Available online at www.sciencedirect.com
ScienceDirect
Appetite 49 (2007) 18–29
Appetite
www.elsevier.com/locate/appet

Research Report

Experienced and remembered pleasure for meals: Duration neglect but minimal peak, end (recency) or primacy effects

Elizabeth Rode[a], Paul Rozin[a,*], Paula Durlach[b]

[a]Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104-6241, USA
[b]US Army Research Institute for the Behavioral and Social Sciences, USA

Received 27 April 2004; received in revised form 17 November 2005; accepted 4 September 2006

Abstract

Rated liking for dishes consumed during a meal was compared with recalled liking in two studies using actual meals and one with an imagined meal. The effects on memory of the most pleasant dish, the first and last dishes, a rising vs. falling hedonic profile, and the time spent eating a dish were evaluated for similarity to effects seen in memories of pain. Across the three studies, there was consistent evidence for duration neglect (no effect of increased duration/exposure of the favorite component), and some weak evidence that patterns rising in liking are preferred to those falling in liking. In all three studies, there was no evidence for peak, primacy or recency effects. The existence of duration neglect implies that, with respect to memories of a meal, small portions of a highly favored dish will have roughly the same memorial effect as large portions.

Keywords: Liking for foods; Memory of liking; Portion size memory; Meal; Pleasure

Introduction

Pleasures and pains are mental events that exist in three temporal domains: in anticipation, in the experience itself, and in memory. This fundamental distinction has been introduced to the study of the psychology of affect by Daniel Kahneman and his associates (Kahneman, Wakker, & Sarin, 1997; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). This formulation, and in particular, the distinction between experienced and remembered pleasure, is central to the study of affect for three reasons.

1. There are important distortions in the relationship between anticipated, experienced, and remembered pleasure. In particular, memories are often markedly discrepant from the actual experiences they represent (Kahneman et al., 1993, 1997).
2. In many instances, the amount of time spent remembering an affective experience (a valenced event) is much longer than the experience itself; this is surely the case for brief experiences of pleasure or pain.
3. At the critical time of a choice, individuals usually are choosing between the memories for the experiences they have had with respect to the two or more choices. Insofar as those memories are discrepant from the actual experiences, the representation in memory may be the more important predictor.

Based on work in the domain of pain—that is, negative affect—Kahneman and colleagues (Fredrickson, 2000; Fredrickson & Kahneman, 1993; Kahneman et al., 1993) suggest that the relation between experienced and remembered affect is complex. In particular, they report that the peak and end of the experience are disproportionately important in the memory representation, and that the duration of an experienced episode has a minimal influence on the affective memory for it—a phenomenon they call "duration neglect" (Kahneman et al., 1993, 1997).

Dan Ariely and his colleagues (Ariely & Carmon, 2000; Ariely & Levav, 2000; Ariely & Zauberman, 2000) have carried out systematic studies on remembered pain, in the context developed by Kahneman and Fredrickson. They employ actual pain procedures in the laboratory, or real world pain experiences (people involved in bone marrow

*Corresponding author.
E-mail address: rozin@psych.upenn.edu (P. Rozin).

0195-6663/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.appet.2006.09.006

transplants). This work supports some of the principles developed by Kahneman and Fredrickson, and demonstrates the importance of some variables in modulating these effects. The work confirms that the end rating is important, along with the trend (rising or falling). Peak and average were less successful predictors of remembered pain (Ariely & Carmon, 2000). The greater potency of the end is also in keeping with the finding of Loewenstein and Prelec (1993) that people generally prefer hedonically rising profiles to falling profiles with the same average level.

Ariely and his colleagues (Ariely & Levav, 2000; Ariely & Zauberman, 2000) have also demonstrated duration neglect, and shown that its appearance depends framing the episode in question as a single experience. In one of the few studies addressing these issues for pleasant experiences, participants continuously rated emotional intensity of musical selections. Remembered pleasure was found to be substantially correlated with peak and end (or more generally, recency), and duration neglect was clearly demonstrated (Rozin, Rozin, & Goldberg, 2003).

In this study, we extend the conceptual and empirical analysis developed by Kahneman, Fredrickson, Ariely and their colleagues to food. There are three arguments for this extension:

1. It is important to determine whether the principles that have been derived from pain studies apply more generally, and in particular to positive affective experiences.
2. Food, meals, and eating comprise one of the most frequent and important experiences that people have.
3. As eating is structured in many societies, particularly contemporary Western/developed cultures, it naturally offers a sequence of experiences (dishes) that is highly appropriate for temporal analysis. As it were, the meal is made for this type of research.

The meal is the natural psychological/ecological unit of eating (Pliner & Rozin, 1999). It is very likely that a meal that is unique or relatively distinctive, and perhaps particularly good or bad, is experienced for more time in recall over subsequent days, months or even years, than in the actual experience. This is unlikely to be the case for more routine meals. In making decisions about purchasing foods in stores, or ordering meals in restaurants, we are certainly referring to our memories and not our actual experiences (to which we have no direct access, after the fact). Even in some of the standard procedures for rating food (and other) preferences in the laboratory, where the foods in question are present and sampled, the participant is comparing memories of albeit very recent experiences. In spite of these arguments, there is very little research on remembered pleasure for meals, and none informed by the advances made by Kahneman and colleagues.

Rogozenski and Moskowitz (1982) examined the relative importance of different meal components on overall meal evaluation for members of the US army. Participants provided like/dislike ratings of each of 144 foods, presented by written description (e.g., "macaroni and cheese"). With these results in hand, the investigators created a set of written descriptions of meals, which participants then rated on the same hedonic scale. The main finding from this work was that the entrée is the strongest predictor of the overall rating of a given meal.

Anderson and Norman (1964) examined the importance of order of dish in overall meal evaluation. Undergraduates rated their likings for 140 verbally presented foods. From these ratings, "meals" were constructed, personalized for each participant. The meals consisted of a sequence of six dishes, described simply by their names. They were constructed such that three highly rated foods were followed by three low-rated foods, or the opposite. The participants were *read* the menu of a meal (e.g. HHHLLL) and they rated how appealing this meal was. The participants did not eat a meal. The meals that began with three highly rated foods were rated more highly than meals composed of the same foods in reverse order (Anderson & Norman, 1964). In other words, in contrast to the "end" (recency) effect that has emerged from the pain (and music) studies, there was a primacy or "beginning" effect in this study. Of course, in studies of memory (for content, as opposed to affect), there is extensive evidence for both primacy and recency effects (e.g. Murdock & Bennet, 1962).

In the present studies, we build upon the prior work, on both remembered affect in general, and memory for meals. These are among the first studies applying the new Kahneman formulation in a positive, food context. Unlike the prior two studies on meal memory, which "created" hypothetical meals by presenting lists of foods, we rely, in two of our three cases, on memory for real eating experiences, rather than lists of foods. We examine memory for "meals" in two controlled conditions: sequences of different flavors of jelly beans, and sequences of foods that we arrange in the context of a real buffet restaurant. In a third study, we do use lists of foods, but unlike prior research, we had participants imagine consuming the meals bite by bite.

Duration neglect in our studies has a somewhat different meaning than the duration neglect based on continuous pain episodes. In the latter studies (e.g., Kahneman et al., 1993; Redelmeier & Kahneman, 1996), duration means extending the length of an episode; the episodes being compared (e.g., long and short) typically have the same average level of pain. In these studies, average level is not pitted against long vs. short durations. In our studies, of a structured meal, extending the duration of a *favorite* course not only increases the duration of the entire meal, but it raises the average hedonic level. So, in our studies, duration neglect means no effect of duration even when it raises the average hedonic level of the experience.

On the basis of the prior research, we entertain the following hypotheses or possibilities:

1. *Duration neglect*: The amount of a particular food eaten will make relatively little contribution to the overall

remembered hedonic rating of a meal. This is represented in these studies principally by comparison of single or double portions of a particular food/dish (usually the favorite item) within a meal.
2. The initial (beginning) dish will contribute disproportionately to the overall hedonic memory.
3. The final (end) dish will contribute disproportionately to the overall hedonic memory.
4. The most liked or disliked (''peak'') dish will contribute disproportionately to the overall rated liking for the meal.
5. A pattern of increasing (rising) liking across the dishes in a meal will yield higher remembered pleasure than an otherwise equivalent pattern of decreasing (falling) liking across the meal.

Study 1: jelly bean "meals"

We constructed ''mini-meals'' of six or eight jelly beans, with different flavors at the beginning, middle and end. The meal, consumed in a period of a few minutes, consisted of two jelly beans of flavor A, followed by two of flavor B, followed by two or four of flavor C. Individual flavors were rated for liking at the time of the ''meal,'' and then, an hour later, the overall jelly bean meal was rated.

Method

Participants

The participants were 173 undergraduate students in an introductory psychology class at the University of Pennsylvania.

Procedures

A pilot study was conducted on many different flavors of jelly beans informed by sales information from the manufacturer, ''Jelly Belly.'' Six flavors of jelly beans that represented a wide range of overall liking and that were easily distinguishable by taste and color were included, and were divided into two sets of three: Cream Soda, Licorice, and Cherry, or Buttered Popcorn, Blueberry, and Cantaloupe.

The study was carried out in class, on all participants simultaneously. Since we could not determine the individual hedonic functions for each student in advance, and then construct an appropriate jelly bean meal for each, we relied on the large number of participants to produce a wide range of hedonic profiles from the sequence of randomly chosen three flavors for each participant.

During class, participants received a large, opaque envelope, which included two answer sheets. Within each envelope were three smaller individually labeled envelopes of three different flavors of jelly beans. The first and second envelopes contained two jelly beans and the third contained two or four jelly beans. The envelopes were numbered one, two, and three. Within each envelope, the jelly beans were identical in color and flavor, and the flavor was different for each of the three bags. The order of consumption of the three types of jelly beans was designated by the colors of the jelly beans (also labeled on the envelope with the first letter of the relevant color) and was varied at random from one participant to another. For those participants who had a four-jelly-bean envelope, it was always consumed last. Approximately, half of the participants received one of the sets of three jelly beans and the remainder had the other set.

Instructions were given orally to the entire class step by step to control the pace of the study. The answer sheet indicated the order in which each participant was to eat the jelly beans. If it said, for example, Blue, Yellow, Orange, the participant would eat the two blue jelly beans in the B envelope first, then the two yellow beans in the Y envelope, and then the two or four orange beans in the O envelope. Participants were instructed to open the first envelope, and, in keeping with the spoken instructions, eat one of the jelly beans. Several seconds later they were then asked by the instructor to eat the second jelly bean of that flavor. After finishing these two jelly beans they were instructed to rate how much they liked or disliked the flavor of the jelly beans they just ate by a mark on one of the numbers 1–9 with each number labeled from ''dislike extremely'' (1) to ''like extremely'' (9), with the midpoint (5) labeled as ''neither like nor dislike'' (the widely used Hedonic Scale of Peryam and Pilgrim). Then they were instructed to open the next two envelopes and eat and rate the jelly beans in the same manner. The third (last) envelope contained four jelly beans for 67 of the 173 participants. If the participant had two jelly beans he/she ate them as instructed and then was asked to rate them. If the participant had four jelly beans, he/she followed the same directions for the first two jelly beans but was not asked to rate them. Rather, he/she was asked to continue to eat two more in the same rhythm as before, and then rate the flavor after the last jelly bean.

After an hour of lecture passed, students were asked to remove a second rating sheet from the envelope. This sheet had a 100-mm line on it, hashed at the midpoint. Participants were asked to write ''dislike extremely'' on the left end, ''neutral'' in the center and ''like extremely'' at the right end and then to indicate how much they liked the whole jelly bean ''meal'' that they had eaten about an hour ago. The lines on the final sheet were not initially labeled so as to reduce the chance of participants anticipating that the whole jelly bean meal would be rated. A rating line was used for the overall experience in order to prevent participants from mentally averaging the nine-point scale ratings of the three sets.

Results

Hedonic patterns

Depending on how much a participant liked each of the three randomly ordered jelly beans in his or her set, each participant experienced a particular hedonic pattern. Fortunately, as one might expect, the mean hedonic ratings

Table 1
Study 1: Mean ratings of individual types of jelly beans or total remembered "meal" for the various groups

Group[a]	*N*	Envelope 1	Envelope 2	Envelope 3	Average	Maximum	Minimum	Overall remembered[b]
All	173	5.9	5.4	5.2	5.5	7.5	3.1	17.8
Rise	26	3.4	6.0	7.6	5.7	7.6	3.4	14.5
Fall	30	7.4	5.8	2.5	5.2	7.4	2.5	16.6
Mid-peak	18	3.6	7.3	3.5	4.8	7.3	2.8	−5.5
Start peak	14	7.4	3.0	4.5	5.0	7.4	3.0	10.7
Mid-trough	33	7.0	2.4	6.7	5.3	7.5	2.4	21.5
End trough	16	7.1	7.8	2.2	5.7	8.0	2.2	18.9

[a]The items from rise to L3 are mutually exclusive categorizations.
[b]These ratings were on a line end-anchored on 'dislike extremely' (scored as −50) and 'like extremely' (+50), different from all the ratings during eating in this Table, which were on the hedonic scale of nine categories from extreme dislike (scored 1) to extreme liking (scored 9).

for the first (5.9), second (5.4) and third (5.2) bags were quite similar. For both groups (ending with two or four jelly beans), the patterns (Table 1) were classified according to the following criteria. If there was a monotonic increase in rating of the three bags, this was categorized as a rising pattern: there were 26 such cases (means for the three bags in sequence: 3.4, 6.0, 7.6). A monotonically falling pattern was shown in 30 cases (means: 7.4, 5.8, 2.5). The remaining cases were categorized by the location of the *most discrepant rating*, called a peak if it was positive with respect to the other two, and a trough if it was negative. Thus, a trial of 7, 3, 2 would be categorized as start peak; 4, 5, 8 would be rated as end peak; 1, 6, 6 as start trough; and 5, 1, 6 as mid-trough. By this criterion, although it happened infrequently, a rating such as 5–6–6 would be considered a start trough, while a rating of 6–8–7 would not be classified. There were too few exemplars of some of the troughs and peaks to permit analysis (we cut off at an *n* of less than 14) but there were 14 start peak trials, 18 mid peak, 33 mid-trough and 16 end trough (Table 1).

Duration neglect

We ask first whether duration neglect holds under the conditions of the study. We address this first because, if there is no difference between the two- and four-terminal jelly bean conditions (that is, if duration neglect holds), we can combine these two groups, making subsequent analyses both simpler and more powerful.

The four jelly bean participants ate twice as many terminal jelly beans as the two jelly bean participants, and did so in twice the time (since the onset of eating of each jelly bean was prompted by the instructor). The difference between these groups in overall liking ratings from memory therefore is a direct test of duration neglect. Since we used different scales to measure liking for the jelly beans as consumed and later memory for the experience, we cannot do direct comparisons of "on-line" ratings (e.g., averaged value of the immediate liking ratings for the three jelly beans) and remembered ratings.

There are three types of tests for duration neglect. First, in the absence of duration neglect, i.e. with an effect of longer exposure, the overall remembered flavor rating should correlate more highly with online ratings for the third bag with four beans than for the third bag with two beans. In fact, in support of duration neglect, the correlations of the ratings of the terminal 2 and 4 beans with remembered pleasures are the same, 0.39.

An alternative version of this idea is that, if we double-weight the last bean bag in the four jelly bean group in calculating the average experience, this weighted average should correlate more highly with the overall remembered rating than the unweighted average. There is a numerical difference (not significant) but it is *not* in the predicted direction: for the four jelly bean group, the correlation between the remembered-flavor rating and the average of the three bags is 0.74, while the correlation with the double-weighted last bag average is 0.70.

A third set of predictions concerns the absolute value of the remembered flavor's rating. If the last (third) envelope is the highest (lowest) rated envelope, then it should result in a higher (lower) overall rating for the four-bean group. This is because in last-high patterns (rising or end peak), the doubling of the beans in the last envelope increases the best part of the experience (from 1/3, 2 of 6 beans, to 1/2, 4 of 8 beans). The opposite holds, of course, for last envelope lowest (falling or end trough). Unfortunately, both the absolute value of the ratings for the third envelope and the average envelope ratings vary from participant to participant. We extracted all possible pairs of two- and four-bean participants with the same third envelope rating, the same average rating (within 0.2), and either both high end or both low-end patterns. There were 15 pairs that met this requirement; in eight of the 15 cases, the four-bean condition had a *smaller* effect on absolute value than the corresponding two-bean condition, again non-significant, but this is in the direction *opposite* to what would be predicted by an effect of increasing exposure.

Thus all three comparisons support the operation of duration neglect; there is no evidence of *any* greater influence of the doubled length of experience of the jelly beans.

Recency/end vs. primacy/beginning

A recency effect would be supported by a higher correlation with remembered pleasure for hedonic scores of the last jelly beans (envelope 3) than for the first (envelope 1) or second (envelope 2) set. The actual values (Table 2) place the first envelope highest ($r = 0.48$), followed by the third (0.39) with the second only slightly behind it (0.38). Although these correlations are not significantly different, the order of the correlation values implies a primacy effect, not a recency effect.

Also, if recency matters, then the rising pattern should show a higher overall remembered-flavor mean liking score than the falling pattern, assuming the two have the same average value. We again used the procedure of pairing, this time pairing rising and falling patterns having the same average (combining the four- and two-bean groups, since no difference has emerged between them). There were 14 cases with rising–falling pairings and matched averages. In eight of the 14 (not significant), the rising pattern had a higher remembered rating. However, when we created a difference score for each pair, by subtracting the overall liking rating of the rising member of the pair from the overall liking rating of falling member of the pair, mean difference was –8.6, in the direction of a higher rating for the falling set; however, this was not significant by single value *t*-test. Hence, there was no evidence here for a recency/end effect and only very weak evidence in favor of primacy.

Peak/trough

For both groups combined (four or two terminal beans), the value of the most liked jelly bean (peak value, labeled as MAX in Table 2) has the second highest correlation with remembered liking, after the average of all three jelly beans (Table 2). This is consistent with peak being, if anything, superior to position (first, middle or last bean) in predicting remembered liking of the meal. We can also consider cases in which the peak was most distinct (the 37 three-bean sequences classified as having a peak), since if there is a peak effect, it should be clearer with distinct peaks. The correlation between the rating of the peak bean and overall remembered rating of the bean meal for just those 37 cases was 0.63, essentially the same as the value 0.62 for the maximum value (even if it was not distinct enough to qualify as a peak) for all participants.

The same logic holds for the trough. For both groups, the trough value (MIN) is the third best predictor (0.58), after average and MAX (Table 2). This is consistent with MIN being, if anything, superior to position. We can also consider cases in which the low was most distinct (trough classifications, $n = 57$), assuming again that a more distinct trough might have a bigger effect. The correlation between the value of the trough bean and the remembered bean meal liking is 0.54, actually slightly lower than the 0.58 correlation for MIN (lowest bean, whether or not it is a trough) for all participants.

Finally, we can examine pairs in which one member has a peak and the other a trough, and both have the same average hedonic ratings. If peaks have a disproportionately positive effect and troughs a disproportionately negative effect, the remembered flavor liking scores for peak sequences should be more favorable than those for trough sequences. The data do not support this claim at all. In 15 of the 24 pairs (n.s.), the trough had a higher remembered liking than the peak, with the mean of the remembered meal liking for the 24 troughs actually 9.0 mm higher (on the 100 mm from dislike extremely to like extremely) than the mean for the corresponding peaks (paired *t*, n.s.).

There is, therefore, at best mixed evidence supporting the hypothesis that the peak or trough experience has a disproportionate effect in predicting remembered liking.

Study 2: imagined meals

In Study 2, participants imagine eating a specified meal but, instead of the common procedure of simply listing the components in order, we asked them to imagine eating the meal bite by bite. In this way, we can manipulate duration of imagined meal as well as the sequence of components, as occurs in actual eating situations.

Method

Participants

Participants were seven males and 13 females, of which 14 were college students, and the remainder older adults. Each participant generated data from six imagined meals for a total of 120 meals. There were two sessions, each lasting from a half-hour to an hour. The pay for the study was 15 dollars upon completion of the second session.

Table 2
Study 1: Correlations between various like/dislike line ratings during eating and overall remembered like/dislike category scores

Group	*N*	Envelope 1	Envelope 2	Envelope 3	Average	Maximum	Minimum
All	173	0.48	0.38	0.39	0.77	0.62	0.58
Rise	26	0.74	0.65	0.80	0.83	0.80	0.74
Fall	30	0.58	0.67	0.37	0.71	0.58	0.37
Mid-peak	18	0.67	0.71	0.47	0.68	0.71	0.58
Start peak	14	0.65	0.77	0.73	0.80	0.65	0.77
Mid-trough	33	0.63	0.55	0.52	0.74	0.62	0.55
End trough	16	0.65	0.15	0.28	0.60	0.26	0.28

Procedure

In the first session, participants were presented with a book containing pictures of 83 different foods, which they rated one at a time. The pictures of the foods were taken from food and gourmet magazines and were close to life-size color photographs of the ready-to-eat dishes. The participant rated each one of these foods/dishes based on the picture and personal knowledge of the food on a scale of −50 (dislike extremely) to +50 (like extremely) with 0 as neutral. Then, the participants were given a sample imaginary meal composed of pictures similar to those just rated. The imaginary meal consisted of an appetizer, drink, side-dish, main dish and dessert. The pictures of these items were presented in front of the participant, course by course. The experimenter directed the participant as to when to take an imaginary bite of each food, thinking of the appearance, flavor, aroma and texture. Thus, the participant actually "ate" the meal using his/her imagination, beginning with the appetizer. After the appetizer, the side and main were "eaten" in the order indicated by the participants as their normal pattern of eating (e.g., alternating between the main and the side finishing both at the same time, eating the main first and then the side, or eating the side first and the main). The dessert was "eaten" last. The participant indicated every time he/she wished to "drink" and "sipped" the beverage on that occasion.

For example, the dialogue of the main course might sound like the following. The experimenter says, "Take a bite of your chicken." Some participants might elect to act out the motions pretending to take a forkful of chicken and chew. After a few seconds' pause the experimenter might say, "Take another bite of your chicken." The participant would comply and then might reply, "drink", reaching for his or her imaginary drink. The dialogue would continue in this manner throughout the meal.

The standard size of the side, appetizer, and dessert consisted of four bites in this experiment. The main dish consisted of eight bites. The first session ended with completion of the sample meal.

In the second session the participant "ate" six meals composed of the foods that were rated in the first session. They did not rate the individual foods in the second session, but only gave an overall rating of each meal after it was completed. About 1 week separated the first and second sessions. The imaginary meals were designed separately for each participant, based on the ratings from session one, so as to follow a rising, falling or flat pattern of hedonic ratings (Table 3). The second, third, and fourth meals were randomly assigned as standard portions rising, falling or flat (note that the standard portions included a double-sized main course). The first and fifth meals were randomly assigned as either the first double rising and the fifth double falling (Pattern A, Table 3) or first double falling and fifth double rising (Pattern B, Table 3). In the double rising pattern, the number of bites of the dessert in the rising pattern doubled (from four to eight), while in the falling pattern the number of bites of the appetizer doubled (from four to eight). Thus, in either case, the most preferred dish was doubled in size. The sixth meal was an exact repeat of the first meal with either the appetizer or dessert doubled as previously described.

The participants rated each food in Session 1 from dislike extremely (−50) to like extremely (+50), with 0 specified as neutral. During the second session, the meals were rated in the same way. Each meal was designed to have a range of at least 40 points in the ratings of the foods that composed the meal (except for the "level" meal) and a positive overall rating. All meals for a participant were designed to have about the same average of the ratings of components and no component with a rating below −10 was "served". Components were ideally unique to each meal but in a few cases a component was used in two meals. The main and side dishes were chosen to have ratings as close together as possible. A drink to accompany all meals was chosen for each participant with an average rating for the rest of the meal.

A sample rising meal might be the following. First a salad, mixed greens with mandarin oranges (rated zero by the participant). Then a main, fried chicken (rated 21), and a side, green beans (rated 23), would be served. For dessert, pumpkin pie, (rated 42), might be served. A beverage, water, (rated 22) would be consumed throughout this meal as well as the other five meals.

After the end of each meal there was a 3-min interval followed by the rating of the meal. During this interval the experimenter filed the pictures from the previous meal and "prepared" for the next meal. During this same time, the participants often just sat and waited, sometimes talking, or asking questions unrelated to the study (related questions were discouraged and not responded to). After this interval the participant indicated how much he/she liked the meal he/she just completed from −50 for dislike

Table 3
Study 2: Patterns for the imagined meal in Session 2

	Meal number					
	1	2	3	4	5	6[a]
Pattern A	Double rising	Fall, rise or level	Fall, rise or level	Fall, rise or level	Double falling	Double rising
Pattern B	Double falling	Fall, rise or level	Fall, rise or level	Fall, rise or level	Double rising	Double falling

[a]Same pattern as at the first meal.

extremely to +50 for like extremely, with 0 as neutral. The experiment was completed at the end of the sixth imaginary meal.

Analysis of data

By virtue of the design of this study, all meals had approximately the same average hedonic value, as calculated from their component dishes. The only potential exception to this would be a consequence of the fact that our average equalization equally weighted all components of the meal. In the doubled appetizer (falling2) or doubled dessert (rising2) conditions, a weighted average would produce a somewhat higher value. However, the results of Study 1 indicate that this manipulation should have no effect. For this reason, we have proceeded in this analysis under the assumption that all meals had an equal average value. Of course, this assumption would make it easier to reveal an effect of duration (disconfirm duration neglect). However, the type of correlations used in the first study, comparing experienced to remembered pleasure, are inapplicable, since there is minimal variation in total experienced pleasure by virtue of the design of this study. For this reason, our analysis compares remembered pleasure of pairs of meals of different patterns but equal value in average components.

One remembered rating for one meal was way out of line with all others (−20), for meals designed to have positive mean ratings of components of about +20 (its average component rating was 17.2). This rating (for a double rising meal) was eliminated and so appropriate comparisons using this category of meal lost input from one participant.

Results

Consistency: meals one and six

In this design, the first and sixth meals were identical. This allows for some determination of the consistency of the overall meal rating following the same meal, a few weeks apart. We use the mean absolute difference between two meals as the measure of consistency, and compare the mean absolute difference between the identical meals to the mean absolute difference between one of the identical meals and three of the other meals experienced by the same participant. If there is some consistency in remembered ratings, the difference should be smaller between identical meals one and six than between meals one and three, or four and six, or three and four. We calculated the appropriate values for these four meal pairings.

As predicted, the mean absolute 1–6 (identical) difference of 7.8 is smaller than all of the other three differences (13.1, 13.2, 15.0). The identical pairing is significantly smaller (t-test, $p<0.05$ or better) for all three comparisons. This argues for some consistency (test–retest validity) in the overall ratings of meals.

Duration neglect

A total experienced liking of each meal was calculated as the sum of the ratings of its components, with the main double-size course doubled. An average liking was derived by dividing this total liking by the number of liking "time units", which we estimate to be six (drink, appetizer, main course counted as two, side dish and dessert). We can make a corresponding calculation for the doubled most liked course, adding its value again to the total, and dividing this number by seven to obtain its average. If duration of exposure were effective, total liking would be 40–45 mm higher for the doubled pattern and average liking would be about 3 mm higher. In fact, rising2 averaged 0.41 units higher than rising (s.d. = 16.5), $t(16) = 0.59$, and falling2 averaged 0.44 mm higher than falling (s.d. = 15.1), $t(17) = 0.12$. These tiny effects give no evidence of an effect of duration, and hence support duration neglect.

Rising vs. falling pattern

Rising and falling patterns did not differ reliably. For the standard (non-doubled) rising minus falling, the mean difference in hedonic ratings was 6.15 mm (s.d. = 14.2), $t(19) = -1.94$, n.s. For doubled rising minus double falling, the mean difference was 3.5 mm (s.d. = 12.2), $t(14) = -1.12$, n.s.

Component type and position

There was no general trend in remembered pleasure of the meals over the sequence of meals. The fourth meal had the highest mean rating, but none of the differences are significant.

There were no major differences in the correlations between the different components (appetizer, dessert, main, etc.) and remembered ratings, although the correlation is notably low (−0.07) for the drink rating. The main course (which is always of "double" size) had a slightly higher correlation (0.32) with the overall remembered score than any other meal component.

Peak

The correlation of the peak component with the remembered pleasure (median value of 0.30 across the five meal conditions) was about the same as the values for the different meal components, suggesting no substantial peak effect. Both the rising and falling patterns were designed to have a peak, unlike the flat pattern. The rising meal was significantly higher than the flat meal (mean difference = 7.2, s.d. = 13.8), $t(18) = 2.25$, $p<0.05$. However, the remembered pleasure of the flat meal fell between the rising and falling meals. So the evidence for a peak effect is minimal.

Study 3. Chinese buffet restaurant meal

Introduction

The jelly bean study was well controlled and involved real eating of foods, but the "foods" were of only one sort and did not resemble the components of a normal meal. The imagined meal study simulated a normal meal but

there was no actual ingestion. In this study, we achieved experimentally controlled eating of a substantial meal in a natural situation. Participants ate a number of meals at a Chinese buffet restaurant. In a session some days before the first of a set of five experimental meals on separate days, participants rated a large set of the foods served at the restuarant. Based on these ratings, the experimenter constructed different five-dish meals to produce the desired hedonic patterns.

Method

The study was conducted at a buffet-style Chinese restaurant located on the campus of the University of Pennsylvania. Twenty subjects completed seven lunches (sessions) at the restaurant. Each session lasted between 30 min and an hour depending on the rate of eating, extent of conversation at the table, the session meal pattern and other variables. The participants were five females and 15 males; 15 were students at the University of Pennsylvania and five were staff at the University.

During the first and last sessions, the participants tasted roughly 25 different foods (the number rated is approximate because there were slight changes in the foods offered on the buffet from day to day) and rated how much they liked each of these foods by a mark on a 100-mm line between "like extremely" (scored +50) at the right end and "dislike extremely" at the left end (scored −50), with a slash at the midpoint labeled "neutral" (scored 0). Each food was presented in sufficient quantity that a few bites could be tried, and participants were instructed to take three full bites before rating it. The eaters were served water as a beverage during all sessions. The foods were primarily Chinese but some more American foods were also included.

Two different sets of meals with five components (courses) were constructed, rising (increasing liking from first to last course) or falling (decreasing liking from first to last). The two meals had no component in common. The range between the highest rated (most liked) component and the lowest (least liked) was at least 20 points (Table 4). Both the rising and falling meals were designed to be as flat as possible in hedonic ratings for the four components other than the peak or trough, but the first four (or last four) of these components were arranged in a rising/falling order, as appropriate to the design.

In each of the five central sessions the participants were served the foods from the rising or the falling menu. Two meals were the rising menu and three meals were the falling menu. Before each meal the participants filled out a questionnaire that included the UWIST mood scale (Matthews, Jones, & Chamberline, 1990) as well as questions about the previous meal. On a 100-mm line like those used to rate foods in the first session, the participants marked how much they enjoyed the food, atmosphere, company and overall experience of the meal consumed in the *previous* session from "not at all" (scored 0) to "extremely" (scored 100).

The participants were then served their meal, as dictated by the design of the study (Table 5). In all meals an equal amount of each component was served (except for doubling of the size of one dish in the critical duration neglect manipulation) and the foods were served on one plate; the conditions were therefore identical in food presentation but differed in the order in which the participants were instructed to eat the meal.

Table 4
Study 3: Median hedonic values of components (positions 1–5 in the meal) of rising and falling sequences, based on initial ratings of the components[a]

Meal type	Median value for each dish position in meal					
	1	2	3	4	5	Mean[b]
Falling	25.0	23.5	20.5	18.0	0.0	17.9
Rising	6.0	9.0	12.5	17.5	40.5	16.9

[a]These values could also be calculated as the mean of the medians for the first and seventh session ratings. However, for 2–3 participants, an item in the original menu was not on the menu for the final tasting and so a final rating could not be established for all the participants.
[b]Mean of the medians for dish positions.

Table 5
Study 3: Design of Chinese restaurant menu components and their patterns of hedonic values

Session	1	2	3	4	5	6	7
Menu components	25 foods	Falling[a]	Rising[b]	Falling[a]	Rising[b]	Falling[a]	25 foods
Pattern	—	Falling, rotate, free	Rising, rising[b]	Rotate, free, falling	Rising[b] rising	Free, falling, rotate	—

[a]Falling menu meals 2, 4, and 6 were randomly assigned for each participant to falling, rotate or free patterns.
[b]Rising menu meals 3 and 5 were randomly assigned to standard rising pattern or to rising2, in which there was a double portion of the final (most liked) dish.

Across the five experimental meals, only two menus were served, either the set of foods in the "falling" menu (at the second, fourth, and sixth meals) or the set included in the "rising" menu (at the third and fifth meals). Within these two categories, however, different types of meal were presented in random order. The three types of falling-menu meals were the falling, rotate and free-eat meals.

In the falling meal the participant was instructed to consume all of the highest rated component (identified simply by food type, not by its rating position) before moving on to the next component and then to consume the next most liked component, and so forth.

In the rotate meal the participant was instructed to consume one bite of the highest rated component, then one bite of the next highest, and to continue in this pattern until the lowest rated component was sampled. This cycle was repeated until all of the food was consumed. The participant was asked to try to complete eating each component on the same final cycle.

During the free-eat meal, the participants were free to eat their foods in any order they chose. The experimenter, sitting at the adjacent table, observed and recorded the pattern of consumption.

The rising meal was similar to the falling meal in structure: the participant ate the lowest rated food first, then the next lowest rated food, and so forth. The other rising meal had the same components and the same sequence but the favorite (last) component was doubled in portion size. The structure of these patterns was not described to the participant during the study.

At the end of each meal, the participants rated their enjoyment of the meal (food), the company, and the atmosphere as specified above. In the last (seventh) session, the participants completed the initial questionnaire again, including ratings of the previous meal, and then again rated liking for the 25 foods presented in the first session. After completing the food ratings, participants completed a questionnaire about their eating habits.

Fortunately, the designated foods for the two meal sequences for each participant were only rarely not in the buffet on the day they were supposed to be served. In the rare case that they were absent, the missing food was replaced by a food, which the participant had given the same liking rating on the initial session. Note that the only ratings made by participants during the experimental meals concerned their total experience; they did not rate the components during these meals. The component ratings came from the values produced in the initial and final sampling sessions.

The interval between sessions varied from two to five days. All meals were consumed at a restaurant table simultaneously with two to five other participants who were present on that day. For the most part, participants ate with the same group of people at each meal and in some cases some of the participants knew each other. Each participant was served a unique meal and consumed it in the particular pattern that was assigned to him/her for that day. After the meal and the ratings, the participant was allowed to eat additional food if still hungry. When this occurred, the participant usually added soda and a dessert. Conversation about matters other than the food was encouraged among the participants. The full design of the study sessions is shown in Table 5.

Results

Rising and falling hedonic patterns

As the construction of rising and falling meals for each participant was constrained by that participant's ratings and since we did not allow the same item to appear in both hedonic patterns, there is not a perfect mirror relation between the two types of meal (Table 4). Note the especially large jump in the final dish under the rising condition.

Consistency of component ratings from first to seventh session

For the falling-menu meal, the components before the five meal sessions were rated 4.5 points (s.d. = 11.2) higher than the same components after the 5 meals, $t(18) = 1.73$, n.s. The mean absolute difference in scores between before and after was 9.4 points. For the rising meal, the before components were a negligible 0.3 (s.d. = 9.0) higher in liking than the same components after, $t(18) = 0.15$, while the mean absolute difference between the first and seventh meals was 7.4. A reasonable consistency is important because a few of the 25 foods that were tested in the initial session and used in the meal sessions had disappeared from the menu by the time of the final session's re-rating of these foods. As a result, we have used the initial ratings to calculate the hedonic value of each of the meals.

Overall meal memory: immediate and delayed ratings

Ratings made immediately after the meal and those made a few days later yielded about the same values. There is no evidence for a particular bias (either increasing or decreasing evaluation) that ensues over the time period between immediate and delayed evaluation. Across all 20 participants and the 5 meals (100 observations), the mean difference between the delayed and immediate overall ratings of liking was only 1.2.

Duration neglect

For the immediate overall liking ratings, the meal with the double last/most-liked portion (rising2 $\times$) was slightly but not significantly *lower* than the rising meal difference (mean = −3.4, s.d. = 21.8), $t(19) = -0.71$, n.s. For the delayed ratings, the difference was even smaller and in the opposite direction (mean = 2.6, s.d. = 17.7), $t(19) = 0.66$, n.s. This failure to detect an effect of these analogs of duration of stimulation is consistent with the lack of effect reported by Kahneman and others for aversive stimuli.

As another measure of an effect of duration of stimulation on memory, the correlation between the rating

of the last (highest) meal component and the overall rating might be higher for rising2 ×, since there is twice as much of this component in the rising2 × meal. There are four relevant correlations of the last dish with the rising pattern (last dish rating before sessions and last dish rating after sessions with initial and delayed overall meal ratings) and a corresponding four correlations between the last dish and the doubled rising pattern. The mean of the four correlations is $r = 0.22$ for the rising pattern, and $r = 0.19$ for the double rising pattern. The difference is, of course, not significant, but is also in the direction opposite to what would be predicted if duration mattered.

Another prediction of duration neglect is that the difference between the final course rating and the overall rating will not be significantly smaller for the double final course (even though it contributes twice as much to the online experience). In fact, for the immediate memory, the rating of the final course dish was actually slightly but not significantly closer to the single rising meal (by a mean difference of −3.45). For the delayed overall meal rating, the results were even less slightly in the direction suggesting a greater influence of the double meal (mean difference of 2.60). The results are not even in the direction of increased duration having a bigger effect on memory.

Rising vs. falling patterns and primacy and recency

The rising and falling patterns are close in overall hedonic value of the components (Table 4) but there is an unfortunate difference in the actual patterns. The high point (40.2) for the rising is substantially higher than the high point for the falling (25.0). Note that, by design, there is a standout peak in the rising pattern, and a standout trough in the falling pattern.

The end (recency) principle implies that the rising pattern of purely positive affect should be rated retrospectively as more positive. As can be seen from Table 6, this is true in terms of direction for both immediate and delayed ratings of the actual rising and falling sequences. However, neither difference approaches significance: rising minus falling, immediate mean = 4.8, s.d. = 19.6, $t(19) = 1.11$; delayed mean = 2.8, s.d. = 20.4, $t(19) = 0.60$. The very high rating of the last dish in the rising meal (in comparison to the first dish in the falling meal) would bias the results (if there was a recency effect) towards showing a significant effect; yet there is far from any significance in the differences. Thus this rising/falling contrast fails to detect a recency effect in likings for remembered dishes.

Table 6
Study 3: Overall ratings by meal type ($N = 20$)

Meal type	Immediate		Delayed	
	Mean	SD	Mean	SD
Rising	21.4	17.1	20.0	16.8
Rising2 ×	18.0	15.5	22.6	13.0
Falling	16.6	16.1	17.8	14.2
Free eat	23.2	15.4	19.5	18.6
Rotate	23.0	11.0	23.4	9.2

Among the set of falling hedonic patterns (falling, rotate, free-eat), which all contained the exact same components, the general pattern was that rotate and free-eat had higher ratings of liking than the falling pattern, which fits the claim that the falling pattern is least desirable. However, the most reliable difference among the six possible comparisons (three for immediate ratings, three for delayed) was the Falling pattern being marginally worse than the Rotate pattern in the immediate condition (Fall minus Rotate: mean = −6.5, s.d. = 13.0, $t(19) = -2.23$, $p < 0.04$. Given the multiple t-tests, this is at best marginal support for a recency effect in this design.

Overall, the results provide weak indications that alternation or free choice yields a slightly better experience, with rising below that, and falling lowest of all.

Discussion

These three studies all deal with sequences of foods in "meals" but use very different definitions of meals, very different procedures and a variety of measures. The findings are consistent across this range of designs when the effects of duration neglect and rising-falling patterns are measured in the same way. The same line rating of overall liking was made in all three studies and so composite values from each study can be compared (Fig. 1).

A total measure of the effect of the difference in duration between two and fort jelly beans can be obtained from all cases in which the third set of jelly beans was either peak high or there was a rising pattern: in either case, the third bean is highest and so we are comparing single or double durations of the highest value. This corresponds to studies two and three, where the doubling in duration was the highest value.

As detailed in the description of results for the three studies and summarized on the left of Fig. 1, there is consistent lack of evidence for an effect of duration of exposure, supporting Kahneman's evidence for "duration neglect."

For rising and falling, we have direct tests in all three studies. (In Fig. 1, when more than one pair of conditions are relevant, the scores are averaged.) The greater liking for rising than for falling patterns, reported in most prior research (but not in Anderson & Norman, 1964), is not supported.

Also, unlike the evidence on aversive experiences, and positive experiences with music, we find no clear signs of recency (end) or peak effects. We also get no indications of the primacy effect which is a common finding in memory studies but has not been reported either in previous studies of remembered hedonic value.

Our most striking finding is the presence of duration neglect. This phenomenon has major implications for the structuring of a high quality of life, since it suggests that from the memory perspective, amounts or durations of

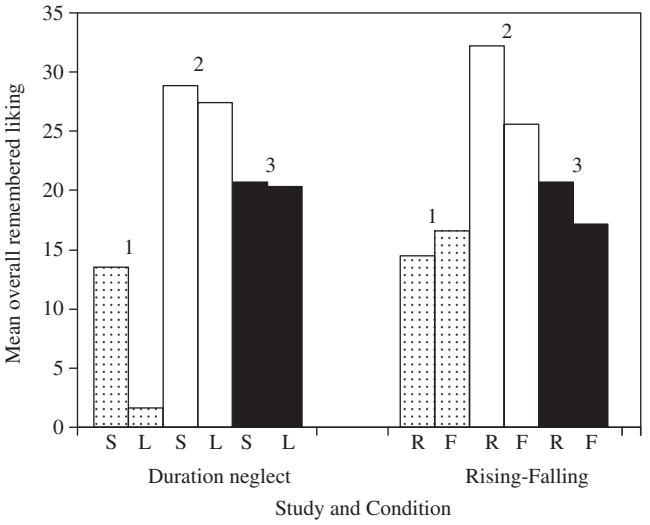

Fig. 1. Summary of findings on duration neglect and rising vs. falling patterns across the Studies 1 (⬚ = jelly beans), 2 (□ = imagined meal) and 3 (■ = chinese meal). S = short duration, L = long duration, R = rising pattern, F = falling pattern.

very positive experiences may have a minimal effect on remembered pleasure. A few bites of a favorite dish in a meal may do the full job for memory (though obviously not for experienced pleasure). However, our results leave open the important question: what is a long or short duration of eating within a meal? That is, we may have varied duration in an irrelevant range. What, in particular, is a segment for a person who eats rather randomly from among the three dishes on her plate?

It is surprising that, more often than not (Fig. 1), adding more of a very pleasant part of a meal made the meal slightly worse. That is, if anything, we have indications of a reverse duration effect. Such a duration negativity effect for positive events makes little sense. However, it should be noted that our designs added extra positive experiences to the comparison meal. This means that the participant experienced extra desirable food usually at the end of the meal, when satiation may have operated against appreciation. Perhaps a better design would be to remove the same amount of food from the middle of the meal as is added to increase duration of the positive segment at the end of the meal.

Future research may benefit from the findings in these studies, especially if designed with greater statistical power. With respect to duration neglect, it would be desirable to explore a wider range of durations. Also, the work of Ariely and his colleagues suggests important framing considerations in the determination of duration effects. For example, at the extreme, would not a duration effect be likely if a favored food was presented as both appetizer and dessert?

Work on primacy, recency and peak effects would perhaps benefit from more salient peaks. The hedonic peaks (and troughs) that our participants experienced were only modest. This is likely to be the case for real meals in real settings, because people usually select the components. However, there are some meals in which there is one and only one outstanding dish. We expect that in those cases, there would be a predominant peak effect, and indeed, that some weeks later, the only dish remembered from the meal would be the outstanding one.

It is notable that primacy effects, though common in memory research, do not appear in the studies on memory for negative affect, or in the present study. However, they did appear in Anderson and Norman's (1964) study of liking for listed sequences of food that composed a meal. It is also possible, of course, that remembered pleasure is affected by different features in the pain vs. food domains, or more generally, negative and positive domains.

We propose that the meal is a good natural unit to investigate many of the issues raised in this paper. Each of our studies made some compromises in design but the compromises differed from study to study. The jelly bean study had by far the largest *N*, but the least pre-programmed "meal" profiles. The other studies might be criticized in retrospect for small numbers of participants, in the face of principally null results (though duration neglect happens to be supported by lack of difference). Nevertheless, the results of jelly bean study reinforce the conclusions from the smaller studies.

We hope therefore that these studies have added some "meat" to the small number of studies of the important issue of the relationships between experienced and remembered pleasure. This is only the first step in understanding how meals are remembered and, more generally, how pleasure is represented in memory.

Acknowledgments

We thank Unilever Research for providing the funds that supported this research.

References

Anderson, N. H., & Norman, A. (1964). Order effects in impression formation in four classes of stimuli. *Journal of Abnormal and Social Psychology*, *69*, 467–471.

Ariely, D., & Carmon, Z. (2000). Gestalt characteristics of experiences: The defining features of summarized events. *Journal of Behavioral Decision Making*, *13*, 191–201.

Ariely, D., & Levav, J. (2000). Sequential choice in-group settings: Taking the road less traveled and less enjoyed. *Journal of Consumer Research*, *27*, 279–290.

Ariely, D., & Zauberman, G. (2000). On the making of an experience: The effects of breaking and combining experiences on their overall evaluation. *Journal of Behavioral Decision Making*, *13*, 219–232.

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*, 577–606.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality & Social Psychology*, *65*, 45–55.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*, 401–405.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experience's utility. *Quarterly Journal of Economics*, *112*, 375–405.

Loewenstein, G. F., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review*, *100*, 91–108.

Matthews, G., Jones, D. M., & Chamberline, A. G. (1990). Refining the measurement of mood: The UWIST mood adjective checklist. *British Journal of Psychology*, *81*, 17–42.

Murdock, & Bennet, B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488.

Pliner, P., & Rozin, P. (1999). The psychology of the meal. In H. Meiselman (Ed.), *Dimensions of the meal: The science, culture, business, and art of eating* (pp. 19–46). Gaithersburg, MD: Aspen Publishers.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, *116*, 3–8.

Rogozenski, J. G., Jr., & Moskowitz, H. R. (1982). A system for the preference evaluation of cyclic menus. *Journal of Food Service Systems*, *2*, 139–161.

Rozin, A., Rozin, P., & Goldberg, E. (2003). *The feeling of music past: How listeners remember musical affect*. Manuscript submitted for publication.