DO STIMULUS FREQUENCY EFFECTS OCCUR WITH LINE SCALES?¹

Gert Haubensak
Justus Liebig University of Giessen
gert.haubensak@t-online.de

Abstract

Two experiments are reported in which the participants judged the number of Xs contained in a frame. Experiment 1 used a five-point category scale whereas Experiment 2 used a 150-mm line scale. The line scale ratings were all lower than the category ratings but stimulus frequency effects occurred both times. The results are inconsistent with theories that ascribe frequency effects to the special nature of category scales. The results, however, are consistent with the hypothesis that frequency effects arise from two tendencies, the tendency to select a response closer to one’s reference point than it should be (the response contraction bias; Poulton, 1989) and the tendency to keep to one’s first judgments (the consistency principle; Haubensak, 1992).

Stimulus frequencies are known to affect judgments. When the frequency distribution is positively skewed (the smaller stimuli presented more frequently), the same stimuli are rated higher than when the distribution is negatively skewed (the larger stimuli presented more frequently), even though both distributions have the same range (Parducci, 1956).

Nearly all studies of the frequency, or skewing, effect have used category scales. Category scales require the subjects to assign a large number of different stimuli to a relatively small number of discrete categories (usually 3 - 9). Some think that this feature is responsible for the effect. The magnitude of the effect declines with the number of categories indeed (Parducci & Perrett, 1971; Parducci & Wedell, 1986). Stimulus frequency effects should then be completely absent when continuous line scales are used, because the subjects have an almost infinite number of “categories” to choose from.

There is mixed evidence for skewing effects with line scales, however. Anderson failed to find such effects in 1982, but Schifferstein and Frijters (1992) found some in sweetness ratings. Sweetness ratings, however, may be something special. The popular sip-and-spit procedure requires the subjects to rinse their mouths with demineralized water in a standardized manner after each stimulus. This procedure may cause bias to the results because the subjects tend to incorporate the water into their context for judgment (Schifferstein, 1992). In addition, relatively large intertrial intervals have to be used in order to prevent adaptation effects. (At least a half minute is recommended). To avoid these kinds of problems, I chose a different attribute, numerosness. The participants were to judge the number of crosses contained in a frame. Experiment 1 used a five-point category scale whereas Experiment 2 used a line scale that had no breaks or interruptions.

**Experiment 1: Five-Point Category Scale**

**Method**

**Subjects.** Thirty-two younger students of psychology (15 men and 17 women) participated in fulfillment of course requirements.

**Apparatus.** The stimuli were presented on a 21-in. Sony F512 color monitor driven by a commercial PC. The participants sat approximately 60 cm from the screen. The room was dimly lit.

**Stimuli and Procedures.** The stimulus material consisted of 8 computer-generated frames. Each frame contained a number of diagonal crosses, 3 mm wide and 3 mm high. The crosses (Xs) were randomly distributed over the 625 possible points of a 25-column x 25 row matrix (150 mm wide x 150 mm high) with the restriction that each successive block of five rows contained just one fifth of the crosses specified. The number of Xs was a positive exponential function of the ordinal number of frame. Frame 1 contained 10 and Frame 8 contained 275 Xs. (For the exact numbers see Table 1.) The positions were reshuffled on every presentation. Thus, no image was presented to a subject more than once.

Below the matrix, five radio-buttons labeled -, -/0, 0, 0/+ and + were arranged in a horizontal row (see Figure 1 below). The task was to assign each frame to one of the five categories represented by the radio-buttons. The verbal meaning of the labels was explained to the subjects before the presentations started. The symbols meant *few, slightly fewer than average, slightly more than average,* and *many.*

The participants were divided into two groups at random. Group POS (n = 8) received the frames in a positively skewed distribution. Group N (n = 8) received the same frames in a negatively skewed distribution. The exact frequencies are given in Table 1. Presentation order was random. The participants made their judgments by clicking the appropriate button with the mouse. The mouse click made the frame disappear and after 3 s the next one appeared.

<table>
<thead>
<tr>
<th>Frame</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Xs</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>45</td>
<td>70</td>
<td>110</td>
<td>175</td>
<td>275</td>
</tr>
<tr>
<td>Presentation frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group POS</td>
<td>20</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Group NEG</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**Results**

First, I calculated the mean rating for each frame. Figure 3 shows the results for both Groups POS and NEG. The ordinal numbers on the abscissa are roughly proportional to the log
number of Xs contained in a frame. There is clear evidence of a stimulus frequency effect. Next, I calculated the judgment level by averaging the mean ratings across frames. The difference between the positive and the negative set was 0.60 category units. The standard deviation of the levels was 0.30 units in the positive set and 0.26 units in the negative set.

Experiment 2: 150-mm Line Scale

Method

Experiment 2 was identical with Experiment 1 except that the participants made their judgments by adjusting the slider of a horizontal track bar. The track bar, 150 mm wide and 4 mm thick, was locate below the matrix (see Figure 2). The left end of the dark-shaded working area was labeled - (“few”) and the right end + (“Many”). The working area itself was not provided with any tick marks. There was room (2 mm) left on either side. Thus the participants could extend their evaluations beyond that area if necessary. Again, the meaning of the symbols + and - was explained to the participants at the outset. To mark a position on the scale, the slider had to be moved back and forth with the mouse. When satisfied with their settings, the participants clicked the ok command button to the right of the track bar. Upon this, the slider returned to its starting point (the left end of the line) and the frame disappeared from the screen and after 3 s the next one appeared.

Results

Only 6 of the 16 subjects made use of more than 90 percent of the line. One subject restricted herself to slightly less than 50 percent of the total range. In Experiment 1, all the subjects had made full use of the categories. To make the results comparable, I transformed the segment of the line scale actually used by a subject into a five-point category scale so that the endpoints of the segment corresponded to the midpoints of Category 1 or 5. In terms of an equation,
Abbildung 3. Experiment 1. Five-point category scale. Ratings of positive set shown by open circles, negative set by open squares. Number of Xs is an exponential function of the ordinal stimulus number on the abscissa.

Abbildung 4. Experiment 2. 150-mm line scale. Settings mapped onto five-point category scale. Frequency effect even stronger than in Experiment 1.

Discussion

The results are inconsistent with the finding that frequency effects decline with the number of categories (Parducci & Wedell, 1986). Perhaps line scales are not just category scales with a very large number of categories. The results are compatible with the consistency model.

\[
c' = \text{trunc}(\frac{c - c_{\text{min}}}{c_{\text{max}} - c_{\text{min}}} + 1.5)
\]
of judgment (Haubensak, 1992). The model assumes that people are cautious with their first judgments. Thus they select a response closer to their reference point than it should be (cf. Poulton, 1989). With category scales, the reference point is the center of the scale. Usually the point is labeled “medium” or “average”. As a consequence, the larger stimuli are rated too low and the smaller stimuli too high if they come first in the presentations. Frequency and spacing (density) effects occur because the larger stimuli are more likely to come first when the distribution is negatively skewed than when it is positively skewed. Since people attempt to be consistent, they continue to rate the stimuli too high or, respectively, too low.

With line scales, the subjects’ reference point may not be the center of the line scale but the left end of the scale. After all, we are accustomed from childhood to read and write from the left to the right. As a consequence, the first line scale responses are shifted to the left. With negatively skewed distributions, this shift will primarily affect the judgment of the larger stimuli. Therefore, mean judgments are lower than when the stimuli are positively distributed.

References


