Sequence effects in hedonic judgments of taste stimuli

HENDRIK N. J. SCHIFFERSTEIN and W. ERNO KUIPER
Agricultural University, Wageningen, The Netherlands

Sequential dependencies in taste research may be different from those obtained in other modalities, due to the long interstimulus intervals and the intermediate rinses. In two experiments, subjects judged the pleasantness of 50 aqueous solutions on 150-mm line scales. During data analyses pseudo-sequence effects arose, because data were aggregated over individuals and because the first trials of the experimental sessions deviated from the rest. After correcting for the pseudo-sequence effects, robust regression analyses revealed small but significant sequential dependencies. The current response deviation was positively related to previous response deviations and negatively related to previous subjective, internal representations.

In experiments in which subjects provide direct estimates of perceived intensity, responses obtained at the current trial t (R_t) are often positively correlated with previous responses (R_{t-1}) and negatively correlated with previous stimulus intensities (I_{t-1}). Empirical investigations of these sequential dependencies have mostly been conducted using auditory or visual stimuli (e.g., DeCarlo & Cross, 1990; Jesteadt, Luce, & Green, 1977; Mori & Ward, 1990; Ward, 1979). Extrapolating these outcomes to taste and smell research may be premature, because the typical characteristics of experiments on chemosensation can affect the pattern of sequential dependencies.

In chemosensory research, the time intervals between successive stimulus presentations are large (30 sec or more) compared with those used in research on hearing and vision. In addition, subjects usually rinse their mouths with water between taste stimuli. Both these measures are taken to prevent sensory adaptation. However, they can also affect the pattern of sequential dependencies. For example, DeCarlo (1992) showed that the size of the interstimulus interval affects the size of sequential effects. In addition, due to the large interstimulus interval, relatively few stimuli are evaluated in one session. Consequently, individual analyses lack statistical power, and data from different subjects are combined, under the assumption that sequential dependencies are similar for all subjects. These typical characteristics of data collection and analysis may affect the type and size of sequence effects.

The main objective of the present study is to provide a detailed description of sequential dependencies in affec-tive responses to gustatory stimuli. Although several authors have indicated that sequential effects are likely to occur in hedonic judgments (e.g., Berlyne, 1973; Helson, 1973), such analyses are lacking in the current literature. The idea of sequential dependencies is captured by Bacon, Rood, and Washburn's (1914) "law of affective contrast" which states that "the pleasure of an agreeable experience is heightened if it is preceded by a disagreeable experience, and an impression in itself unpleasant may be felt as pleasant if a more unpleasant state has been its antecedent" (p. 290). In Beebe-Center's (1929) "law of affective equilibrium," however, "the affective value of the experiential correlate of a stimulus varies conversely with the sum of the affective values of those experiences preceding this correlate which constitute with it a unitary temporal group" (p. 64). According to Beebe-Center's experiments on odorants and Harris's (1929) study on colors, the hedonic contrast should be attributed to the entire set of preceding stimuli (context effect), and not to separate, consecutive stimulus presentations (sequential effect).

In the present study, sequential dependencies in hedonic judgments are studied for stimulus sets eliciting multiple taste qualities (Experiment 1) or one, sweet quality (Experiment 2). In our analyses, we start out from a two-stage stimulus–organism–response (S–O–R) model of psychological judgment. In the first stage, each stimulus S_{it} is represented on an internal, hedonic continuum (S–O: the psychohedonic function). In the second stage, the internal representation I_{it} is transformed into an observable response R_{it} (O–R: response output function). The subscripts refer to a particular stimulus (i) presented at a specific trial (t).
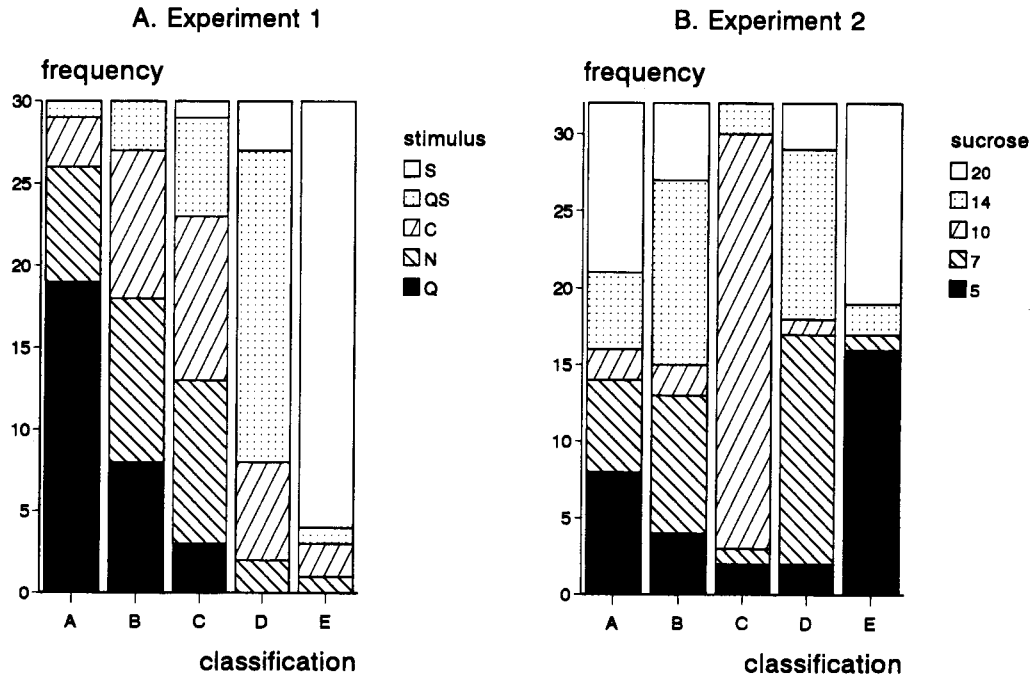
METHOD

Subjects

In Experiment 1, thirty untrained paid volunteers, 12 men and 18 women, ranging in age from 18 to 49 years (median age 22.5) participated. In Experiment 2, thirty-six paid or unpaid volunteers, 15 men and 21 women, ranging in age from 18 to 49 years (median age

A. Experiment 1

frequency

stimulus
☐ S
☐ QS
▨ C
◩ N
■ Q

classification

B. Experiment 2

frequency

sucrose
☐ 20
☐ 14
▨ 10
◩ 7
■ 5

classification

Figure 1. The frequency with which each stimulus is classified as A, B, C, D, or E (A = *least preferred*, E = *most preferred*). Panel A shows the results for Experiment 1. Q, QHCl; N, NaCl; C, citric acid; and S, sucrose. Panel B shows the results for Experiment 2, in which the stimuli were sucrose solutions varying in concentration level (% w/v).

22) participated. They were students and employees of the Agricultural University. The subjects were naive with respect to the substances used and the purpose of the study.

Stimuli

In both studies, five different stimuli were used. Experiment 1 employed a heterogeneous stimulus set containing stimuli that varied considerably in pleasantness. Mean pleasantness ratings were expected to increase in the following order: 120 μM quinine hydrochloride (QHCl, Sigma Q-1125) < 0.16 M NaCl (Merck 6404) < 2.5 mM citric acid (Merck 244) < a mixture of 60 μM QHCl and 10% w/v sucrose (Merck 7653) < 10% w/v sucrose (Schifferstein, 1995). In Experiment 2, the stimuli differed in sucrose concentration only (Merck 7653): 5%, 7.1%, 10%, 14.1%, and 20% w/v. Group mean hedonic responses for sweetened lemonade are found to peak at 10% sucrose. The standard error of this mean is large, indicating that subjects differ considerably in their preferred sucrose level (Pangborn, 1980). Therefore, the two stimulus sets are expected to differ with regard to the number of subjects with equal preference orderings.

Solutions were prepared in demineralized water at least 24 h before tasting, and were stored in a dark, refrigerated room at 4°C for no longer than 3 days.

Procedure

Subjects were instructed to judge the pleasantness of each stimulus on a 150-mm hedonic line scale. The left and right ends of the scale were labeled *not pleasant at all* and *extremely pleasant*, respectively. The subjects were requested to rinse their mouths thoroughly with demineralized water after each stimulus. The stimuli were presented at room temperature (~20°C) in polystyrene medicine cups. Each cup contained 10–12 ml of solution. One stimulus was presented every 60 sec in Experiment 1, and every 50 sec in Experiment 2. Every solution was judged 10 times by each subject

during a 1-h session. The order in which the stimuli were presented was randomized and differed between subjects.

RESULTS

Preliminary Analyses

For each subject, a one-way analysis of variance was performed to check whether the stimuli differed in experienced pleasantness. In Experiment 1, the stimulus main effect was significant for every subject [$F(4,45) >$ 5.57, $p \leq .001$]. The results of 4 subjects in Experiment 2 whose test yielded no significant outcome ($p >$ .05) were discarded.

Individual preference orders were determined for the five stimuli from the mean individual hedonic ratings. The data were recoded: Stimulus codes were replaced by Stimulus A–E (A = *least preferred* and E = *most preferred*). Figure 1 shows the frequency with which each stimulus was classified as A, B, C, D, and E. As expected, Figure 1 shows that the degree of agreement among subjects was larger in Experiment 1 than in Experiment 2. Experiment 2 suggests a bimodal distribution of subjects, with some subjects preferring high sucrose concentrations and some subjects preferring low concentrations. After recoding, the data were aggregated over subjects.

Stimulus Analysis

The mean response on trial t, averaged over solutions, subjects, and replications, was calculated as a function of the preceding stimuli presented at $t - 1$ up to $t - 6$ (see,

A. Stimulus analysis

hedonic response

B. Response analysis

hedonic response
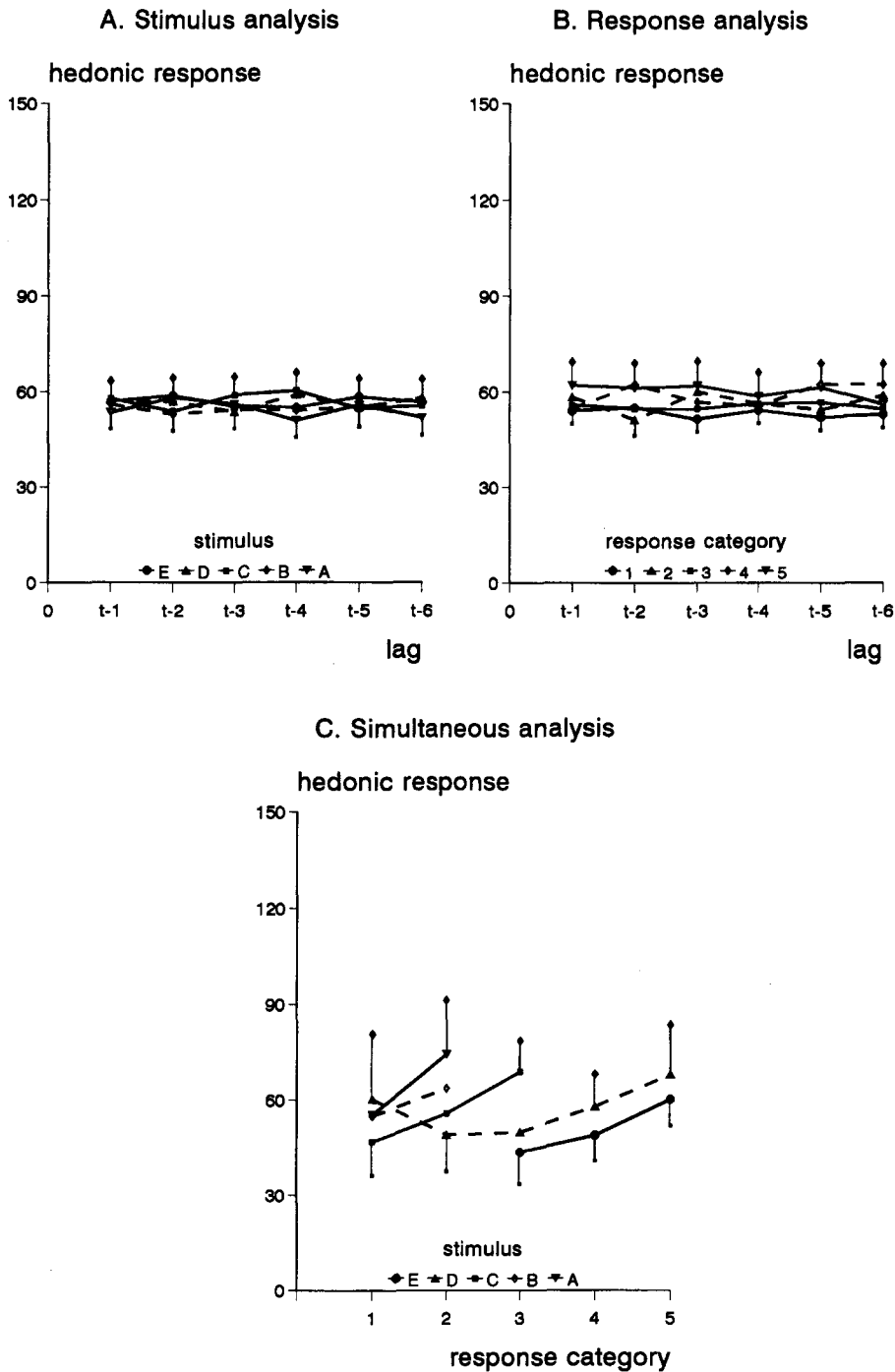

C. Simultaneous analysis

hedonic response

Figure 2. Sequential stimulus and response effects for hedonic judgments of the stimulus set eliciting multiple taste qualities (Experiment 1). The mean response on trial t (R_t), averaged over stimuli, is depicted as a function of the stimulus level (panel A) or the response level (panel B) on trial $t - k$, and as a function of both the preceding stimulus and response on trial $t - 1$ (panel C). The stimuli range from A (*least preferred*) to E (*most preferred*). The response categories correspond to $1 = R < 30$; $2 = 30 \le R < 60$; $3 = 60 \le R \le 90$; $4 = 90 < R \le 120$; $5 = 120 < R$. Mean responses are depicted only if the number of observations exceeds 25. For the outermost stimuli, the sizes of 95% confidence intervals are indicated by error bars (± 2 *SE*).

A. Stimulus analysis

B. Response analysis

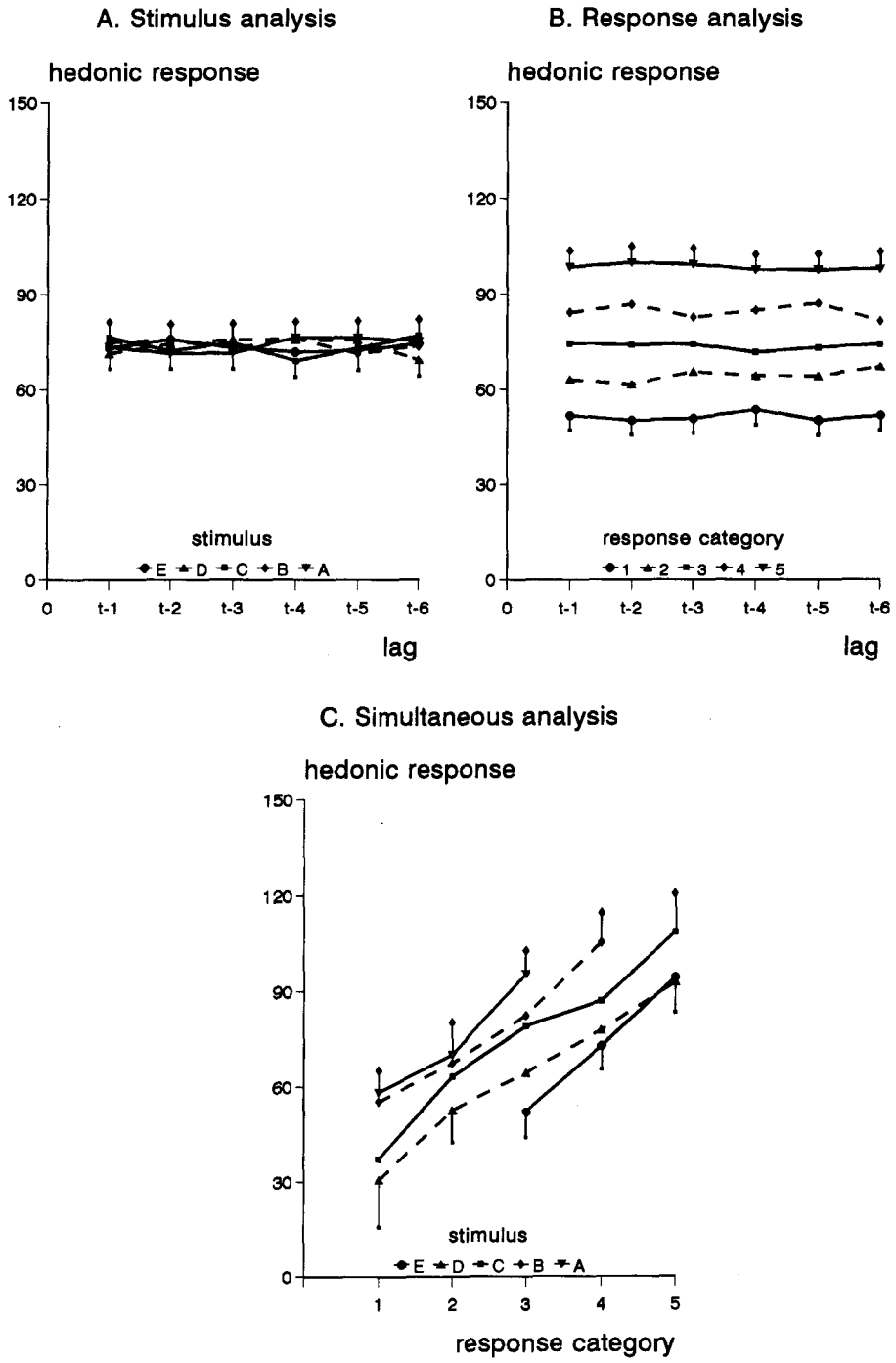
C. Simultaneous analysis

Figure 3. Sequential stimulus and response effects for hedonic judgments of the unidimensional
stimulus set containing sucrose solutions only (Experiment 2). See Figure 2 for further explanation.

e.g., Holland & Lockhead, 1968). These calculations do
not show systematic sequence effects for Experiment 1
(Figure 2A) or Experiment 2 (Figure 3A).

Response Analysis

The mean response on trial t was calculated as a func-
tion of previous responses for trial $t - 1$ up to $t - 6$ (see,

e.g., Holland & Lockhead, 1968). To reduce the number
of possible responses, the responses were divided into
five response categories. Each category represented one
fifth of the response scale. Responses on the left part of
the scale (<30 mm) received Category 1, responses on
the second part of the scale (>29 mm and <60 mm) were
classified in Category 2, and so on.

The curves in Figure 2B (Experiment 1) are somewhat separated; the curves for high previous responses seem to be somewhat higher than those for low previous responses. In Experiment 2, this effect is more pronounced. Figure 3B shows that a high response on a preceding trial coincides with a high response on the present trial, and that this effect hardly decreases with increasing time lag.

The Simultaneous Effects of Stimulus and Response

The mean responses at trial t were calculated as a function of the stimulus (S_{t-1}) and the response category (R_{t-1}) at trial $t - 1$. A simultaneous analysis aims to disentangle the effects of previous stimuli from those of previous responses (Petzold, 1981). The outcomes show a positive dependency between consecutive responses and a negative dependency between the present response and the previous stimulus (Figures 2C and 3C): The current response increases with increasing levels of the previous response (the slope of the curves), and it decreases with increased preference for the previous stimulus (the separation between the curves). Since the curves are steeper in Figure 3C than in Figure 2C, these analyses suggest that the response effects are larger for the unidimensional stimulus set than for the multidimensional set.

Differences Between Subjects

Figure 3B suggests that a response given six or more trials back exerts a substantial effect on the momentary response selection process at trial t. However, the pattern observed can also be accounted for by idiosyncratic differences in response selection. If some subjects use higher ratings than others, a high response on trial $t - 1$ will coincide with a high response on trial t when the data are aggregated over individuals (Haubensak, 1992). If so, the mean individual responses in Experiment 2 should be more variable than those in Experiment 1. Indeed, the mean individual responses in Experiment 1 ranged from 27.4 to 81.1 $(n = 30)$, whereas they varied from 19.5 to 131.6 $(n = 32)$ in Experiment 2.

To correct for idiosyncratic differences in scale usage, the individual response distributions can be brought back to a mean of 0 and a standard deviation of 1 by calculating

$$R_{ijt-st} = (R_{ijt} - \bar{R}_j)/\sigma_j, \tag{1}$$

where R_{ijt-st} is the momentary standardized response, R_{ijt} is the momentary response, and $\bar{R}_j$ and σ_j are the overall mean and the standard deviation of the responses of a particular subject j. To perform the response analysis and the simultaneous analysis, boundaries for the response categories of the standardized responses need to be determined. These boundaries can be obtained from the standard normal distribution: 20% of the standardized responses should be < -0.841, 40% < -0.255, 60% < 0.255, and 80% < 0.841. Figure 4 shows the results for the stimulus analysis, the response analysis, and the simultaneous analysis after standardization of the re-

sponses in Experiment 2. Most of the "sequence effects" evident in Figure 3 have now disappeared.

The Adoption Stage

For a meaningful analysis of sequential dependencies, we first have to make sure that the same processes operate throughout the entire experimental session. Therefore, we first inspected the ratings for individual stimuli as a function of trial number, after standardization (Equation 1). The average number of observations for each mean equals the number of subjects divided by five (the number of different stimuli). Due to a lack of observations, no mean could be calculated for the least appreciated stimulus (A) on Trial 8 in Experiment 2.
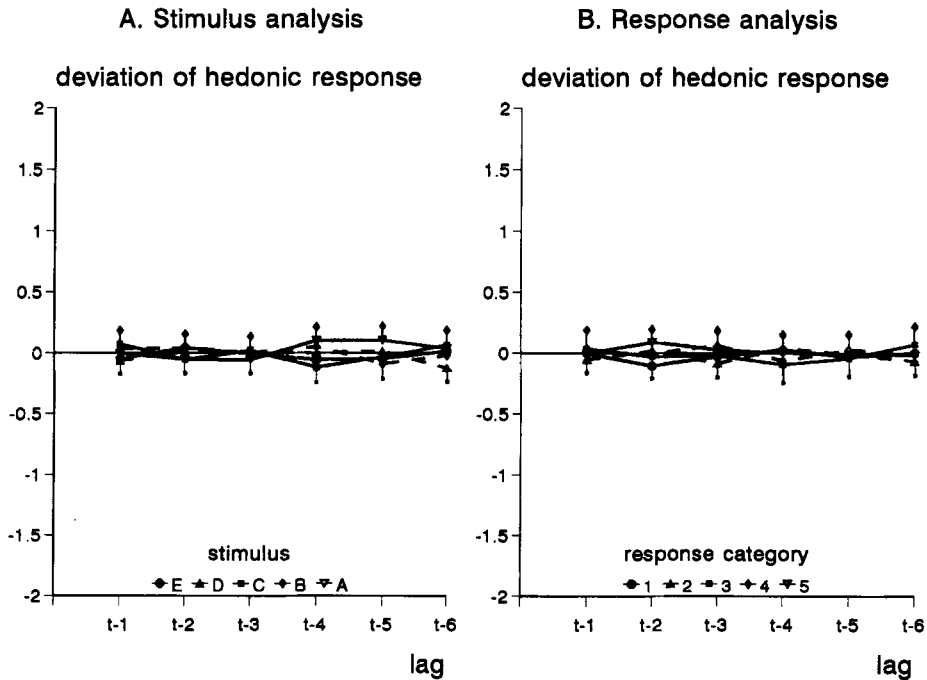
If naive subjects develop a response strategy during the first trials of a session, the initial responses are expected to deviate from the remainder of the experiment. We tested whether the 20 initial observations can be considered outliers (see Godfrey, 1988). For each stimulus S_i, we performed a regression of the response R_i on a constant and 20 dummy variables. Each dummy variable equaled 1 at trial t and 0 otherwise $(t \leq 20)$. The instances in which the two-tailed t statistics for the coefficients attached to the dummy variables were significant $(p < .025)$ are indicated in Figure 5. These observations are considered outliers.

Figure 5 shows that the subjects' response behavior changes gradually during the first trials: The number of outliers decreases with trial number. The effect is more pronounced in Experiment 2 than in Experiment 1. These results suggest that a sensory evaluation session can be considered a two-stage event. In the first stage, the subjects probably use the way in which they perceive the response scale and the perceived taste sensations to construct a response strategy. In this stage, response variability decreases. During the second stage, the mean ratings for the individual stimuli remain constant. To analyze the trial-to-trial dependencies in this stationary stage, the observations obtained during the adoption stage have to be deleted from the sample.

Data Cleaning and Model Selection

In determining the duration of the adoption stage, we used the rule that the mean standardized responses to one or more stimuli had to deviate significantly on at least two consecutive trials $(p < .025)$. In this way, we fixed the length of the adoption stage at 3 trials for Experiment 1 and at 12 trials for Experiment 2. These observations were deleted before the sequential effects were calculated: $(50 - 3) \times 30 = 1,410$ and $(50 - 12) \times 32 = 1,216$ responses were left in Experiments 1 and 2, respectively. We standardized these data again using Equation 1 to delete any primacy effects. Note that after this standardization, the mean standardized responses still differ across stimuli and individuals.

To investigate the sequential dependencies during the stationary stage, we performed regression analyses. We utilized Schifferstein and Frijters's (1992a) idea that not

A. Stimulus analysis

deviation of hedonic response

B. Response analysis

deviation of hedonic response

C. Simultaneous analysis
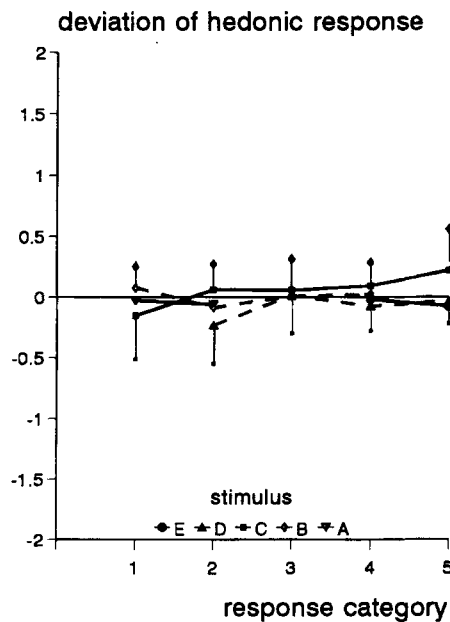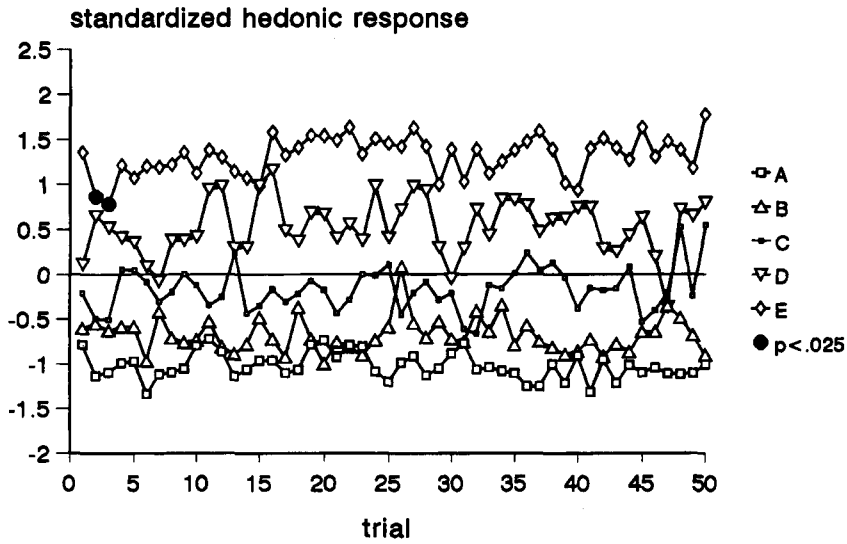
deviation of hedonic response

Figure 4. Sequential stimulus and response effects for the unidimensional stimulus set (Experiment 2), after standardization of the raw data. The standardized response on trial t, averaged over stimuli, is depicted as a function of the stimulus level (panel A) or the response level (panel B) on trial $t - k$, and as a function of both the preceding stimulus and response on trial $t - 1$ (panel C). The stimuli range from A (*least preferred*) to E (*most preferred*). The response categories correspond to $1 = R_{st} < -0.841$; $2 = -0.841 \leq R_{st} < -0.255$; $3 = -0.255 \leq R_{st} < 0.255$; $4 = 0.255 \leq R_{st} < 0.841$; $= R_{st} \geq 0.841$. Mean responses are depicted only if the number of observations exceeds 25. For the outermost stimuli, the sizes of 95% confidence intervals are indicated by error bars ($\pm 2 SE$).

A. Experiment 1

standardized hedonic response

B. Experiment 2

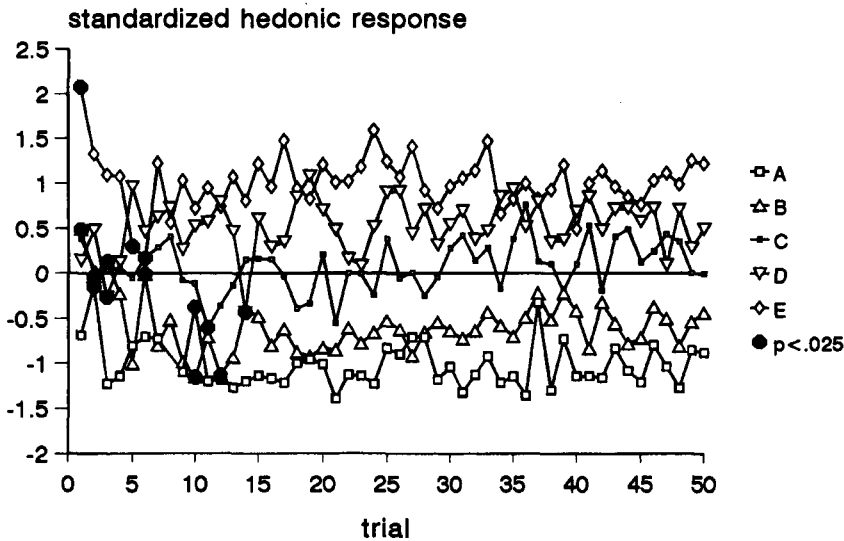### standardized hedonic response

Figure 5. Mean standardized hedonic responses for each stimulus, plotted as a function of trial number. Panel A shows the responses to the multidimensional stimulus set (Experiment 1), and panel B shows the responses to the unidimensional set of sucrose solutions (Experiment 2). The stimuli range from A (*least preferred*) to E (*most preferred*). Responses were averaged over subjects.

responses, but successive response deviations, are related. In addition to preceding response deviations, the model includes previous stimulus intensities, because DeCarlo and Cross (1990) found effects of previous stimuli separate from autocorrelated residuals. The model is given by

$$RD_t = \alpha_0 + \sum_{k=1}^{K} \left[\gamma_k RD_{t-k} + \beta_k I_{t-k} \right] + u_t, \qquad (2)$$

where RD_t denotes the momentary response deviation $(R_t - I_t)$ at trial t, k denotes the lag number, α, β, and γ are regression coefficients, and u_t is an error term.

Some authors have used log (S_t) as an approximation of the internal representation I_t. However, deviations from an assumed S–I function may result in outcomes that resemble a pattern expected for sequential dependencies (Gregson, 1976). Therefore, we made no assumptions about the individual S–I functions, and we
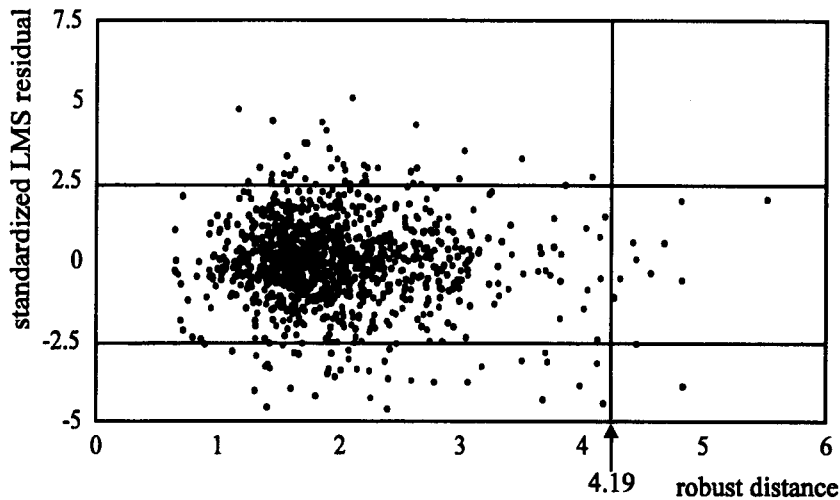
A. Experiment 1
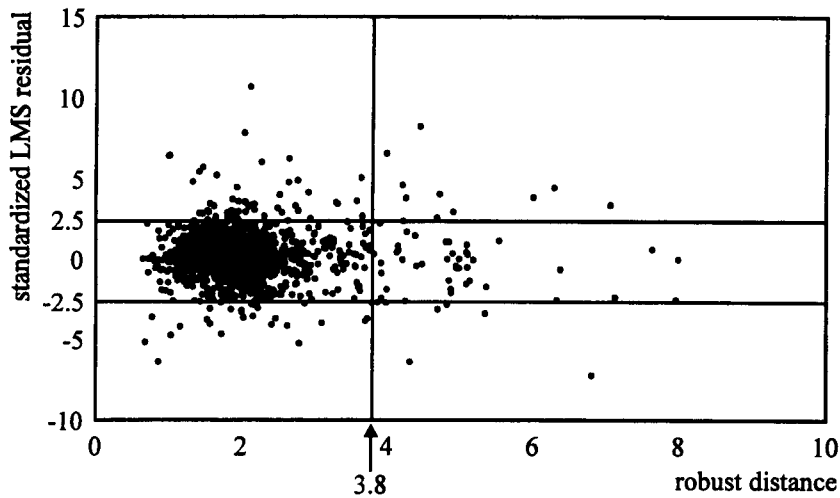
B. Experiment 2

Figure 6. Plots of robust residuals versus robust distances obtained by applying robust estimation methods to the data of Experiment 1 (panel A) and Experiment 2 (panel B), respectively. The lines show the tolerance limits defined by Rousseeuw and van Zomeren (1990). LMS, least median of squares.

used the mean response individual j gave to stimulus S_i as the estimate of I_{ij} ($\overline{R}_{ij}$).

The coefficients obtained for Equation 2 can be easily transformed into coefficients for a model employing only R_{t-k} and I_{t-k}:

$$R_t - I_t = \alpha_0 + \sum_{k=1}^{K}\left[\gamma_k R_{t-k} + \left(\beta_k - \gamma_k\right)I_{t-k}\right] + u_t. \quad (3)$$

The only additional restriction is that the coefficient for I_t must equal 1. If the standard deviations of the individual average responses per stimulus differ from the standard deviation of the average individual response, the coefficient for I_t could be significant when it enters

the right-hand side of Equation 3. However, this was not the case in any of the following analyses.

Time series analysis assumes at least weak stationarity (Maddala, 1992), implying that the mean response over time should be constant. A trend dominates all features of the time series and will affect the pattern of sequential dependencies. To correct for a possible trend, a trend term that equaled 1 on the first trial accounted for by the dependent variable, 2 on the second trial, and so on, was included in the regression model as an explanatory variable.

The Stationary Stage

Assuming that sequential dependencies are similar for all individuals, we calculated the ordinary least squares

Table 1
Regression Estimates (RE) With t Values of the Parameters in Equation 2 for the Ordinary Least Squares (OLS) and the Generalized M (GM) Estimation Procedures

| | Experiment 1 ($N = 1{,}320$) | | | | Experiment 2 ($N = 1{,}088$) | | | |
| | OLS | | GM | | OLS | | GM | |
	RE	t	RE	t	RE	t	RE	t
Intercept	0.006	0.577	−0.001	0.089	−0.059	−1.599	−0.070	−2.067
Trend	—	—	—	—	0.0031	1.688	0.0042	2.546
RD_{t-1}	0.099	3.686	0.111	3.593	0.071	2.327	0.053	1.614
I_{t-1}	−0.054	−4.466	−0.041	−4.451	−0.148	−6.541	−0.125	−5.922
RD_{t-2}	0.010	0.368	0.079	2.814	0.030	0.991	0.051	1.639
I_{t-2}	−0.018	−1.506	−0.024	−2.590	−0.052	−2.232	−0.013	−0.606
RD_{t-3}	0.084	3.150	0.060	1.986	0.055	1.850	0.039	1.324
I_{t-3}	−0.033	−2.685	−0.017	−1.910	−0.066	−2.853	−0.051	−2.433
RD_{t-4}	—	—	—	—	0.061	2.081	0.022	0.739
I_{t-4}	—	—	—	—	−0.035	−1.509	−0.012	−0.611
R^2	0.041		0.033		0.076		0.068	
F	9.2				9.9			
SE	0.395		0.263		0.592		0.464	
DW	2.071		2.078		2.016		2.020	
χ_2^2	1831				70			

Note—N, effective sample size; F, tests for significance of R^2; SE, standard error of the regression; DW, Durbin-Watson statistic; and χ_2^2, Bera-Jarque statistic.

(OLS) estimates for Equation 2 for $k = 5$. As the estimates of γ_5 and β_5 were both individually and jointly insignificant, we reduced k by 1 and estimated Equation 2 again. We repeated this procedure until at least one of the coefficients for the largest time lag became significant. The estimation results are presented in Table 1. Parameter estimates differ significantly from zero when the t values are larger than 1.96 (two-tailed t test, $p < .05$). The estimates for α_0 are expected to approach zero because the explanatory variables all have zero means.

The R^2 values are significant, suggesting the existence of sequential effects in both experiments. The values of the Durbin-Watson statistic lie around 2, as they ought to for dynamic models that are adequately specified (Durbin & Watson, 1950, 1951). Nevertheless, the Bera–Jarque statistics reject the null hypothesis, indicating that the u_ts do not follow a multivariate normal distribution (Jarque & Bera, 1980). This result is caused by too many large residuals. Since lagged values of RD are used as explanatory variables, outliers affect both the dependent and independent variables of Equation 2.

To identify influential observations, we made a scatter diagram of the robust residuals versus robust distances along the lines of Rousseeuw and Van Zomeren (1990) using the computer program RIPE (Lucas, 1993; Figure 6). All observations outside the tolerance limits are outliers. The observations in the upper right and lower right segments of each diagram make a relatively large contribution to the values of the estimated regression coefficients. These "bad leverage points" lead to regression results that suggest a relationship between the dependent variable and the independent variables, when such a relationship does not exist. Twenty of the 146 outliers in Experiment 1 and 1 of the 81 outliers in Experiment 2 were bad leverage points.

To preserve our estimation results from influential observations, we estimated the two regression equations again by a robust estimation procedure, employing the generalized M (GM) estimator (see Lucas, 1996). Roughly stated, the GM estimator weights the non-deterministic independent variables and the estimated residuals in an iterative procedure to obtain estimates not exposed to outliers. The weighting procedure transforms the outliers so that they become part of the majority of the data. All calculations were performed in the computer program RIPE (Lucas, 1993; Table 1).

If we compare the results for Experiment 1, the robust method yields estimates that differ considerably from the OLS estimates. The robust estimates and their t values decline monotonically with increasing lag of the regressor. Sequential dependencies are significant up to RD_{t-3} and I_{t-2}. The robust estimation method also led to different results in Experiment 2. The most important difference with the OLS estimates is that the robust estimates no longer suggest that subjects are primarily responsive to the response deviation four trials ago. The coefficients for the lagged response deviations are positive and decrease as a function of lag. However, none of the individual coefficients for the lagged RDs is significant. The coefficients for the preceding stimulus intensities remain difficult to interpret, since the coefficient for I_{t-3} is larger than the one for I_{t-2}. Experiment 2 yields a positive trend, implying that the average RD value increases during the session. The trend causes the coefficient for the intercept to be significant. To summarize, the robust estimation procedure yields parameter estimates that are more in line with a declining influence of preceding events as a function of lag than does the OLS procedure.

Second-Order Dependencies

It should be noted that regression analyses may not be appropriate because the variables in the regression equation may be dependent. Many authors (e.g., Baird, Green, & Luce, 1980; DeCarlo & Cross, 1990; Jesteadt et al.,
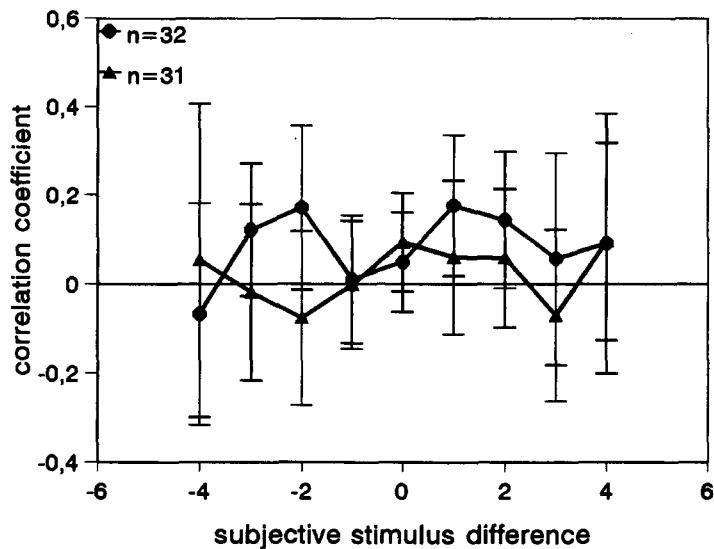
A. Experiment 1
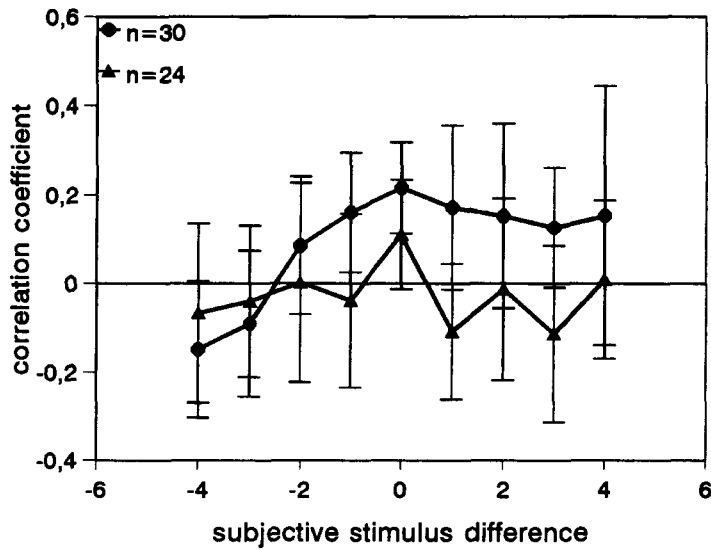
B. Experiment 2

Figure 7. Second order dependencies for Experiment 1 (panel A) and Experiment 2 (panel B). Correlation coefficients between RD_t and RD_{t-1} ($\pm 2\,SE$) are plotted as a function of subjective stimulus differences ($I_t - I_{t-1}$). The subjective stimulus differences were grouped in nine categories: $I_t - I_{t-1} < -2.1$; $-2.1 \le I_t - I_{t-1} < -1.5$; $-1.5 \le I_t - I_{t-1} < -0.9$, and so on. The curves with the reduced number of subjects (n) show the outcomes after subjects who produced bad leverage points were deleted from the sample.

1977) have found second-order dependencies in their data: The size of the correlation between consecutive responses depended on the subjective difference between stimuli. For two identical stimuli, the correlation coefficient was largest, and it decreased with increasing stimulus differences. These findings formed an inverted V pattern when the size of the correlation between consec-

utive responses was plotted as a function of the stimulus difference. Therefore, we calculated the correlation between successive response deviations (RD_t, RD_{t-1}) as a function of the difference in internal representation ($I_t - I_{t-1}$). Neither of the two experiments showed the characteristic inverted V (see the curves drawn for the largest numbers of subjects in Figure 7). However, sev-

eral correlations were significantly larger than zero, especially in Experiment 1.

To make sure that the results of the robust regression analyses were not affected by second-order dependencies, the correlation coefficients were calculated again after deleting the data from subjects who produced one or more bad leverage points in Figure 6. For Experiment 1, data from 6 subjects were deleted, and for Experiment 2, data from 1 subject were deleted. No inverted V pattern was found. In addition, none of the correlations deviated significantly from zero (Figure 7). Consequently, we conclude that the robust regression analyses are appropriate for analyzing these data.

GENERAL DISCUSSION

The analyses reveal two different judgmental strategies during a 1-h experimental session. During the first trials of a session, subjects adopt a certain way of responding. After a response mapping strategy has been developed, the average responses to a stimulus remain fairly stable. The momentary response deviations during the second part of the session are related to preceding response deviations and the preceding stimuli.

The Adoption Stage

During the initial trials of a session, subjects can obtain all the necessary information from the response scale, but they have only limited access to information from the stimulus set. The information gathered during the beginning of a session is used to develop a way in which perceived sensations are mapped onto responses. These first trials largely determine the responses during the entire session (Haubensak, 1990). To optimize consistency over trials, many subjects adopt a conservative attitude and use the middle response category on the first trial (Poulton, 1979).

In experiments in which stimulus context was manipulated, the length of the adoption stage extended over fewer than 13 trials (Morris & Rule, 1988; Parducci & Wedell, 1986; Vollmecke, 1987). In the present study, estimates for the length of this stage were 3 (Experiment 1) and 12 (Experiment 2) trials. These estimates were obtained from the decreases in response variability with trial number. The length of this stage presumably depends on the subjects' initial expectations concerning the relationship between stimulus and response. In the present study, the heterogeneous stimulus set probably matched the subjects' ideas of *not pleasant at all* and *extremely pleasant* better than the set of sucrose solutions. The subjects in Experiment 2, therefore, needed more trials to redefine the meaning of the scale descriptors in the context of the perceived sensations. The estimates for the length of the adoption phase may be relatively large for experiments on contextual effects, since these studies typically use stimulus sets that do not correspond to subjects' initial expectations with respect to perceived dynamic range and stimulus frequency.

Most studies reporting sequential dependencies have employed a training procedure or a practice session (e.g., Baird, Berglund, & Olsson, 1996; Bruvold, 1970; Green, Luce, & Duncan, 1977; Mori & Ward, 1990; Ward, 1985; Ward & Lockhead, 1970) or several practice trials within the experimental session (e.g., DeCarlo, 1992, 1994; DeCarlo & Cross, 1990; Jesteadt et al., 1977; Riskey, Parducci, & Beauchamp, 1979; Treisman & Williams, 1984). Practice trials are probably most effective in eliminating the effects of the adoption stage from experimental data. Subjects may start any session with their initial response behavior and use their previously developed response strategy only after they have made sure the stimulus context is similar to the one presented at the preceding session (see Stang, 1975). It should be noted that some studies have reported analyses of sequential effects without eliminating the results of the first trials, and that these may have contaminated their results (e.g., Gregson, 1983; Schifferstein & Frijters, 1992a; Stang, 1975; Ward, 1979).

The notion of an adoption stage is particularly relevant for taste research, since several investigators in the taste realm have argued that subjects cannot evaluate more than 10 samples per session (see Sauvageot, 1990). Means and standard deviations of the responses obtained in such short sessions are likely to depend on trial number to a considerable extent.

Gradual Changes in the Stationary Stage

In Experiment 2, a significant trend was found in the regression analyses (Table 1). However, the impact of this trend on the stationary stage is small and probably has no practical relevance. The absence of trends in the stationary stage may be expected for intensity judgments, but is less logical for hedonic judgments. Mere exposure to an initially novel stimulus increases familiarity and liking (e.g., Becknell, Wilson, & Baird, 1963; Zajonc, 1968, 1980), whereas the pleasantness of a well-known food decreases after subjects eat the food until satisfied (e.g., Pliner, Polivy, Herman, & Zakalusny, 1980; Rolls, Rolls, Rowe, & Sweeney, 1981).

The tastes perceived in the present study were probably not extremely novel, although these particular stimuli may have been. Therefore, a decrease in pleasantness ratings with trial number seems feasible. In studies on sensory-specific satiety (Rolls et al., 1981) researchers have typically examined the decrease in pleasantness after instructing subjects to eat certain foods until they were satisfied. In the present study, however, the stimuli were not ingested. The present findings are in accordance with Cabanac (1971), who reported that the pleasantness of sweet solutions decreases for subjects who swallow sucrose solutions but remains unchanged for subjects who expectorate them. Cabanac accounted for these findings by a mechanism based on the subjects' physiological need. However, a cognitive mechanism is equally plausible: When subjects judge the pleasantness of a stimulus without swallowing the sample, they may re-

gard "pleasantness" an abstract *stimulus* attribute independent of their own physiological condition.

Sequential Dependencies in the Stationary Stage

Contrary to most of the other continua investigated, consecutive judgments for taste stimuli show a negative dependency between the current response and the preceding stimulus, even after correction for autocorrelation in the error term (see DeCarlo, 1992; DeCarlo & Cross, 1990; Schifferstein & Frijters, 1992a). This discrepancy between modalities may originate from differences in intertrial time intervals. DeCarlo (1992) has shown that the positive dependency between R_t and I_{t-1} disappears when the interstimulus interval is increased from 2 or 6 sec to 15 or 20 sec. In taste psychophysics, where interstimulus intervals of 30–60 sec are frequently used, the positive dependency is likely to be absent or may turn negative.

A negative dependency is also found when a test stimulus is tasted after a series of stimuli low or high in target intensity. The size of this dependency increases approximately linearly with the number of repetitions of the preceding stimulus (Kroeze, 1983; Schifferstein & Frijters, 1992b; Schifferstein & Oudejans, 1996). However, this type of dependency may result from another mechanism than the one observed in the present study. The series of extreme (high- or low-intensity) stimuli may encourage the subjects to adjust their judgmental reference frames. In that case, subjects are continually adopting new response strategies and have not reached the stationary stage.

Second-Order Dependencies

The present data do not show an inverted V pattern when the correlation coefficient between R_t and R_{t-1} is plotted as a function of the subjective stimulus difference. This finding contrasts with findings from most empirical studies investigating intensity responses. The absence of the inverted V could be due to the weakness of the intertrial dependencies. When the present analyses were performed on Schifferstein and Frijters's (1992a) sweetness intensity data, the R^2 value was 0.12, which is higher than the values reported in Table 1 for hedonic judgments.

A second possibility is that identification underlies second-order dependencies. The awareness that two consecutive stimuli are identical could induce a response heuristic: Identical stimuli get the same response. The correlation decreases with increasing stimulus differences because the probability of being identified as "same" decreases with an increasing difference. Two stimuli that are equally pleasant are not necessarily the same. If this line of reasoning is correct, therefore, pleasantness ratings will not exhibit the inverted V pattern. However, such a mechanism is unlikely, given that Ward (1982) found inverted V patterns for consecutive responses when subjects judged the intensities of alternating sounds and lights.

REFERENCES

BACON, M. M., ROOD, E. A., & WASHBURN, M. F. (1914). A study of affective contrast. *American Journal of Psychology*, 25, 290-293.

BAIRD, J. C., BERGLUND, B., & OLSSON, M. J. (1996). Magnitude estimation of perceived odor intensity: Empirical and theoretical properties. *Journal of Experimental Psychology: Human Perception & Performance*, 22, 244-255.

BAIRD, J. C., GREEN, D. M., & LUCE, R. D. (1980). Variability and sequential effects in cross-modality matching of area and loudness. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 277-289.

BECKNELL, J. C., WILSON, W. R., & BAIRD, J. C. (1963). The effect of frequency of presentation on the choice of nonsense syllables. *Journal of Psychology*, 56, 165-170.

BEEBE-CENTER, J. G. (1929). The law of affective equilibrium. *American Journal of Psychology*, 41, 54-69.

BERLYNE, D. E. (1973). The vicissitudes of aplopathematic and thelematoscopic pneumatology (or the hydrography of hedonism). In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, preference: Their nature, determinants, and role in behavior* (pp. 1-33). New York: Academic Press.

BRUVOLD, W. H. (1970). Rated acceptability of mineral taste in water: III. Contrast and position effects in quality scale ratings. *Journal of Experimental Psychology*, 85, 258-263.

CABANAC, M. (1971). Physiological role of pleasure. *Science*, 173, 1103-1107.

DECARLO, L. T. (1992). Intertrial interval and sequential dependencies in magnitude scaling. *Journal of Experimental Psychology: Human Perception & Performance*, 18, 1080-1088.

DECARLO, L. T. (1994). A dynamic theory of proportional judgment: Context and judgment of length, heaviness, and roughness. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 372-381.

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

DURBIN, J., & WATSON, G. S. (1950). Testing for serial correlation in least squares regression, I. *Biometrika*, 37, 409-427.

DURBIN, J., & WATSON, G. S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika*, 38, 159-178.

GODFREY, L. G. (1988). *Misspecification tests in econometrics*. Cambridge: Cambridge University Press.

GREEN, D. M., LUCE, R. D., & DUNCAN, J. E. (1977). Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception & Psychophysics*, 22, 450-456.

GREGSON, R. A. M. (1976). Psychophysical discontinuity and pseudo-sequence effects. *Acta Psychologica*, 40, 431-451.

GREGSON, R. A. M. (1983). The sequential structure of odour mixture component intensity judgements. *British Journal of Mathematical & Statistical Psychology*, 36, 132-144.

HARRIS, A. J. (1929). An experiment on affective contrast. *American Journal of Psychology*, 41, 617-624.

HAUBENSAK, G. (1990). Primacy effects in absolute judgments. In H. G. Geissler (Ed.), *Psychophysical explorations of mental structures* (pp. 104-113). Toronto: Hogrefe & Huber.

HAUBENSAK, G. (1992). Sequenzeffekte in absoluten Urteilen? Kritisches zur Methode [Sequence effects in absolute judgments? A critique on the method]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 39, 101-113.

HELSON, H. (1973). A common model for affectivity and perception: An adaptation-level approach. In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, preference: Their nature, determinants, and role in behavior* (pp. 167-188). New York: Academic Press.

HOLLAND, M. K., & LOCKHEAD, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception & Psychophysics*, 3, 409-414.

JARQUE, C. M., & BERA, A. K. (1980). Efficient tests for normality, heteroscedasticity and serial dependence of regression residuals. *Economic Letters*, 6, 255-259.

JESTEADT, W., LUCE, R. D., & GREEN, D. M. (1977). Sequential effects

in judgments of loudness. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 92-104.

KROEZE, J. H. A. (1983). Successive contrast cannot explain suppression release after repetitive exposure to one of the components of a taste mixture. *Chemical Senses*, **8**, 211-223.

LUCAS, A. (1993). *RIPE—Robust inference plus estimation* [Computer program]. Rotterdam: Erasmus University, Tinbergen Institute.

LUCAS, A. (1996). *Outlier robust unit root analysis.* Doctoral dissertation, Erasmus University, Tinbergen Institute.

MADDALA, G. S. (1992). *Introduction to econometrics* (2nd ed.). New York: Macmillan.

MORI, S., & WARD, L. M. (1990). Unmasking the magnitude estimation response. *Canadian Journal of Psychology*, **44**, 58-68.

MORRIS, R. B., & RULE, S. J. (1988). Sequential judgement effects in magnitude estimation. *Canadian Journal of Psychology*, **42**, 69-77.

PANGBORN, R. M. (1980). A critical analysis of sensory responses to sweetness. In P. Koivistoinen & L. Hyvönen (Eds.), *Carbohydrate sweeteners in food and nutrition* (pp. 87-110). London: Academic Press.

PARDUCCI, A., & WEDELL, D. H. (1986). The category effect with rating scales: Number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 496-516.

PETZOLD, P. (1981). Distance effects on sequential dependencies on categorial judgments. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1371-1385.

PLINER, P., POLIVY, J., HERMAN, C. P., & ZAKALUSNY, I. (1980). Short-term intake of overweight individuals and normal weight dieters and non-dieters with and without choice among a variety of foods. *Appetite*, **1**, 203-213.

POULTON, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin*, **86**, 777-803.

RISKEY, D. R., PARDUCCI, A., & BEAUCHAMP, G. K. (1979). Effects of context in judgments of sweetness and pleasantness. *Perception & Psychophysics*, **26**, 171-176.

ROLLS, B. J., ROLLS, E. T., ROWE, E. A., & SWEENEY, K. (1981). Sensory specific satiety in man. *Physiology & Behavior*, **27**, 137-142.

ROUSSEEUW, P. J., & VAN ZOMEREN, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, **85**, 633-639.

SAUVAGEOT, F. (1990). Fatigue in sensory evaluation: Myth or reality? In G. Charalambous (Ed.), *Flavors and off-flavors* (pp. 577-594). Amsterdam: Elsevier.

SCHIFFERSTEIN, H. N. J. (1995). Contextual shifts in hedonic judgments. *Journal of Sensory Studies*, **10**, 381-392.

SCHIFFERSTEIN, H. N. J., & FRIJTERS, J. E. R. (1992a). Contextual and sequential effects on judgments of sweetness intensity. *Perception & Psychophysics*, **52**, 243-255.

SCHIFFERSTEIN, H. N. J., & FRIJTERS, J. E. R. (1992b). Sweetness does not habituate during a sip-and-spit experiment. *Physiology & Behavior*, **51**, 331-336.

SCHIFFERSTEIN, H. N. J., & OUDEJANS, I. M. (1996). Determinants of cumulative successive contrast in saltiness intensity judgments. *Perception & Psychophysics*, **58**, 713-724.

STANG, D. J. (1975). When familiarity breeds contempt, absence makes the heart grow fonder: Effects of exposure and delay on taste pleasantness ratings. *Bulletin of the Psychonomic Society*, **6**, 273-275.

TREISMAN, M., & WILLIAMS, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, **91**, 68-111.

VOLLMECKE, T. A. (1987). *The influence of context on sweetness and pleasantness evaluations of beverages.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

WARD, L. M. (1979). Stimulus information and sequential dependencies in magnitude estimation and cross-modality matching. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 444-459.

WARD, L. M. (1982). Mixed-modality psychophysical scaling: Sequential dependencies and other properties. *Perception & Psychophysics*, **31**, 53-62.

WARD, L. M. (1985). Mixed-modality psychophysical scaling: Inter- and intramodality sequential dependencies as a function of lag. *Perception & Psychophysics*, **38**, 512-522.

WARD, L. M., & LOCKHEAD, G. R. (1970). Sequential effects and memory in category judgments. *Journal of Experimental Psychology*, **84**, 27-34.

ZAJONC, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology Monographs*, **9** (2, Pt. 2), 1-28.

ZAJONC, R. B. (1980). Feeling and thinking. Preferences need no inferences. *American Psychologist*, **35**, 151-175.