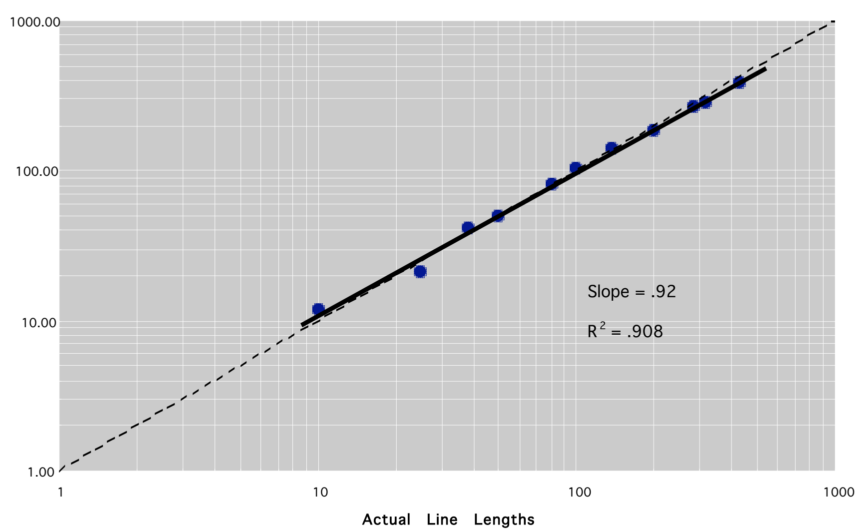


Researchers have found the principle underlying this linear relationship – *equal stimulus ratios produce equal subjective ratios* – holds for many aspects of our five senses. This *Power Law* is represented as $R = kS^b$, where R = the subjective perception of the magnitude of a stimulus, S = the actual magnitude of the stimulus, k = a constant, and b = the empirically derived exponent that characterizes the relationship. Transforming to logs, the Power Law can be represented as a simple linear function: $\log(R) = b\log(S) + \log(k)$. The value of b can be calculated through a simple linear regression.

The Power Law defines the relationship between stimulus magnitude and perception to be linear (when plotted on a log-log graph), but it does not define the slope of that line. Appendix B displays the slopes (characteristic exponents) obtained through controlled experiments for a variety of stimulus modalities. The exponents show the proportional increase in perceived stimulus magnitude when the actual stimulus magnitude has been doubled. For example, the exponent for line length is 1.0. This means that doubling the length of a line will produce a doubling of the numeric estimates (subjects will perceive a doubling of length). The exponent for loudness of 0.67 means that when the volume of a sound is doubled, subjects perceive a two-thirds increase in loudness.

The results displayed in Figure 1 provide evidence of the internal validity of the magnitude estimation process – subjects' numeric estimation of the proportional increases in line lengths <u>is</u> a power function and <u>does</u> have an exponent of 1.0. Reversing the exercise by having subjects produce lines of varying lengths relative to a reference line (see Appendix C) yielded similar results (Figure 2). The correlation of .953 provides evidence that the Power Law holds for this exercise. The observed slope (characteristic exponent) was calculated as 0.92, slightly smaller than the predicted exponent of 1.0 for perceived visual length. This slope appears to be lower than expected because of the regressive nature of the line production exercise. Subjects tended to draw the longer lines shorter than their actual lengths and the shorter lines longer than their actual lengths.

Figure #2: Line Production Exercise



Both line length exercises provide evidence that the Power Law holds in the psychophysical domain of perceived visual line length. This was easy to determine empirically, since the horizontal axis in psychophysical exercises are metric (they have a known scale). But how can we verify the Power Law (which is the basis for magnitude scaling) holds for social data (attitudes, beliefs, perceptions)? Since social data do not typically have an agreed-upon metric, what do we put on the horizontal axis? Stevens (1957) developed a solution to this problem, called *Cross-Modality Matching*. The logic of Cross-Modality Matching is this: If the exponents for each sensation listed in Appendix B does in truth represent its characteristic exponent, then any one of those sensory modalities could be used as a magnitude response measure. For example, suppose we asked subjects to estimate the relative lengths of lines using two response measures. In the first measure, we have subjects squeeze a hand-grip to estimate the length of the lines. If a subject believed a line was twice as long as the reference line, the subject would squeeze the hand-grip twice as hard. The characteristic exponent for hand-grip force has been determined to be 1.70. We then show the subjects a reference line and play a tone (the reference tone). The subjects then adjust the volume of the tone to match their perceptions of the magnitude of lengths for each line. For example, if a subject believes a line is twice as long as the reference line, the subject would adjust the volume of the reference tone so that it is twice as loud as the reference tone. According to Appendix B, the characteristic exponent for loudness is 0.67. Then, according to the Power Law:

$R_1 = kS_1^{b_1} = kS_1^{1.70}$ for the numeric estimates and $R_2 = kS_2^{b_2} = kS_2^{0.67}$ for the loudness estimates.

By equating the line lengths across both the numeric and loudness estimates, we can set these equal to each other, take their logarithms, and cancel the constants to find:

$$(1.70)\log(S_1) = (0.67)\log(S_2) \quad \text{so that} \quad \log(S_1) = \frac{0.67}{1.70}\log(S_2) = .394\log(S_2)$$

This final result can be rewritten in general terms as $\log(S_1) = \frac{b_2}{b_1}\log(S_2)$

This final result, a simple linear relationship between two response measures for the same set of stimuli, is the key to Cross-Modality Matching. If the Power Law is valid and if the exponents derived from magnitude estimation are truly characteristic, then any two response measures with established exponents could be used as stimulus magnitude estimates. If we then plot these two response measures against each other on a log-log graph, the slope of the resulting line should be equal to the ratio of the two characteristic exponents of the response measures. In this simple example, if we were to plot the geometric means of the loudness estimates on the horizontal axis and the geometric means of the hand-grip estimates on the vertical axis, the resulting plot should be linear with a slope equal to the ratio of the two characteristic exponents (0.394 in this case).

The nonmetric nature of social data is thus solved through Cross-Modality Matching. If subjects use numeric estimation (shown to have an exponent of 1.0) and line production (also shown to have an exponent of 1.0) as their responses to the same set of stimuli, we can validate the resulting magnitude scale by

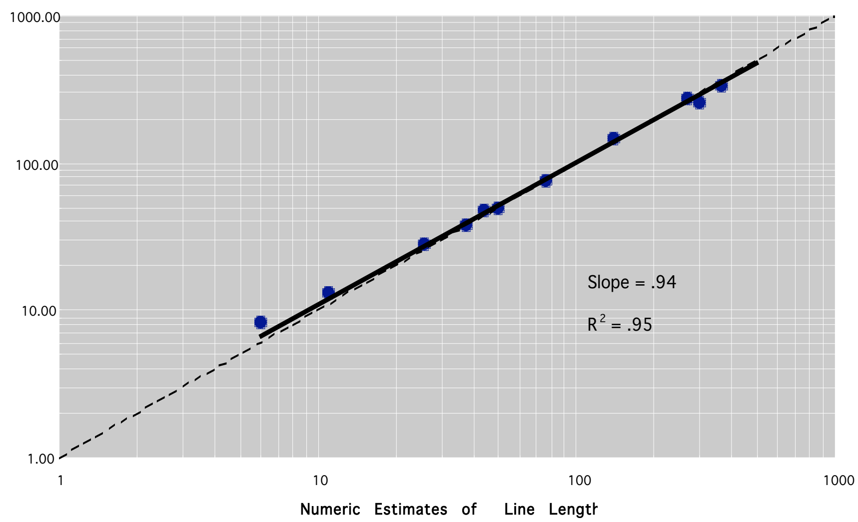
(1) determining if the slope of each response method is 1.0 (as in Figures 1 and 2), and
(2) determining if the slope of the response methods plotted against each other is equal to the ratio of their characteristic exponents (1.0 in this case).

Once these two criteria have been met, we can use the geometric means of subjects' responses to the stimuli as the (ratio-level of measurement) scale values for the stimuli.

Application: Perceived Hostility Towards the United States

Replicating a 1997 study conducted by Sulfaro & Crislip, magnitude scaling was used to determine the perceived level of hostility twenty countries have towards the United States. The data for this analysis come from surveys administered to 83 undergraduate students who were offered extra-credit for their participation. First, as a calibration/training exercise, the subjects completed the line length exercises discussed previously (Appendices A & C; Figures 1 & 2). The results of these exercises demonstrate the subjects understood the estimation procedures and that the Power Law held in both cases. Figure 3 shows the resulting Cross-Modality Matching plot, demonstrating, once again, differences in perceived magnitude in line lengths were estimated accurately.

Figure #3: Cross-Modality Matching

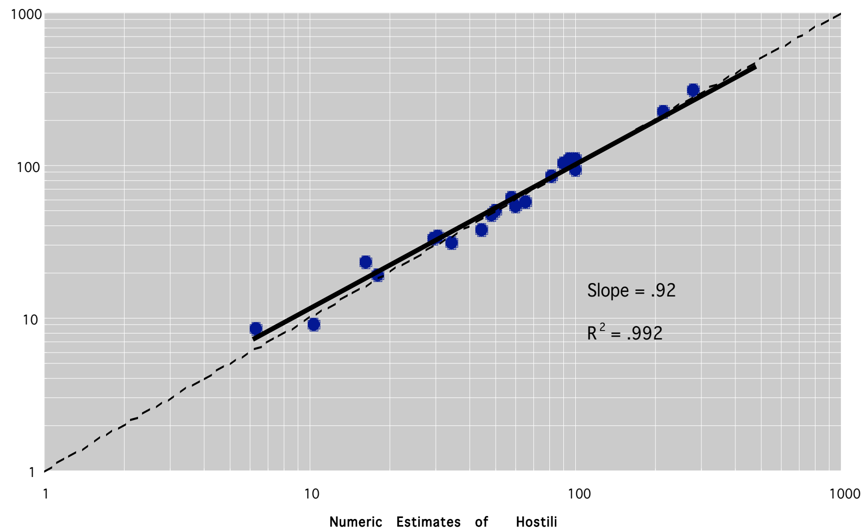


The subjects were then presented with a list of 20 countries from around the world and asked to rate their perception of the level of hostility each country has towards the U.S. First, subjects were asked to complete a categorical scaling of the stimuli (Appendix F). Subjects were then administered two magnitude scaling exercises. France was used as a reference and given a hostility rating of 50. Subjects were first asked to provide a numeric estimate of the perceived hostility for each country. Afterwards, the subjects were asked to produce lines whose lengths were proportional to the each country's perceived hostility. The survey instruments are found in Appendices D & E.

The geometric means were calculated for each country for both the numeric estimation and line production exercises. These means are listed in Appendix G. As you can see, ten countries were perceived to have more hostility towards the U.S. than France. Predictably, Iraq and Afghanistan were perceived to have the highest level of hostility (Iraq has over 5 time the amount of hostility towards the U.S. as France has). Australia and Sweden were perceived to have the lowest amount of hostility towards the U.S. The numeric estimates and line production estimates were in close agreement, with a correlation of .996. Interestingly, the countries with the highest and lowest estimated hostility levels tended to have the lowest amount of variability in the estimates provided by the subjects. Those interested in a discussion of subjects' perceived hostility towards the U.S. should read the article written by Sulfaro & Crislip (1997).

Figure 4 shows the log-log plot demonstrating the Cross-Modality Matching for perceived hostility. The slope of the line was found to be 0.92, which is close to the expected value of 1.0 (the ratio of characteristic exponents). This provides evidence of the validity of the resulting scale values.

Figure #4: Perceived Hostility Towards the U.



The scale values obtained from the numeric estimation and line production exercises were then compared to the ordinal scale values obtained from the categorical scales. Appendix H displays the standardized scale values obtained for each stimulus. Figure 5 displays a plot of the magnitude scale values against the categorical scale values. The curvilinear relationship between the two scales represents the limitations placed on scale values by the limited number of categories available for subjects to place stimuli.

Discussion

Magnitude scaling, as presented in this project, provides a number of advantages over categorical scaling. First, the scale values obtained from magnitude scaling represent a true ratio level of measurement (with an arbitrary zero point). Second, it does not limit (or influence) respondents by providing a nearly unlimited number of categories for placing stimuli.

Magnitude scaling also has some drawbacks. In the hostility survey, subjects took approximately 2-3 minutes to complete the categorical scaling. The same subjects took over ten minutes to complete the line production and numeric estimate exercises. Also, measuring the lines produced by each subject took a considerable amount of effort (which could be reduced by using a computer scanner). One could reduce the time needed to construct a magnitude scale by having subjects only complete the numeric estimations (and skipping the line production exercise). This would, however, not make it possible to conduct a Cross-Modality Matching analysis on the data. While magnitude scaling does take more time to complete and is a bit more taxing for subjects, its benefits seem to outweigh its drawbacks.

The book, *Magnitude Scaling: Quantitative Measurement of Opinions* written by Milton Lodge (1981) provides other uses of magnitude scaling in the social sciences. With a few modifications, the magnitude scaling methods can be used to scale subjects (instead of stimuli) on a unidimensional continuum. Lodge also outlines procedures one can use when subjects' numeric estimates are not in agreement with their line production estimates. Finally, Lodge describes methods that can be used to determine whether or not each individual understands the process used to determine numeric estimates.

As a final application of magnitude scaling, subjects were asked to rate characters from *The Simpsons* in terms of how funny they are. These results can be found in Appendix I.

Appendix B

Examples of Characteristic Exponents for the Power Law

Stimulus Modalities	Exponent	Description
Visual Length	1.00	Projected Line
Visual Area	0.70	Projected Square
Loudness	0.67	Sound pressure of 3000-Hz tone
Duration	1.10	White noise stimuli
Electric Shock	3.50	Current through fingers
Heaviness	1.45	Lifted weights
Pressure on Palm	1.10	Static forces on skin
Muscle Force	1.70	Hand-grip contractions
Warmth	1.60	Metal contact on arm
Cold	1.00	Metal contact on arm
Smell	0.60	Heptane
Taste	0.80	Sacchrine
Taste	1.40	Salt
Taste	1.30	Sucrose
Lightness	1.20	Reflectance of gray papers

Source: Stevens (1957)

Examples:

1. The exponent for visual length (1.0) means that doubling the length of a line will result in a doubling of the perceived estimates of length.

2. The exponent for loudness (0.67) means that when the volume of a sound is doubled, subjects perceive a 0.67 increase in loudness.

Appendix G

Geometric Means (Scale Values) of Perceived Hostility Towards the U.S.

Country	Numeric	Line Length
Iraq	279.16	309.44
Afghanistan	214.96	221.88
Saudi Arabia	99.73	93.91
Pakistan	99.49	110.45
Cuba	94.28	110.59
China	89.70	102.01
Bosnia	81.41	84.39
Russia	65.36	57.36
Israel	59.00	54.75
Germany	57.45	62.48
France	50.00	50.00
Japan	48.52	47.80
India	44.24	38.41
Panama	34.43	31.58
Spain	30.48	34.50
Mexico	29.13	33.35
Canada	18.08	19.16
Great Britain	16.35	23.03
Sweden	10.31	9.03
Australia	6.30	8.62

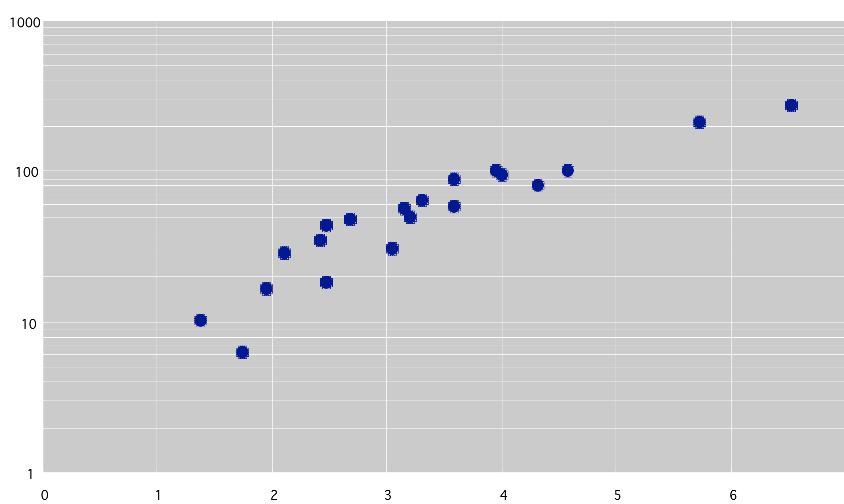
Higher numbers = more perceived hostility

Appendix H

Standardized Geometric Means of Perceived Hostility Towards the U.S.

	Standardized		
Country	Numeric	Line Production	Categorical
Iraq	3.07	3.17	2.47
Afghanistan	2.12	1.99	1.87
Saudi Arabia	0.42	0.25	0.49
Pakistan	0.41	0.48	0.98
Cuba	0.34	0.48	0.53
China	0.27	0.36	0.21
Bosnia	0.15	0.13	0.77
Russia	-0.09	-0.24	0.00
Israel	-0.18	-0.28	0.21
Germany	-0.21	-0.17	-0.12
France	-0.32	-0.34	-0.08
Japan	-0.34	-0.37	-0.48
India	-0.40	-0.50	-0.64
Panama	-0.55	-0.59	-0.68
Spain	-0.60	-0.55	-0.20
Mexico	-0.62	-0.57	-0.93
Canada	-0.79	-0.76	-0.64
Great Britain	-0.81	-0.71	-1.05
Sweden	-0.90	-0.90	-1.49
Australia	-0.96	-0.90	-1.21

Figure #5: Magnitude Scaling vs. Categorical Scalir

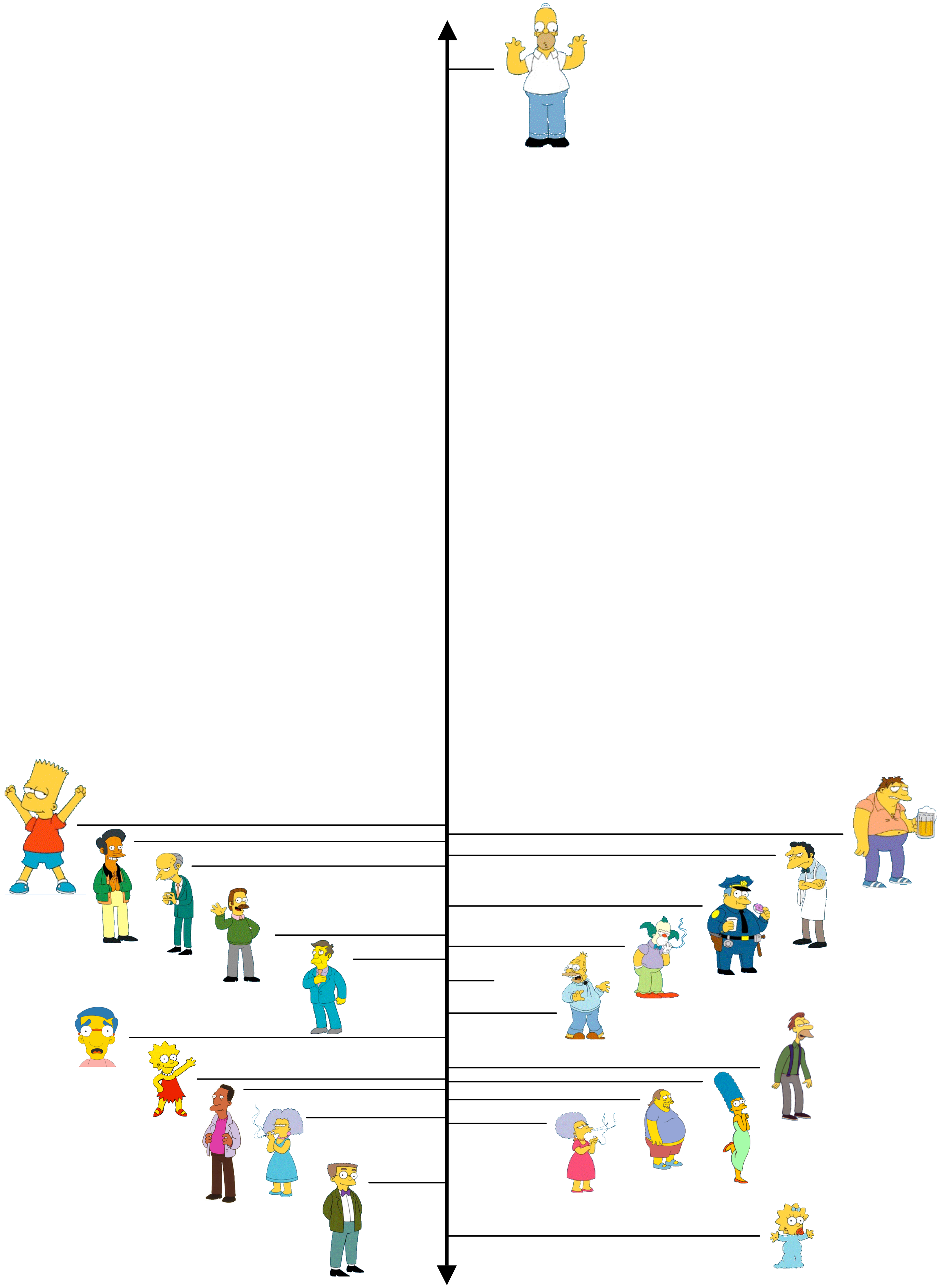


Appendix I

How funny are characters from *The Simpsons*?

Character	Geometric Mean
Homer	140.19
Bart	50.00 (reference)
Barney	48.59
Apu	47.80
Moe	46.44
Burns	44.82
Wiggum	41.76
Flanders	39.37
Krusty	38.54
Skinner	37.88
Abe	32.17
Milhouse	29.44
Lenny	26.44
Lisa	24.16
Marge	23.91
Carl	23.01
Comic	22.53
Selma	20.75
Patty	19.38
Smithers	14.97
Maggie	6.92

The following page shows a visual representation of these scale values.



References

Lodge, Milton (1981). *Magnitude scaling: quantitative measurement of opinions*. Beverly Hills: Sage

Stevens, S. (1957). *On the psychophysical law*. Psychological Review, 64

Sulfaro, V. & M. Crislip (1997). *How Americans perceive foreign policy threats: a magnitude scaling analysis*. Political Psychology, 18, (1)