DEPTH OF SEQUENTIAL DEPENDENCIES IN PSYCHOPHYSICAL JUDGMENT

Peter Petzold
University of Giessen

Abstract

In a first experiment, subjects had to rate the size of squares. We found that the depth of sequential dependencies depended on the judgment task. Whereas for magnitude estimation only the immediately preceding stimulus-response event was included in the judgment process, events up to two trials back were incorporated for category judgment. In a second experiment, squares of two categories differing in color were presented. Under these conditions, category-specific sequential effects were found. Sequential dependencies were stronger when the current stimulus and the preceding stimulus belonged to the same category. More specifically, the depth of sequential effects was affected by the categories of previous stimuli. For category judgment, when the stimuli one and two trials back did not share the category with the current stimulus but the stimulus three trials back did, the event three trials back affected the judgment. However, if the stimuli one and two trials back belonged to the same category as the current stimulus, the event three trial back exerted no influence.
The results are discussed in a framework in which preceding events play the role of referents.

There is some evidence indicating that not only the immediately preceding stimulus-response pair affects the judgment of the current stimulus, but also events more than one trial back. For instance, analyzing response errors in absolute judgment, Ward and Lockhead (1971) found that sequential effects extended to at least five trials back. On the other hand, Jesteadt, Luce and Green (1977) concluded from regression analyses of magnitude estimations that sequential dependencies were limited to the immediately preceding event.
What is the reason for the different results concerning the depth of sequential effects? Possibly, the depth depends on the judgment task. Whereas Lockhead and Ward referred to an experiment on absolute judgment, Jestead, Luce and Green focused on experiments on magnitude estimation. The present study is meant to examine whether the depth of sequential dependencies is really affected by the task.

Experiment 1

Method
Subjects had to judge the size of squares that varied from 50 mm to 83 mm in steps of 3 mm. Subjects performed in one of two judgment tasks: category judgment on a 5-category scale and magnitude estimation. One session consisted of three blocks lasting about 10 minutes each. There were 200 trials per block.
Measurement of sequential dependencies
To determine the depth of sequential dependencies, it may be determined up to which lag g partial correlations between current responses r(t) and preceding responses r(t-g) as well as between r(t) and preceding stimuli s(t-g) are significant. In the present experiment, partial correlations between r(t) and r(t-g) and between r(t) and s(t-g) turned up to be significant up to $g = 8$. Does this mean that stimulus-response events up to 8 trials back may be included in

the judgment process? To answer this question, we have to take into account that pseudosequential effects may appear as artifacts. Gregson (1976) and Haubensak (1992) have pointed out that judgments averaged over subjects show sequential dependencies even if individual data do not when some individuals tend to judge the stimuli generally high and others tend to judge generally low. Such pseudosequential effects may even appear in analyzing individual data averaged over an experiment if systematic response shifts occur in the course of the experiment. Therefore, artifacts produced by averaging the data must be eliminated in determining the depth of sequential dependencies.

There are two ways to disentangle sequential effects and pseudosequential effects.

1) We can estimate the magnitude of pseudosequential effect and compare them with the values obtained in an experiment. As long as the empirical values are greater than those calculated for pseudoeffects, we can assume that the corresponding stimulus-response events are actually included in the judgment process. We made the following steps. First, individual mean judgments were calculated for each stimulus within each of the three blocks of the experiment separately. Second, to produce stimulus-response series without sequential effects, these mean judgments were used as scale values in a computer program based on a Thurstone-like model. This was done for each individual and for each block of the experiment. Third, the series from the three blocks were combined for each individual to calculate partial correlations between r(t) and r(t-g) that are due to drifts of the judgment scale. Then the calculated pseudoeffects were compared with the empirical findings.

2) A second way to eliminate artifacts produced by averaging data was proposed by Schifferstein and Kuiper (1997). One can correct for differences in scale usage by calculating standardized responses which are brought back to a mean of 0 and a standard deviation of 1. Following this procedure, we determined mean m and standard deviation σ of the responses r and calculated standardized responses $r_s = (r - m)/\sigma$ for each subject and each part of a session separately. Then, these standardized responses were used to analyze sequential dependencies.

<u>Results</u>

Empirical correlations and correlations calculated from the simulation of pseudosequential effects are represented in Figures 1and 2. Mean values averaged over the subjects are shown. The curves differ for category judgment and magnitude estimation.

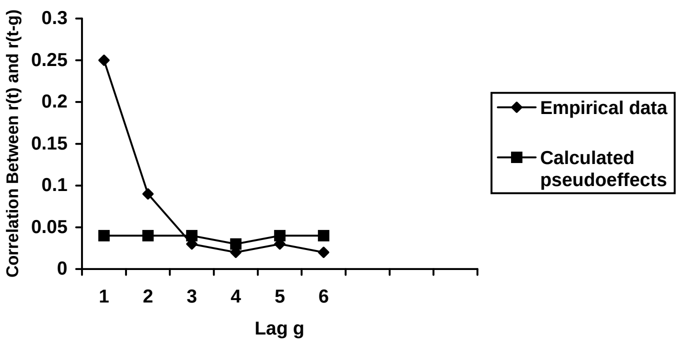

Figure 1. Empirical partial correlations between the current response and preceding responses of the lag g and calculated pseudoeffects for category judgment

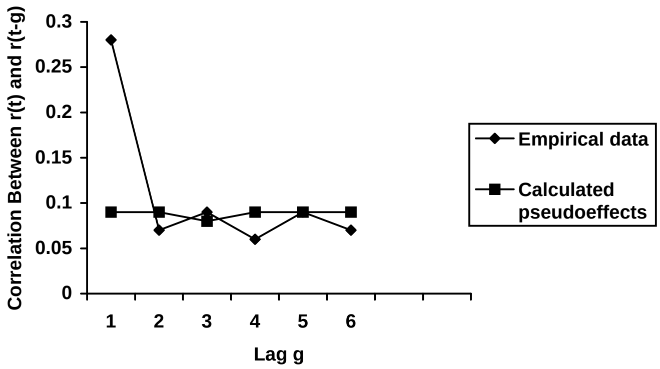

Figure 2. Empirical partial correlations between the current response and preceding responses of the lag g and calculated pseudoeffects for magnitude estimation

In the case of category judgment, the empirical correlations for lags 1 and 2 have higher values than those produced by pseudosequential effects. For higher lags, the values found in the experiment correspond approximately to the values of the correlations produced by a drift of scale. From this follows that sequential effects up to lag 2 are caused by the judgment process and that the correlations for events more than two trials back are produced by averging the data. In the case of magnitude estimation, the empirical value is higher than the calculated pseudosequential effect only for lag 1 and the correlations reach the level of pseudoeffects already at lag 2. These findings sugest that the depth of sequential effects is different for category judgment and for magnitude estimation. Whereas events up to lag 2 seem to be included in the judgment process for category judgments, only events of lag 1 appear to serve as referents for magnitude estimation.

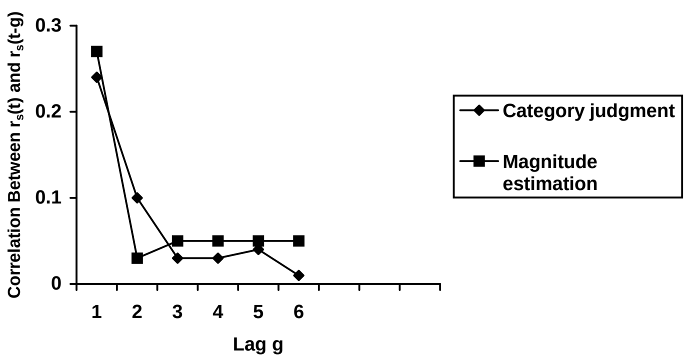

Figure 3. Partial correlations between the current standardized response and preceding standardized responses in dependence on the lag g.

An equivalent result is obtained by using the second method for eliminating pseudosequential effects. Correlations of standardized responses are represented in Figure 3. The curves indicate that the correlation between successive standardized responses decrease to near zero for lags greater than 2 in case of category judgment and for lags greater than 1 in case of magnitude estimation. The reason for not reaching the zero level entirely is probably that there is a small shift of the judgment scale also within a block. Therefore, the pseudoeffects are not completely eliminated by the procedure applied.

Discussion

The results of the present experiment suggest that the depth of sequential dependencies is two for category judgment and one for magnitude estimation The reason for the different depth may be that there are different decision processes in the two judgment tasks. In the case of category judgment, the relative distance of the stimulus from the two end-points of a subjective range seems to determine the response. Consequently, two prior stimuli can be involved in the formation of the range. This assumption is supported by interactions between the stimulus-response events one and two trials back (Petzold & Haubensak, 2001).For magnitude estimation the relation to one prior event as referent seems to be decisive as is proposed by the response ratio rule (Luce & Green, 1974) or the dynamic judgment model by DeCarlo & Cross (1990). The exact nature of the difference in the depth of sequential effects needs further clarification. Especially, it should be explored whether the difference found in the present experiment is typical for categorical judgment and magnitude estimation. Perhaps, there are other conditions that affect the depth as well. Such an additional condition could be the categorization of stimuli. This condition was considered in a further experiment.

Experiment 2

Method

The stimuli were squares of different size. Two sets of stimuli were presented, filled and empty squares, which differed in range . For half the subjects, 8 filled squares ranged from 50 mm to 71 mm in steps of 3 mm and 8 empty squares ranged from 62 mm to 83 mm, for the other half filled squares and empty squares were reversed. The overlapping area of the two stimulus categories (squares from 62 to 71 mm) ensured that the effect of stimulus category on judgments could be studied for four squares of equal size which differed only in the category.

The task was to rate the size of squares on a 5-step scale. Subjects were instructed to compare each filled square only with the other filled squares and each empty square only with the other empty squares.

Results

Effect of Stimulus Categories on Sequential Dependencies.

To measure sequential dependencies, partial correlations between $r_s(t)$ and $r_s(t-1)$ were calculated. These correlations were determined separately for each combination of categories to which the current stimulus s(t) and the preceding stimulus s(t-1) may belong. To make sure that the mean distance between the current stimuli and the preceding stimuli was equal in all cases, only those stimuli were included in the analysis whose values fell into the overlapping area of the two stimulus categories (squares from 62 mm to 71mm). After a z-transformation, mean values of the correlation between $r_s(t)$ and $r_s(t-1)$ averaged over subjects and categories of s(t) were calculated (Figure 4, left side). One can see that the correlations are higher for identical categories than for different categories. An analysis of variance revealed that this difference is significant.

The influence of the presentation two trials back was analyzed in the same way as the immediately preceding presentation. After performing a z-transformation the of correlations, mean values of the partial correlations between $r_s(t)$ and $r_s(t-2)$ were calculated. (Figure 4, right side). As in the case of the immediately preceding event, there is a significant difference between the correlations of equal and different stimulus categories. This equivalence in the influence of stimulus categories on sequential dependencies for one and two trials back supports the finding that both the immediately preceding event and the event two trials back meet the same function in the judgment process.

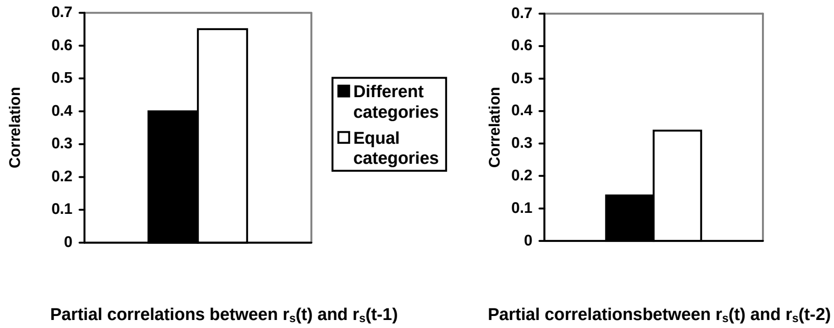

Figure 4. Partial correlation between the current response and preceding responses for equal and different categories of the current stimulus and the preceding stimulus.

Interaction between s(t-1) and s(t-2).

There was an interaction between the membership to categories of the immediately preceding stimulus, s(t-1), and of the stimulus two trials back, s(t-2). The correlation between r_s(t) and r_s(t-1) was lower if s(t-1) and s(t-2) belonged to the same category as s(t) than if only s(t-1) had the category in common with s(t). In the first case, the mean value averaged over subjects and the category of current stimuli was 0.38, in the latter case, 0.60. A corresponding result has been found for the correlation between r_s(t) and r_s(t-2). The values were lower if s(t-1) and s(t-2) belonged to the same category as s(t) than if only s(t-2) shared the category with s(t). The values were 0.14 and =.34, respectively. These findings indicate that the events one and two trials back operate concurrently in the judgment process.

The influence of the event three trials back.

The data showed that in the present task also the event three trials back affected the judgment of the current stimulus. However, this was only the case if the stimulus s(t-3) had the same category as the current stimulus s(t) and the category of the stimuli s(t-1) and s(t-2) differed from that of s(t). This is demonstrated in Figure 5. We find for events up to three trials back that the partial correlation between the standardized current responses and the standardized responses of the lag g is higher when only the stimulus of the lag g has the same category as the current stimulus. This result suggests that in the present judgment task also the event three trials back is included in the judgment process equivalent to the events one and two trials back

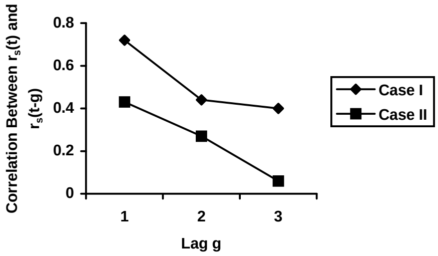

Figure 5. Partial correlations between the current standardized response and preceding standardized responses in dependence on the lag l for two cases. Case I: Only the stimulus s(t-g) has the same category as the current stimulus s(t). Case II: Also other preceding stimuli share the category with s(t).

Discussion

The findings of the present experiments show that the depth of sequential effects is dependent on the judgment task. Whereas in magnitude estimation only the immediately preceding event affected the judgment of the current stimulus, events up to two trials back were included in the judgment process for category judgment. Also the existence of two stimulus categories modified the depth of sequential effects. In case of category judgment, the depth was at least three when s(t-3) had the same category as s(t) and s(t-1) or s(t-2) had not. One can speculate that appropriate traces of preceding events are searched as referents. If the events one or two trials back have not the same category as s(t) and are considered not appropriate, the search goes further up to at least three trials back. These findings are consistent with the predictions of the multiple standards model (Petzold & Haubensak, 2001).

References

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General, 119*, 375-396..

Gregson, R. A. M. (1976). Psychophysical discontinuity and pseudosequence effects. *Acta Psychologica*, *40,* 431-451.

Haubensak, G. (1992). Sequenzeffekte in absoluten Urteilen? Kritisches zur Methode. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *39*,101-113.

Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of the Experimental Psychology, Human Perception and Performance, 3,* 92-104.

Luce, R. D., & Green, D. M. (1974). The response ratio hypothesis for magnitude estimation. *Journal of Mathematical Psychology*, *11*, 1-14.

Petzold, P., & Haubensak, G. (in press). Short-term and long-term frames of reference: A multiple standards model. In Ch. Kaernbach, E. Schröger, & H. Mueller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition*. Mahwah, N. J. : Erlbaum

Schifferstein, H. N. J. & Kuipers, W. E. (1997). Sequence effects in hedonic judgments of taste stimuli. *Perception and Psychophysics*, *59*, 900-912.

Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception and Psychophysics*, *9*, 73-78.

Acknowledgment. This study was supported by the Deutsche Forschungsgemeinschaft (Ha 936/6).