Teorie & Modelli, n.s., XII, 1-2, 2007 (237-246)

# Test of Anderson's model of numerical rating

*Mirko Dai Prà*[*] (Padua)

In the method of numerical rating the participant rates the subjective value $\psi$ of predefined stimuli. Each stimulus is presented together with two fixed stimuli, the anchors. The subjective value $\psi_L$ of one anchor is smaller then the smallest $\psi$, and the subjective value $\psi_U$ of the other anchor is larger than the largest $\psi$. The measure of $\psi$ is a number $R$ selected by the participant from an ordered set of equidistant numbers. The smallest number $R_L$ and the largest number $R_U$ of this set are defined to be the measure of $\psi_L$ and of $\psi_U$, respectively. Participants select $R$ so that $R$ is proportional to $\psi$ considering that $R_L$ is the measure of $\psi_L$ and that $R_U$ is the measure of $\psi_U$.

Anderson (1982) proposed the following model for the selection of $R$. Participants would select each $R$ so that

$$\frac{R - R_L}{R_U - R} = \frac{\psi - \psi_L}{\psi_U - \psi}. \tag{1}$$

Equation 1 implies that

$$R = \psi \frac{R_L - R_U}{\psi_L - \psi_U} + \frac{R_U\ \psi_L - R_L\ \psi_U}{\psi_L - \psi_U}, \tag{2}$$

that is, when $\psi_L$, $\psi_U$, $R_L$, and $R_U$ are fixed, the model predicts that $R$ is a linear function of $\psi$. The present study was designed to test Equation 2.

## Experiment 1

### *Method*

*Participants*. Ten university students with mean age of 22 years, with standard deviation of 2.1 years, took part in Experiment 1. One was male.

[*] Department of General Psychology, University of Padua.

*Stimuli*. The experiment was run in a darkened room. The stimuli were achromatic squares with side length of 1.3 cm presented in the middle of a 32 × 19 cm rectangular area of the frontoparallel achromatic screen of a monitor (Apple Multiple Scan 1705) controlled by a computer (Macintosh 7200). A head-and-chin rest held viewing distance at 70 cm. The stimulus was between two anchors, which were achromatic squares equal in shape and size to the stimuli. The stimulus and the anchors were aligned horizontally. The gap between the stimulus and each anchor was 1.4 cm. The luminance of the screen was 0.1 cd/m$^2$. The luminance of the stimulus was 8, 16, 24, 32, or 40 cd/m$^2$. For each of these values, the luminances of the left and right anchors were 4 and 55, 2 and 75, 0.8 and 100, or 0.3 and 130 cd/m$^2$, respectively. These 20 combinations of stimulus luminance and pair of anchor luminances were presented in random order five times consecutively. That is, there were five consecutive blocks of 20 trials each. On each trial the stimulus and the anchors remained visible until the experimenter typed the participant's response. The subsequent stimulus with two anchors appeared 2 sec after the experimenter typed this response.

*Procedure*. Participants were asked to rate the luminous intensity of stimuli using the integer numbers in the range 1–100, with "1" and "100" being defined as the ratings of luminous intensity of the left and right anchors, respectively. Before the proper experiment began, five trials with stimuli selected at random were run for familiarization with the procedure. The first two blocks of trials were used as training trials. The responses obtained in these trials were discarded.

*Statistical analyses*. Analyses of variance were performed for this and the following experiments, which showed nonsignificant effects of the last three blocks of trials in all experiments. Effects of factors other than blocks and corresponding interactions were highly significant in all experiments.

## Results

Figure 1 shows the results from the last three blocks of trials. (Essentially, these results match those obtained in the first two blocks of training trials.) Symbols show rated luminous intensity, $R$, as a function of stimulus luminance for each pair of anchors. The left and right diagrams show the mean obtained $R$s and the predicted $R$s, respectively.

Predicted $R$s were calculated as follows. Equation 2 expresses $R$ as a function of the known values of $R_L$ and $R_U$ (1 and 100, respectively) and of unknown values of $\psi$. These unknown values may be estimated by the psychophysical law

$$\psi = k\,\phi^n, \tag{3}$$

where $k$ is a proportionality constant, $\phi$ the stimulus luminance, and $n$ the exponent. Using functional measurement and the bisection task, Anderson (1981, pp. 39-40) found that $n$ was 0.14. Equations 2 and 3 imply that

$$R = \phi^n \frac{R_L - R_U}{\phi_L^n - \phi_U^n} + \frac{R_U\,\phi_L^n - R_L\,\phi_U^n}{\phi_L^n - \phi_U^n}. \tag{4}$$

Each $R$ plotted in the right diagram in Figure 1 was calculated by Equation 4 using the luminance values reported above for $\phi$, $\phi_L$, and $\phi_U$ and using the value 0.14 for $n$.

In Figure 1, comparison of the obtained pattern of graphs with the predicted pattern of curves shows that Anderson's model of numerical rating accounts for the two main properties of the data: the rightward convergence and the crossover of the anchor graphs. However, the observed vertical range is greater than that predicted. This may reflect an inappropriate value of $n$, which is subject of considerable uncertainty for luminous intensity. The $R$s predicted by the model match imperfectly the corresponding $R$s

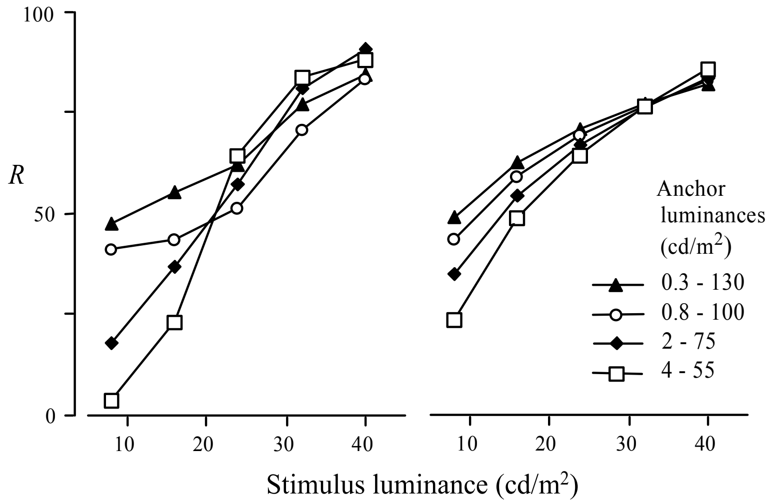


**Figure 1.** Obtained (left) and predicted (right) mean rating $R$ of luminous intensity as a function of stimulus luminance for each pair of anchor luminances.

obtained empirically. This imperfection is expected to occur for the following reasons. (i) Predicted *R*s vary with *n* and there is insufficient agreement about the value of *n*. (ii) Obtained *R*s depend on context effects such as ceiling effects and individual response preferences, which are not encoded in Anderson's model.

## Experiment 2

In the left diagram in Figure 1, graphs converge as the luminous intensity of stimuli increases. This convergence has two possible explanations.

*Explanation 1.* Cataliotti and Gilchrist (1995) hypothesized that the visual system transforms the highest luminance in the visual field to white. Since the luminance of the right anchor was always the highest luminance in the visual field, this hypothesis implies that participants compared the luminous intensities of the stimuli with the white seen in the right anchor. Thus, relative to this anchor, the luminous intensity of the stimulus with the largest $\psi$ was rated about the same for each pair of anchors, causing graphs to converge as the luminous intensity of stimuli increased.

*Explanation 2.* The psychophysical law for luminous intensity is negatively accelerated. Consequently, for each pair of anchors the difference between $\psi_U$ and the largest $\psi$ was smaller than the difference between the smallest $\psi$ and $\psi_L$. (To verify this statement one needs to transform the luminances used in Experiment 1 to the corresponding luminous intensities calculated by Equation 3 with $n = 0.14$ and with $k$ arbitrarily set for example at 1.) Accordingly, for each pair of anchors, the difference between $R_U$ and the mean $R$ assigned to the largest $\psi$ was smaller than the difference between the mean $R$ assigned to the smallest $\psi$ and $R_L$, thus causing graphs to converge as the luminous intensity of stimuli increased.

To test Explanations 1 and 2, Experiment 1 was repeated with the luminance of the background now set higher than each anchor luminance. All stimuli and anchors looked dark. Stimulus darkness increases as a negatively accelerated function of the absolute difference in luminance between stimulus and background. Participants rated stimulus darkness. Explanation 1 cannot predict these ratings since all anchor luminances were less than the highest luminance in the visual field. Explanation 2 predicts that graphs converge as stimulus darkness increases. If this prediction turns out to be true, one can conclude that Explanation 2 is preferable to Explanation 1.

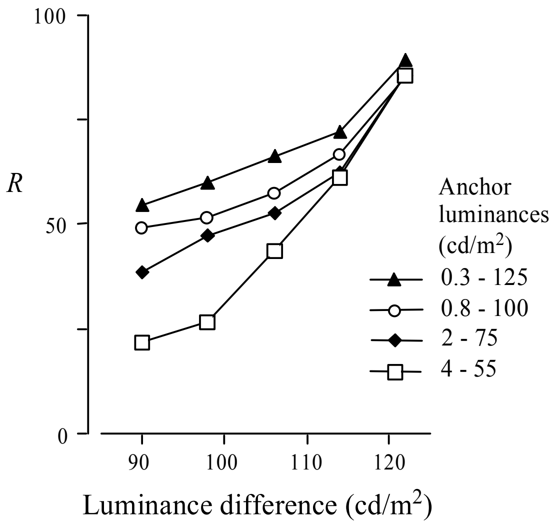


**Figure 2.** Mean rating *R* of darkness as a function of absolute luminance difference between stimulus and background for each pair of anchor luminances.

*Method*

Experiment 2 was a repetition of Experiment 1 with only the changes described below.

*Participants*. The participants were ten university students. None had participated in Experiment 1. Their mean age was 25.4 years, with standard deviation of 5 years. Four were male.

*Stimuli*. The luminance of the background was 130 cd/m$^2$. The highest anchor luminance was 125 cd/m$^2$. For each stimulus the anchor with the lower luminance was on the right of the stimulus.

*Procedure*. Participants were asked to rate the darkness of the stimulus using the integer numbers in the range 1–100, with “1” and “100” being defined as the ratings of darkness of the left and right anchors, respectively.

## Results

Figure 2 shows mean ratings *R* of darkness as a function of the absolute luminance difference between stimulus and background, for each pair of anchors. Graphs converge as stimulus darkness increases. Thus, the convergence of graphs in Figures 1 and 2 is explained by the psychophysical law better than by Cataliotti and Gilchrist's hypothesis.

The psychophysical law varies with the definition of the physical dimension associated to the sensation (Myers, 1982; Weiss, 1981). For example, possible equivalent physical dimensions that could be associated to darkness are the inverse of stimulus luminance or the luminance difference between stimulus and background. Since the *n* for these alternative physical dimensions is not securely known, no attempt is made here to use Equation 4 to predict the results reported in Figure 2.

## Experiment 3

As mentioned above, there is insufficient agreement about the value of *n* for luminous intensity – functional measurement and magnitude estimation provide estimates of *n* of 0.14 and 0.33, respectively (Anderson, 1981; Marks, 1974). A secure estimate of *n* entails a more robust test. Functional measurement and magnitude estimation show that *n* for perceived length of lines presented on a frontal plane is essentially equal to 1 (Anderson, 1977; Pitz, 1965; Schiffman, 1965). Since this *n* is a secure estimate, Experiment 1 was repeated using perceived line length in place of luminous intensity.

### *Method*

*Participants*. The participants were ten university students with mean age of 25.4 years, with standard deviation of 4 years. Four were female. None had participated in Experiments 1 or 2.

*Stimuli*. On each trial, the stimulus and two anchors were presented successively. Stimuli and anchors were black horizontal lines with width of 1 mm presented in the middle of a 102 × 52 cm rectangular gray area of the frontoparallel screen of a plasma monitor (NRC PlasmaSync 50MP2) controlled by a computer (Intel Pentium 4). The duration of each stimulus and each anchor was 1 sec. The shorter anchor appeared first, followed by the longer anchor and then by the stimulus. The time intervals between the two anchors and between the second anchor and the stimulus were 1 sec. The participant sat in front of the screen. Viewing distance was about 2 m.

The length of the stimulus was 12.5, 20.5, 28.5, 36.5, or 44.5 cm. For each of these lengths, the length of the first and second anchors were 9.3 and 55, 7.3 and 64, 5.3 and 73.5, or 3.3 and 83.5 cm, respectively. Thus, there were 20 combinations of stimulus and pair of anchors. Each combination after the first was presented 2 sec after the experimenter typed the participant's response. These combinations were presented in random order five times consecutively. That is, there were five blocks of 20 trials each.

*Procedure*. Participants were asked to rate the length of stimuli using the integer numbers in the range 1–100 with "1" and "100" defined, respectively, as the ratings of the length of the shorter and longer anchors that appeared immediately before the stimulus. The first two blocks of trials were training trials.

**Results**

Figure 3 shows the results obtained in the last three blocks of trials. Symbols show mean ratings $R$ of length as a function of stimulus length, for each pair of anchors. The left diagram shows obtained $R$s. The right diagram shows $R$s calculated by Equation 4 with $n = 1$.

Note that stimuli and anchors were selected such that, for each pair of anchors, the difference between $\psi_U$ and the largest $\psi$ was larger than the difference between the smallest $\psi$ and $\psi_L$. Thus for each pair of anchors the difference between $R_U$ and the mean $R$ assigned to the largest $\psi$ was larger than the difference between the mean $R$ assigned to the smallest $\psi$ and $R_L$, thus causing graphs to diverge as the length of stimuli increased.

It may be seen that Anderson's model accounts for the rightward divergence of the obtained graphs.

**Discussion**

Equation 1 assumes that participants evaluate the ratio between $\psi - \psi_L$ and $\psi_U - \psi$. Another model could assume that participants evaluate the difference between these sensory differences. The new model would read

$$(Q - R_L) - (R_U - Q) = c\,[(\psi - \psi_L) - (\psi_U - \psi)] \tag{5}$$

with $c$ a proportionality constant and $Q$ the integer number selected as a rating of $\psi$. Equation 5 yields

$$Q = c\,\psi + \frac{R_L + R_U - c\,(\psi_L + \psi_U)}{2}. \tag{6}$$

Thus, like Equation 1, Equation 5 – which encodes the operation of evaluation of a difference of sensory differences – predicts that $Q$ is a linear function of $\psi$ when $\psi_L$, $\psi_U$, $R_L$, and $R_U$ are fixed.

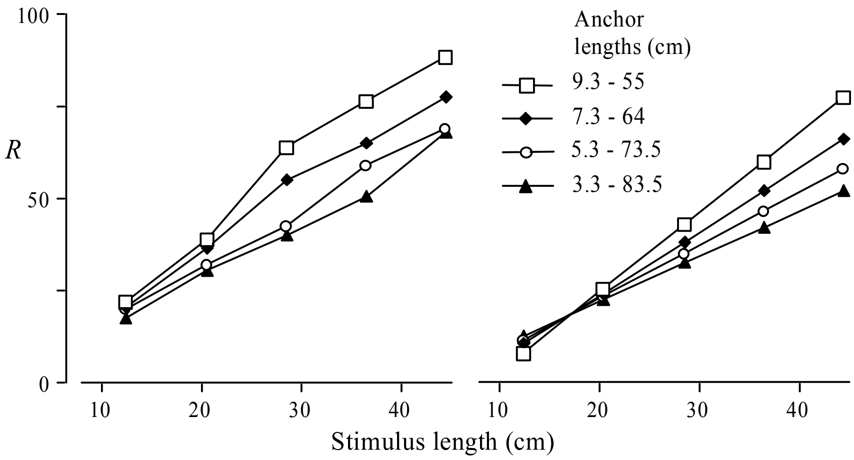


**Figure 3.** Obtained (left) and predicted (right) mean rating *R* of line length as a function of stimulus length for each pair of anchor lengths.

Equation 2 accounts for both the convergence and the divergence of anchor graphs in Figures 1 and 3, respectively. Equation 6 cannot do this because its slope *c* is a constant.

In Figure 3, empirical graphs in the left diagram do not intersect while predicted graphs in the right diagram intersect. This disagreement of results with Equation 2 suggests that participants may have imperfectly evaluated the ratio between $\psi - \psi_L$ and $\psi_U - \psi$.

Participants could have made a compromise between the evaluation of the ratio between $\psi - \psi_L$ and $\psi_U - \psi$ (Equation 1) and the evaluation of the difference between these differences (Equation 5). If this hypothesis is true, the integer number selected as a rating of $\psi$ would be the rounded value of

$$J = w\,R + (1 - w)Q \tag{7}$$

with *w* a weight varying between 0 and 1.

For example, for $c = 1.3$ and $w = 0.6$, Figure 4 shows the rating *J* predicted by Equation 7 as a function of stimulus length for each pair of anchors. Equation 7 predicts nonintersecting diverging curves in agreement with the empirical results reported in the left diagram in Figure 3. Equation 7 accounts for both the crossover and the noncrossover of anchor graphs.

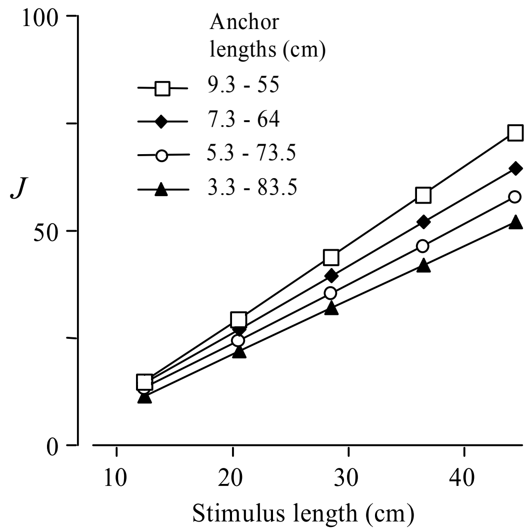


**Figure 4.** Predicted rating $J$ of line length as a function of stimulus length for each pair of anchor lengths.

## References


Anderson, N. H. (1977). Failure of additivity in bisection of length. *Perception & Psychophysics*, *22*, 213-222.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.

Cataliotti, J., & Gilchrist, A. (1995). Local and global processes in lightness perception. *Perception & Psychophysics*, *57*, 125-135.

Marks, L. E. (1974). *Sensory processes*. New York: Academic Press.

Myers, A. K. (1982). Psychophysical scaling and scales of physical stimulus measurement. *Psychological Bulletin*, *92*, 203-214.

Pitz, G. F. (1965). Magnitude scales of line length. *Psychonomic Science*, *2*, 213-214.

Schiffman, H. R. (1965). Size estimation and the size of the measuring unit. *Psychonomic Science*, *3*, 479-480.

Weiss, D. J. (1981). The impossible dream of Fechner and Stevens. *Perception*, *10*, 431-434.

## Abstract

Participants rated numerically the sensory intensity of stimuli for different pairs of anchor stimuli. For each of these pairs there was one graph relating mean rated sensory intensity to stimulus intensity. The pattern of these graphs turned out to be close to that predicted by Anderson's model of numerical rating. A model of wider applicability is proposed which incorporates Anderson's model.

## Riassunto

I partecipanti hanno valutato numericamente l'intensità sensoriale di stimoli per differenti coppie di stimoli àncora. Per ciascuna di queste coppie c'era un grafico che poneva in relazione la media della valutazione della intensità sensoriale alla intensità dello stimolo. La configurazione di tali grafici è risultata simile a quella prevista dal modello della valutazione numerica di Anderson. Viene proposto un modello di più ampia applicabilità che incorpora quello di Anderson.

**Aknowledgment.** I wish to thank Prof. Norman Anderson for useful comments.

**Address.** Mirko Dai Prà, University of Padua, Department of General Psychology, via Venezia 8, I-35131 Padova (daipra@psy.unipd.it).