

Gestalt Characteristics of Experiences: The Defining Features of Summarized Events

DAN ARIELY^{1*} and ZIV CARMON²

¹*Sloan School of Management, Massachusetts Institute of Technology, USA*

²*Fuqua School of Business, Duke University, USA*

ABSTRACT

In this paper we take stock of recent research on how people summarize and evaluate extended experiences. Summary assessments do not simply integrate all the components of the evaluated events, but tend to focus on only a few features (*gestalt characteristics*). Examples of these defining features include the rate at which the transient state components of the experience become more or less pleasant over its duration, and the intensity of the state at key instances, in particular the most intense (*peak*) and the final (*end*) moments. It is not yet sufficiently clear which specific gestalt characteristics dominate summary assessments of experiences, nor how this differs across types of experiences or measurement approaches. To address some of these issues, we describe new research in this area, discuss potential methodological difficulties, and suggest directions for future research. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS summary evaluation of experiences; extended experiences; experience profile; pain

Common experiences such as a bus ride, a meal, and a college course unfold over time through a stream of transient states that may vary in intensity and even in sign from moment to moment. Such extended episodes can be depicted in the form of an *experience profile*, as illustrated in Exhibit 1 (see next page). Decision makers often evaluate the overall goodness or badness of extended episodes to assist future decisions or convey to others (presumably to aid them in making decisions). In some cases people evaluate the goodness or badness of episodes they experienced in the past; in other cases they predict the overall desirability of episodes that they have not yet experienced. Retrospective summary evaluations are an important input into decisions to repeat (or not repeat) past experiences; similarly, prospective summary evaluations may determine whether or not the decision maker pursues that experience.

Such summary evaluations are not only important for decision making, but can also have direct hedonic consequences. Prospective summary assessment of an experience can evoke sensations such as

* Correspondence to: Dan Ariely, Massachusetts Institute of Technology, 38 Memorial Drive, E56-329, Cambridge, MA 02142, USA. E-mail: ariely@mit.edu

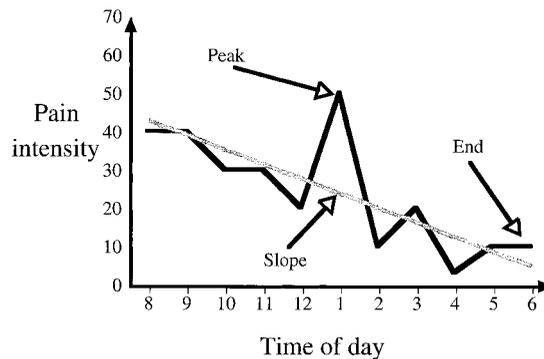


Exhibit 1. An example of an experience profile and three of its gestalt characteristics, based on the data of Subject 17 in the hospital study described later in this paper (assessment of a painful day at a hospital). *Peak* is the maximum intensity, *end* is the intensity at the final moment of the experience, and in this case *slope* is a single measure of the profile's overall linear trend (in gray)

anticipation and dread before the experience ever takes place (Loewenstein, 1987), and retrospective summary evaluations influence the utility that can be derived from consuming memories after the experience occurs. For example, a brief exotic vacation can produce memories to be consumed for the rest of one's life. Summary evaluations can also affect the experience itself. For example, anticipating an unpleasant flight can not only result in negative feelings long before we set foot on the airplane, but can also exacerbate the impact of events in the flight such as momentary turbulence.

When people summarize experiences they do not simply combine the intensity of their actual experiences. Instead, they appear to extract only a few defining features (gestalt characteristics) of these sequences, which they combine into overall summary evaluations of the sequences (see Kahneman *et al.*, 1993; Carmon and Kahneman, 1996). Such features include particularly salient characteristics such as the most intense state (peak) and the final state (end) of the experience. As another example, a variety of empirical studies show significant conflicts with the well-known discounted utility model, a prominent economic view on the assessment of consequences that occur over time. Loewenstein and Prelec (1991, 1993), for example, demonstrate that contrary to the prediction (based on time discounting) that one should prefer that better outcomes precede worse ones, decision makers prefer improving sequences in which worse outcomes precede better ones.

The value of the research stream on the evaluation of experiences extends well beyond illustrating anomalous patterns of behavior. Since summary evaluations play a central role in common judgments and decisions, studying how these assessments are formed and what information they express is clearly important. In particular, it seems conceptually valuable to identify the specific features of experiences that are central to summary assessments and understand their impact. For example, we believe it is important to understand why and how a few moments in an experience appear to be over-weighted and others under-weighted in summary evaluations. From a practical perspective, understanding how people assess and judge experiences can also be beneficial. It can help providers of experiences, for example, serve their customers more effectively, simply by better allocating existing resources (e.g. Carmon, Shanthikumar, and Carmon, 1995).

Our goal in this paper is to take stock of research on the summary assessment of experiences, presenting some of what we now know about this topic as well as what we would like to learn. To that aim, we briefly review a growing body of recent research, and then offer our view on emerging conclusions, methodological considerations, and directions for future research.

WHAT WE KNOW ABOUT GESTALT CHARACTERISTICS OF SUMMARIZED EXPERIENCES

Summary assessments of experiences have been studied with several methodological approaches and in a variety of empirical contexts (for details see below). An important overarching conclusion is that in such assessments people do not simply integrate the transient states they experienced as the events unfolded. Rather, two types of defining features of the affective profile of an experience (i.e. gestalt characteristics) appear to dominate overall retrospective evaluations. One reflects the change over time in the intensity of the transient state components. Prominent examples of such characteristics include the trend of the profile (Loewenstein and Prelec, 1993; Ariely, 1998) and its rate of change (Hsee and Abelson, 1991; Hsee, Abelson, and Salovey, 1991). The other type of gestalt characteristic reflects the intensity of the transient experience at particular key points in time. Specifically, a variety of studies found that the momentary experience at the most intense and final moments (peak and end, respectively) consistently accounted for global retrospective evaluations (Varey and Kahneman, 1992; Kahneman *et al.*, 1993; Fredrickson and Kahneman, 1993; Redelmeier and Kahneman, 1996; Baumgartner, Sujan, and Padgett, 1997). Features of the simple sum of the transient state experiences like duration, on the other hand, were found to have no significant effect on summary assessments. More precisely, sometimes variables like duration do appear to affect summary evaluations, but the *marginal* impact is minor, since they tend to have the same impact on gestalt characteristics like the final transient state.

By and large, the findings of the many studies that have explored summary assessments of experiences appear consistent. In fact, we believe that differences that have been found are mostly due to either situational moderators or differences in the methodological approach. We will elaborate on this idea below, but first we describe a new field study we conducted at a local hospital. The study examined the relative importance of specific gestalt characteristics (end, peak, and slope) of the transient state components, in a naturalistic setting in which assessment of the experience can have significant implications.

A field study on the assessment of a painful day at a hospital

Thirty-seven patients in the bone marrow transplant unit of a local hospital participated in this study. Once every hour (between 8 a.m. and 6 p.m.) a nurse entered the room and asked each patient to rate the pain they experienced on a 0–100 scale (0 represented no pain, and 100 the worst pain they could imagine). At the end of the day (7 p.m.) the nurse asked each patient to rate the overall pain they experienced throughout that day on the same 0–100 scale.

The hourly pain ratings reported ranged between 8 and 64, and tended to increase throughout the day. The average pain rating was 29, and the standard deviation was 16. To examine how patients' overall pain ratings corresponded to gestalt characteristics of their hour-by-hour pain ratings, we regressed the overall ratings on four gestalt characteristics of the corresponding pain profiles. These included: the rate of change in the hour-by-hour pain ratings (the linear trend we named *slope*), the average, maximum, and final pain ratings (see Exhibit 1 for an example of one patient's experience profile; see Exhibit 2 for a summary of the results). The regression model was significant [$F_{(4,32)} = 66.0$, $p < 0.0001$], with an R^2 of 0.89. More importantly, both the intensity of the final state and the slope were significant predictors of the summary evaluations, while the effect of the average pain rating was not significant. Thus, consistent with the findings of earlier studies, we find that overall evaluations do not reflect simple integration of the entire experience, but rather seem heavily influenced by typical gestalt characteristics. Note that because the duration was the same for all patients, the total amount of pain is linearly related to the average pain. Thus we can also conclude

Exhibit 2. Regression coefficients for the hospital study

	Standard coefficient	<i>t</i> -value	<i>p</i> -value
Final (end)	0.525	2.85	<0.01
Slope	0.245	2.49	<0.02
Maximum (peak)	0.147	1.24	0.22
Average	0.119	0.99	0.33

that the total amount of pain experienced throughout the day did not affect patients' overall evaluations.

Unlike findings of earlier studies (e.g. Kahneman *et al.*, 1993; Redelmeier and Kahneman, 1996), in this study the most intense state (peak) was not a significant predictor of overall evaluations. This difference may be due to the moderating effect of a unique aspect of the stimuli we studied on the role of the maximum intensity (peak). More specifically, we believe that the different result may be due to frequent recent exposures our subjects had to similar experiences. Unlike the Kahneman *et al.* studies, in our study subjects had encountered many similar stimuli prior to the study, since they were all long-term patients in the bone marrow transplant unit. Hence, the most intense pain experienced during the study day may not have been clearly distinguishable from pain experienced on previous days. Furthermore, the most intense pain during that day was probably less intense than the most intense pain experienced in previous days.

Additional support for the notion that recent exposure to similar stimuli moderates the impact of the most intense state (peak) on overall evaluations comes from a different study of physically painful events (Ariely, 1998). In that study the slope of the experience profile was the most important predictor of retrospective summary evaluations, while the peak intensity had a significant but much smaller effect. Although subjects had no previous experience with that particular type of stimulus, the two experiments effectively provided substantial experience, since they consisted of exposure to many trials (40 in one study and 70 in another). Thus, the salience and separability of the most intense pain may moderate its impact on the overall evaluation of the experience.

Next we review findings relating to other factors that may account for differences in the impact of particular gestalt characteristics. We begin with additional situational factors, and then discuss methodological considerations in the following section.

Situational moderators of the effects of gestalt characteristics

Chapman (2000) suggests that the preference for improvement in the trend of experiences is moderated by expectations about that trend. For example, subjects preferred that their skin, initially appear young and become increasingly wrinkled with age, rather than the typical preference for improvement over time. This suggests to us that summary evaluation of an experience depends not only on its transient state components in and of themselves, but also on inferences that can be drawn from them (more on this idea in the next section).

Ariely and Zauberman (2000) show that the way subjects summarize an experience depends on whether the experience is perceived to be composed of single or multiple parts (i.e. continuous or discrete). Their results show that the preference for improving trends over deteriorating ones is substantially reduced if the same experience is composed of discrete parts, demonstrating that the cohesiveness of an experience impacts the relationship between its pattern and overall evaluation. This appears reasonable in light of the idea that summary assessments partly reflect inferences about future states. Thus, the trend of an experience may well appear to convey more information about the future

states of continuous experiences than about the future states of multiple discrete or segmented experiences (for more on this idea see the next section).

Carmon and Kahneman (1996) find that global retrospective evaluations of queuing experiences are dominated by the final affective state (end). In that setting, summary assessments effectively ignored the transient state components that precede the ending of the experience. As an example, some queuing events that were dissatisfying up until before their ending but concluded on a positive note were summarized positively. Carmon and Kahneman suggest that this 'end rule' applies to goal-directed experiences (such as competitions, as opposed to process-directed experiences), or ones in which the transient real-time experience is influenced mostly by anticipation or dread of what may happen (such as waiting for a dental appointment, rather than what actually happened). Based on this idea we hypothesize that in some situations summary evaluations may 'reinterpret' what had actually been experienced based on information received after an experience ends.¹

Using hypothetical scenarios such as the following, we examined this idea in an exploratory study. Condition 1 described a 30-year period of a marriage as having been very satisfying. Condition 2 of the same scenario added that several years after that 30-year period, it was discovered that the partner had an affair. Subjects in the two conditions (different people) were asked for their summary assessment of the 30-year period of the marriage. As expected, subjects in Condition 1 summarized the 30-year marriage experience very positively, while those in the second condition summarized it significantly less positively. Interestingly, the mean rating in Condition 2 represented overall dissatisfaction with the 30-year marriage although the entire period was clearly satisfying as it took place. Thus, although subjects realized that the summary was to reflect their overall assessment of the actual experience, their ratings were heavily influenced by information received after it was already over (and could therefore not affect the actual experience). It is as though subjects felt that emotions experienced as the event unfolded were 'not justified' and should hence be discounted in a retrospective summary. While other scenarios showed similar results, future research should examine if this holds for non-hypothetical experiences. If it does, we find it interesting that summary assessments may sometime assign little weight to the actual experience.

WHY THE GESTALT CHARACTERISTICS AFFECT SUMMARY ASSESSMENT

As we suggested earlier, research to date clearly supports the notion that two types of defining features of the experience (gestalt characteristics) are significant predictors of its summary assessment. One type represents configural aspects of the experience such as the trend of its profile, and the other represents the transient state at key moments. A remaining challenge is to identify the specific gestalt characteristics of common types of experiences and understand how they are encoded in memory and how they influence judgments and decisions. We believe that a productive approach toward this is to consider people's underlying task goals in encoding and summarizing experiences. The notion is that taking into account the purposes served by summarizing an experience can provide insight into which of its features will be important to the summary assessment (see also Carmon and Ariely, 1998; Fischer *et al.*, 1998; Wertenbroch and Carmon, 1997). Note that we believe that superficial cues may evoke mindless pursuit of such goals, even when mindful examination of the particular situation clearly does not justify this (to learn more about mindless behavior see Langer, 1989; for more on this idea, see the Directions for Future Research section). Below we examine two types of goals, which we name extrapolation and encoding.

One goal of assessing and summarizing experiences is to facilitate effective decisions by helping to predict future states. We refer to the view of summary evaluations corresponding to this task goal as

¹ We thank Danny Kahneman for stimulating our thinking in this direction.

extrapolation. To illustrate, imagine a patient undergoing a painful and long medical treatment that becomes less painful over time. Based on such a trend of decreasing pain, the patient may extrapolate that the future is likely to be less painful, or even infer that she is closer to full recovery. In addition, an improving trend may also improve the transient experience itself by increasing anticipation or reducing dread. This extrapolation-based explanation can explain why configural aspects of the experience (i.e. the relationship between its intensities, such as trend) and the final state (end) can help in predicting future states.

A secondary goal of summary evaluations may be to cope with people's cognitive limitations, which requires efficient representation of the many detailed characteristics of stimuli such as experienced events. We refer to the account of summary evaluations that corresponds to this goal as encoding. For many situations such parsimonious representation of experiences with a few key characteristics seems highly adaptive, and hence reasonable. The alternative, collecting and representing each of the individual transient states of an experience, may often be too demanding of the person's cognitive resources, and offer relatively minor marginal benefits. A similar argument can be made for limited resources during the information retrieval and response generation process. This encoding view predicts that only a few features of the experience will be encoded and influence summary evaluations (see Kahneman, 1995). Examples of such features are likely to include particularly salient characteristics such as the most intense state (peak) and the final state (end) of the experience.

Recent work on perception of multiple objects (sets) points to similar conclusions. Ariely and Burbeck (1998) found that respondents accurately recognized the mean size of an item in a set but not the size of individual items in the set. Specifically, their subjects briefly observed stimuli consisting of a set of circles, and were then asked whether or not a particular circle appeared in the set. On other trials subjects saw the same sets of circles and were then asked whether a particular circle was larger or smaller than the mean of the set. Results showed that as the number of circles in a set increased, knowledge of individual circles in the set dropped rapidly to chance level, whereas knowledge of the set's mean remained precise. Based on these perception studies and research on experiences we described earlier, we believe that stimuli consisting of many units, whether scattered over time or space, tend to be represented by summary measures rather than by the individual components.

METHODOLOGICAL CONSIDERATIONS

Although research on the assessment of experiences has used somewhat different methodological approaches, the considerations that underlie the selection of a particular methodology have rarely been discussed. Because we believe that such methodological differences are a major cause for discrepancies between the findings of studies, we list dimensions on which the approaches of prior studies differ. We then discuss two methods in some detail.

Dimensions on which studies differ include: the extent to which events were actually experienced versus hypothetically described; the extent to which experiences were meaningful in and of themselves (i.e. experience-directed, such as a luxurious meal) versus a means to an end (i.e. goal-directed, such as waiting in line); whether experiences were assessed before or after they took place; and whether experiences were generally pleasant (e.g. a meal) or unpleasant (e.g. pain). Additional methodological dimensions that we discuss next include: whether the experience-profile (transient states of the experience) was measured continuously (on-line responses) or not; whether the results were analyzed according to these subjective intensity reports (possible only if on-line measures were taken) or according to the manipulated intensity; whether the experience to be summarized is naturalistic (as in the bone marrow transplant experience in our field study) or experimentally manipulated; and whether experiences were assessed with choices or evaluations (see also Hastie and Park, 1986).

Measuring experience profiles

Studying summary evaluations of experiences obviously requires data about both the experience profiles and their summary evaluations. Two approaches have been used to obtain such data. One considers objective measures of the experience as representing the subjective transient states. For example, if a person consecutively lifts five items weighing 80, 70, 60, 50, and 40 pounds, under this approach the experience profile would be (80, 70, 60, 50, 40). A significant advantage of this approach is the tight control over the independent variable (the objective real-time experience). Thus, it is simple to experimentally manipulate the intensity pattern of the transient experiences and compare the evaluation of one pattern to another. A significant shortcoming that may not be immediately obvious is an implicit assumption that subjective perception of the intensity pattern is not different from the objective pattern (in some cases a sufficient assumption is that the two relate monotonically to one another). In the item-lifting experience described above, for example, this may not be true for reasons such as gradual fatigue (or adaptation), which might cause items lifted late in the sequence to feel more (or less) heavy than heavier (lighter) items lifted earlier in the sequence.

An alternative approach addresses these concerns by obtaining real-time ratings of the intensity of the transient state components of the experience (see e.g. Fredrickson and Kahneman, 1993; Carmon and Kahneman, 1996; Redelmeier and Kahneman, 1996). It is the resulting pattern of perceived intensity, not the actual (manipulated) intensity pattern, that is then compared to the overall evaluation. This approach of measuring the transient experience profile is advantageous because there is no need to assume how the manipulated and perceived experiences correspond to one another. However, a potential disadvantage of this subjective approach is that it assumes that people correctly report their transient perceptions. Moreover, it assumes that the very act of measuring the subjective transient experience profile does not alter the experience (Feldman and Lynch, 1988). Much of the evidence on this issue suggests that on-line measurement does not alter the nature of the results (e.g. Fredrickson and Kahneman, 1993; Carmon and Kahneman, 1996). However, some studies suggest that continuously measuring the subjective transient experience profile can increase the effect of the transient experience and decrease the importance of the trend (Ariely, 1998; Ariely and Zauberman, 1998).

Given the potential pitfalls of both approaches, one solution is to use both methods. This could also help assess the magnitude and significance of this potential measurement problem. Alternatively, there may be measurement techniques that are less intrusive than the ones typically used, such as squeezing a handle to express the intensity of momentary pain, or even to obtain physiological measures. Another possibility is that the two approaches (providing momentary and global evaluations or providing only global evaluations) tap different psychological mechanisms. If so, choosing the more appropriate measure can depend on the domain that is being generalized to. Specifically, if the domain naturally invokes momentary evaluations, then using on-line responses can be appropriate. But if the domain does not naturally invoke momentary evaluations, then not using on-line responses may be more appropriate (see Hastie and Park, 1986).

Evaluations versus choices

In the beginning of this paper we suggested that summary evaluations are important because they impact future decisions. Indeed much of the work on summary evaluations of experiences assumes that retrospective ratings are indicative of prospective choices. However, recent work by Ariely and Loewenstein (1998) illustrates that this may not always be true, suggesting that aspects of the particular situation may moderate the extent to which retrospective ratings predict future choices. Specifically, the consistency of rating and choice may depend on the ease of comparison and the goal of the decision-maker. Next we briefly discuss these two factors.

The comparability argument is that ratings tend to assess each experience independently, whereas choices are more comparative in nature. Consequently, to the extent that some attributes of the experiences are more easily judged individually and others are more easily judged comparatively, ratings and choices may very well diverge (for related ideas see Fischer *et al.*, 1998; Hsee, 1996; Nowlis and Simonson, 1997). In the context of experienced sequences, Ariely and Loewenstein (1998) suggest that the duration of an experience is not easily judged independently, and that it is more easily compared across experiences (when there is a comparison standard). Their results show that the effect of the duration on overall evaluations is indeed greater when the experience is easily compared to others versus when it is not. Similarly, we can expect that the transient state at particular instances (such as peak or end) is better compared across experiences, whereas configural properties (such as slope) are better judged separately.

Another factor, implicit goals underlying participants' assessments, may also differ among choices and evaluations (Grice, 1975). For example, when asked to evaluate a dental treatment, respondents may reasonably infer that the underlying purpose is to evaluate that dentist for possible future treatment. In such a setting it seems reasonable not to base the response on aspects that are idiosyncratic to the person's particular experience (such as their own sensitivity, or the duration of their individual treatment), since they might not generalize to the person receiving the recommendation. Choices, on the other hand, reflect the individual's behavioral intentions, and it is therefore reasonable to weigh such factors more heavily in choice (Grice, 1975). In support of this idea Ariely and Loewenstein (1998) find that the duration of experiences only has marginal impact on overall evaluations that are expressed in a simple rating, but a large effect when the response mode reflects behavioral intentions (e.g. willingness to pay).

DIRECTIONS FOR RESEARCH

We believe that an important direction for future research is to pursue a better understanding of the mechanisms by which gestalt characteristics affect summary assessments. One direction is to more explicitly test the extrapolation-based explanation, that sequences are evaluated partly based on what they imply regarding future states. A possible approach could be to explicitly ask respondents for their predictions about future states and examine how these predictions and evaluations change with different experience profiles. Of course, the expectation is that differences in respondents' projections will map onto their summary evaluations for the corresponding experiences. Another approach could be to manipulate the extent to which extrapolations are feasible, the prediction being that the impact of the profiles will be greater when extrapolations are more feasible. Another direction is to empirically test the broader version of the extrapolation account that we suggested, whereby such task goals are sometimes pursued mindlessly even when mindful examination of the particular situation does not justify this. We find it important to study this broader version of the extrapolation account, since there are situations in which people appear to prefer improvement over time even though mindful consideration of the situation seems unlikely (e.g. in Loewenstein and Prelec's 1993 restaurant example). A competing explanation (in which we have little faith, but seems worthy of an empirical test) is that the desire for improvement is simply a crude heuristic used due to information processing limitations. A simple experiment in which subjects' cognitive load would be manipulated could test this idea.

Another direction is to test the encoding-based explanation, whereby selected aspects of the experience are encoded while others are ignored due to selective attention or memory mechanisms. A more specific version of this explanation posits that the experience is represented by only a few of its characteristics due to processing limitations. To test this idea, one could simply ask respondents directly about their knowledge of the sequence they experienced. For example, at the end of a noise

exposure experience one could ask respondents to replicate the maximum, minimum, average and final noise level that they experienced. One prediction is that the combination of the better remembered aspects will account for the overall evaluations of the sequence (see also Hastie and Park, 1986). Another approach could be to manipulate the attention respondents can pay to particular aspects of the experience (e.g. with selective distractions) to test the prediction that moments that receive greater attention will be weighted more heavily in the summary evaluation.

A variant of the above explanations combines the idea of limited cognitive capacity with that of underlying encoding goals (see e.g. Carmon and Ariely, 1998; Fischer *et al.*, 1998) and reporting goals (see e.g. Grice, 1975; Ariely and Loewenstein, 1998). The idea here is that subjects intentionally attend to some aspects of the experience and not to others. A possible approach to testing this idea is to explicitly manipulate the goal of the respondents (see also Lichtenstein and Srull, 1987). For example, respondents could be asked to evaluate each experience for recall, comprehension, recommendation to others, and for comparative judgment (similarity or differentiation).

Before concluding this section, we would like to note that although many studies on summary assessment of experiences investigated unpleasant or pleasurable stimuli, virtually no studies examined experiences combining pleasant and unpleasant events. We believe that understanding how pain and pleasure are combined might prove to be a very interesting avenue for future research (see Cabanac, 1971).

CONCLUSIONS

The dominant effect of the profile's trend on summary evaluations of the corresponding experience has interesting implications for effective design of experiences whose profile can be controlled. For instance, summary evaluations may benefit from an (unneeded) initial low point in the experience profile, since this allows for greater improvement over the duration of the experience. Thus, adding a sour note to the beginning of an experience will obviously detract from the overall mean pleasure, yet may enhance the summary evaluation. Note that such a lower starting point may not be advisable if there is a significant probability that the person may abort the experience prematurely. Interestingly, if we ask decision makers directly if they prefer to add an undesirable start to their experience, they will most likely say no. Yet, such an addition may be 'better for them' in terms of their global evaluations. Such inconsistencies between responses to different types of questions about the same experience raise obvious ethical and practical quandaries (Redelmeier and Kahneman, 1996). For example, should an initial low point be added to the beginning of an experience if it improves the global evaluation? Similarly, should a physician end a painful examination with an unnecessary test that is less unpleasant than the preceding procedures to increase the likelihood that the patient will return for follow-up treatment?

On a final note, how can we use the knowledge we have acquired about assessments of experiences to enhance the joy in our lives? As Hollywood producers have evidently realized long ago, ending episodes on a positive note is clearly desirable. Moreover, it can be useful to remember that intense pleasure and substantial improvement are weighted very heavily, and that experiences in different domains appear to be coded and assessed separately. This advocates pursuing distinct and diverse experiences, each providing opportunities for significant improvement over time.

ACKNOWLEDGEMENTS

We are grateful to Celeste Mayer and the nurses at the bone marrow transplant unit for their help with our data collection, and to Jim Bettman, Gretchen Chapman, Danny Kahneman, and John Lynch for their helpful comments and suggestions.

REFERENCES

- Ariely, D. 'Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations', *Journal of Behavioral Decision Making*, **11** (1998), 19–45.
- Ariely, D. and Burbeck, C. Seeing sets, Working paper MIT, 1998.
- Ariely, D. and Loewenstein, G. Encoding sequences and conveying information in ratings and choice: the case of sequences' characteristics and duration, Working paper MIT, 1998.
- Ariely, D. and Prelec, D. Are retrospective evaluations prospective? Variability in experiences and prediction, Working paper MIT, 1998.
- Ariely, D. and Zauberman, G. 'On the making of an experience: the effects of breaking and combining experiences on their overall evaluation', *Journal of Behavioral Decision Making*, **13** (2000), 219–232.
- Baumgartner, H., Sujan, M. and Padgett, D. 'Patterns of affective reactions to advertisements: the integration of moment-to-moment responses into overall judgments', *Journal of Marketing Research*, **34** (1997), 219–232.
- Cabanac, M. 'Physiological role of pleasure', *Science*, **173** (1971), 1103–1107.
- Carmon, Z. and Ariely, D. 'Focusing on the undesirable: a task-goal view of why value can appear so different to buyers and sellers', Working paper, Fuqua School of Business, Duke University, Durham, 1998.
- Carmon, Z. and Kahneman, D. The experienced utility of queuing: real time affect and retrospective evaluations of simulated queues, Working paper, Fuqua School of Business, Duke University, Durham, 1996.
- Carmon, Z., Shanthikumar, J. G. and Carmon, T. 'A psychological perspective on service segmentation: the significance of accounting for consumers' perceptions of waiting and service', *Management Science*, **41**(11) (1995), 1806–1815.
- Chapman, G. 'Preferences for improving and declining sequences of health outcomes', *Journal of Behavioral Decision Making*, **13** (2000), 203–218.
- Feldman, J. M. and Lynch, J. G. 'Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior', *Journal of Applied Psychology*, **73** (1988), 421–435.
- Fischer, G. W., Carmon, Z., Ariely, D. and Zauberman, G. Goal-based construction of preferences: task goals and the prominence effect, *Management Science*, (forthcoming).
- Fredrickson, B. L. and Kahneman, D. 'Duration neglect in retrospective evaluations of affective episodes', *Journal of Personality and Social Psychology*, **65** (1993), 45–55.
- Grice, H. P. 'Logic and conservation', in Cole and Morgan (eds), *Syntax and Semantics 3: Speech Acts*, New York: Academic Press, 1975.
- Hastie, R. and Park, B. 'The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line', *Psychological Review*, **93** (1986), 258–268.
- Hastie, R., Park, B. and Weber, R. 'Social memory', in Wyer, R. S. and Srull, T. K. (eds), *Handbook of Social Cognition*, Vol. 2, Hillsdale, NJ: Erlbaum, 1984.
- Hsee, C. K. 'The evaluability hypothesis: an explanation of preference reversals between joint and separate evaluations of alternatives', *Organizational Behavior and Human Decision Processes*, **46** (1996), 247–257.
- Hsee, C. K. and Abelson, R. P. 'Velocity relation: Satisfaction as a function of the first derivative of outcome over time', *Journal of Personality and Social Psychology*, **60** (1991), 341–347.
- Hsee, C. K., Abelson, R. P. and Salovey, P. 'The relative weighting of position and velocity in satisfaction', *Psychological Science*, **2** (1991), 263–266.
- Hsee, C. K., Salovey, P. and Abelson, R. P. 'The quasi-acceleration relation: Satisfaction as a function of the change of velocity of outcome over time', *Journal of Experimental Social Psychology*, **30** (1994), 96–111.
- Kahneman D. 'When parts are larger than wholes: violations of monotonicity in judgments and preferences', Bartlett Lecture presented to the meeting of the Experimental Psychology Society, Birmingham, UK, July, 1995.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A. and Redelmeier, D. A. 'When more pain is preferred to less: Adding a better end', *Psychological Science*, **4** (1993), 401–405.
- Kahneman, D. and Schreiber, C. A. Determinants of the remembered utility of aversive sounds, Working paper. Princeton University, 1998.
- Kahneman, D., Wakker, P. P. and Sarin, R. 'Back to Bentham? Explorations of experienced utility', *Quarterly Journal of Economics*, **112** (1997), 375–405.
- Katz, J., Redelmeier, D. A. and Kahneman, D. Memories of painful medical procedures, Paper presented to the meeting of the American Pain Society, Washington, DC, and Public Health, November 1997.
- Langer, E. J. *Mindfulness*, Reading, MA: Addison-Wesley, 1989.
- Lichtenstein, M. and Srull, T. K. 'Processing objectives as a determinant of the relationship between recall and judgment', *Journal of Experimental Social Psychology*, **23**(2) (1987), 93–118.

- Loewenstein, G. 'Anticipation and the valuation of delayed consumption', *Economic Journal*, **97** (1987), 666–684.
- Loewenstein, G. and Prelec, D. 'Negative time preference', *American Economic Review: Papers and Proceedings*, **82**(2) (1991), 347–352.
- Loewenstein, G. and Prelec, D. 'Preferences for sequences of outcomes', *Psychological Review*, **100** (1993), 91–108.
- Nowlis, S. M. and Simonson, I. 'Attribute-task compatibility as a determinant of consumer preference reversals', *Journal of Marketing Research*, **34** (1997), 205–218.
- Redelmeier, D. A. and Kahneman, D. 'Patients memories of painful medical treatments — real-time and retrospective evaluations of two minimally invasive procedures', *Pain*, **66** (1996), 3–8.
- Ross, W. T. Jr and Simonson, I. 'Evaluations of pairs of experiences: A preference for happy endings', *Journal of Behavioral Decision Making*, **4** (1991), 273–282.
- Varey, C. A. and Kahneman, D. 'Experiences extended across time: Evaluation of moments and episodes', *Journal of Behavioral Decision Making*, **5** (1992), 169–185.
- Werthenbroch, K. and Carmon, Z. 'Dynamic preference maintenance', *Marketing Letters*, **8**(1) (1997), 145–152.

Authors' biographies:

Dan Ariely is an Assistant Professor at the Massachusetts Institute of Technology, apprehensively starting his first winter in New England. His expertise and ability in decision making has culminated with his recent marriage to Sumi.

Ziv Carmon is an Associate Professor at Duke University's Fuqua School of Business. He is very happy for Dan and Sumi. For more information about Ziv, see his web page, currently at www.duke.edu/~zc.