Journal of Behavioral Decision Making
J. Behav. Dec. Making, **19**: 191–211 (2006)
Published online 21 June 2006 in Wiley InterScience
(www.interscience.wiley.com). **DOI**: 10.1002/bdm.516

Hedonic versus Informational Evaluations: Task Dependent Preferences for Sequences of Outcomes

GAL ZAUBERMAN[1]*, KRISTIN DIEHL[2] and DAN ARIELY[3]
[1]*University of North Carolina—Chapel Hill, USA*
[2]*University of Southern California, USA*
[3]*Massachusetts Institute of Technology, USA*

ABSTRACT

This work examines how people form evaluations of extended experiences that vary in valence and intensity. It is documented that when people retrospectively evaluate such experiences, not all information is weighted equally. Some prior research demonstrates that earlier parts are weighted more than later parts, while other research shows the opposite. In this paper we suggest that differences in evaluation tasks shift the focus to different aspects of the experience, causing individuals to be differentially influenced by earlier or later parts of the experience. We show that ratings of feelings (hedonic evaluation tasks) lead to stronger preferences for improving experiences than do evaluative judgments (informational evaluation tasks), suggesting that later aspects of the experience are weighted more heavily in affective tasks. In addition, we investigate other evaluation tasks, demonstrating that whether the task is descriptive or predictive and whether the target of the evaluation is the source of the experience or the experience itself also alter the weight given to different parts of the experience. Our studies demonstrate systematic shifts driven by these different evaluation task, revealing changes in overall evaluations as well as changes in the underlying weighting of key characteristics of the experience (i.e., start, end, and trend). Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS order effects; hedonic evaluation; sequences; time; evaluative judgments; weighting; experiences

People are commonly exposed to events that unfold over time and that vary in valence and intensity over the course of their duration. It is well documented that when retrospectively evaluating such experiences, not all information that is unveiled over the course of the experience is weighted equally (e.g., Anderson 1981; Fredrickson & Kahneman, 1993). Instead, certain parts of an experience are more important than others in

* Correspondence to: Gal Zauberman, McColl Building, CB# 3490, The University of North Carolina, Chapel-Hill, NC 27599-3490, USA. E-mail: zauberman@unc.edu

influencing overall evaluations, namely the beginning, the end and the overall trend of the experience. In this paper we propose one factor that can affect the weight placed on different parts of the experience. Specifically, we suggest that different evaluation tasks render different aspects more or less relevant when evaluating sequences, and thus altering the weighting of these aspects.

Researchers studying retrospective evaluations of extended experiences showed that individuals base their overall evaluations on the peak and final intensity of the experience (e.g., Fredrickson & Kahneman, 1993), as well as the trend of the experience, a dimension taking into account the progression of information (Ariely & Zauberman, 2003; Loewenstein & Prelec 1993). Evaluations based on these aspects favor improving over declining experiences. However, researchers studying information-integration (e.g., Anderson, 1981), found that the absolute serial position of each piece of information, not its relative position, influences overall evaluations. Findings in this paradigm suggest that information with an early position is often more influential, producing primacy effects, although under certain conditions recency was also observed (e.g., Hastie & Park, 1986; Hogarth & Einhorn, 1992; Jones & Goethals, 1972).

Whether the earlier or the later parts of an experience have greater influence on overall evaluations could be driven merely by divergent methodologies, different domains, or different dependent measures. However, we suggest that different *evaluation tasks* can produce systematic rather than merely contextual differences. As such, these results are not merely methodological artifacts, but are in fact theoretically informative. The overall theoretical perspective we apply to this question is based on prior research that, in other contexts, has shown that the evaluation task decision makers face determines which pieces of information seem more task relevant and thus are weighted more heavily (e.g., Abele & Petzold, 1998; Fischer, Carmon, Ariely, & Zauberman, 1999; Mussweiler 2003; Trope & Lieberman, 2000).

THEORETICAL BACKGROUND

We next compare evaluation tasks used in research on retrospective evaluations of experiences (e.g., Fredrickson & Kahneman, 1993) and research on order effects in information integration (e.g., Anderson, 1981; Jones & Goethals, 1972) in order to illustrate differences relevant to our research question. Note that our intention is not to provide a comprehensive review of this vast body of work, nor is it to integrate these multifaceted research streams. Rather, the specific evaluation tasks used in each of these areas serve as existing illustrations of our central idea: that certain aspects of the experience, such as the beginning, the end, and the rate of change, are weighted differently depending on the key objective of the focal evaluation task.

Evaluation tasks

In this paper we examine factors that influence the extent to which overall evaluations are affected by the trend and end of an episode compared to its beginning, and thus leading to preferences for improving experiences. Prior research has found evidence for both effects. Since prior work used very different evaluation tasks, these experimental operationalizations may have contributed in a systematic way to seemingly contradictory findings of primacy versus recency.

Findings by Trope and Lieberman (2000) support the idea that aspects more central to the evaluation task are weighted more heavily when forming evaluations. In particular, in their Study 5, participants watching a movie with the goal of learning about a topic placed greater weight on the informativeness of that movie, while those watching a movie in order to get into a good mood weighted the funniness of the movie more heavily in their evaluations.

Further support for the idea that task relevant aspects are weighted more heavily is provided by Fischer et al. (1999). These authors demonstrate that in response tasks in which the underlying evaluation

task is to differentiate between alternatives rather than to equate alternatives, the prominent feature (i.e., the most important attribute), is weighted more heavily. Similarly, Mussweiler (2003) examined assimilation and contrast effects, and showed that dissimilarity testing leads to greater emphasis on differentiating evidence, resulting in contrast effects, while similarity testing produces assimilation effects.

Note that the task's objective does not necessarily need to be explicitly stated (e.g., Trope & Lieberman, 2000), but can be inferred by meta-informational cues such as the format in which information is presented (Abele & Petzold, 1998). When individuals infer the task to be one of discrimination *between categories*, for example, category information is weighted more heavily and evaluations of individual category members are assimilated to the category average. When the task is to differentiate *within a category*, individualizing information is weighted more heavily.

Combined, the findings reviewed in this section support the idea that the evaluation task can influence which aspects of information are weighted more heavily. In a similar way, we propose that whether the trend and end of an experience or its start are more influential may also depend on the specific evaluation tasks individuals engages in. Next, we will review different tasks used in research on retrospective evaluations of experiences and on information integration to illustrate how these research streams differed in the evaluation tasks they employed.

Evaluation tasks used in research on retrospective evaluations of experiences

In a relatively recent stream of work, researchers have sought to understand how people retrospectively evaluate experiences that unfold over time (e.g., Frederickson & Kahneman, 1993). The vast majority of studies in this research stream asked respondents to rate the overall pleasure and pain across different types of experiences, although some studies also asked for choice or willingness to pay responses. For example, Ross and Simonson (1991) asked participants to rate with which of two computer games ''would you be happier?'' (Experiment 1 and 2), as well as ''How attractive is this game to you?'' and ''How attractive would others find this game to be?'' (Experiment 3). Diehl and Zauberman (2005) had participants search through lists of hotels or apartments and then asked ''How satisfied were you with the overall hotel [apartment] search experience?''. Participants in Ariely and Zauberman's (2003) studies viewed the progress of a firm's service quality and were then asked ''Looking back at the whole service experience, how satisfied or unsatisfied are you?''

For negative experiences, Frederickson and Kahneman (1993) asked participants ''Overall, how much displeasure or discomfort did you experience during the film you just saw?'. In other work, patients rated the ''total amount of pain experienced'' while undergoing colonoscopy and lithotripsy (Redelmeier & Kahneman, 1996), or in a milder form, indicated which of two trials of having their hand submerged in cold water ''caused the greater overall discomfort?'' (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). Studies using aversive sounds asked participants, ''Looking back at the sound you just heard, overall how annoyed were you?'' (Ariely & Zauberman, 2000) or ''Rate the overall pleasantness or unpleasantness of listening to that last sound'' (Schreiber & Kahneman, 2000).

Results across many of these studies show that overall retrospective evaluations are most strongly affected by the rate of change (e.g., Ariely & Zauberman, 2003; Loewenstein & Prelec, 1993), the final intensity, and the maximum intensity (e.g., Schreiber & Kahneman, 2000). Consequently, when individuals are asked to provide overall ratings of their feelings they are more likely to prefer a non-enjoyable experience followed by an enjoyable experience (i.e., improving pattern) than an enjoyable experience followed by a non-enjoyable experience (i.e., declining pattern), a result that could be interpreted as a recency effect. In fact, preferences for improving patterns can be so strong that individuals sometimes even prefer to receive more overall pain, as long as the experience ends with a reduced level of pain (Ariely, 1998; Redelmeier & Kahneman, 1996).

Evaluation tasks used in research on information integration

Extensive research on information integration and order effects has examined questions relating to how people form evaluations of extended episodes in a range of domains (e.g., Anderson, 1981; Asch, 1946). However, particularly the ''first phase of this research program was largely devoted to one experimental situation, namely, the personality adjective task'' (Anderson, 1981, p. 99). In these studies participants were asked to form person perceptions based on trait information they received, while presentation order of that information was systematically manipulated. Asch (1946), for example, varied the order in which six adjectives describing a person were heard (Experiment 6). Participants were then asked to form an impression of that person and subsequently indicated which of 18 personality qualities were in accordance with their impression. Similarly, participants in Lichtenstein and Srull's (1987) studies assessed the likeability of a target person after reading a set of statements describing that person. Hastie and Park (1986) asked for ratings of friendliness, likeability, and intelligence of a person described as performing different everyday activities. Participants in studies by Jones, Rock, Shaver, Goethals, & Ward (1968) also provided intelligence assessments: After observing a target person solving a set of multiple choice questions with the frequency of correct answers either increasing or decreasing, participants rated the target's intelligence on a scale ranging from ''well below average Duke student'' to ''well above average Duke student.''

In addition to order effects in person perception, changing the order of information was also shown to alter its persuasive impact. In a hypothetical courtroom setting, Miller and Campbell (1959) investigated the degree of attitude shift induced by differently ordered arguments asking participants to judge whether the ''defendant is more responsible'' or the ''plaintiff is more responsible.'' Relatedly, Hogarth and Einhorn (1992) studied the persuasiveness of different information orderings by asking participants to judge the likelihood that a specific factor was responsible for a given outcome, such as a broken speaker, etc.

Lastly, Jones, Goethals, Kennington, and Severance (1972) investigated how recall for the frequency of actual events was affected by ordering. Participants in their studies experienced different events such as aircraft radar signals (Experiment 1) or aces revealed from a card deck (Experiments 2 and 3) that occurred in prearranged increasing or decreasing frequencies. Participants were later asked to recall the number of target events that had occurred as well as to remember the sequence in which target and non-target events had taken place.

In general, this stream of research has studied how serial position (i.e., information sequence A-B compared to B-A) affects overall evaluations, revealing substantial evidence for primacy effects, yet with some exceptions (e.g., Hastie & Park, 1986; Hogarth & Einhorn, 1992). The evidence for primacy suggests that declining patterns should be preferred over improving patterns due to greater influence of earlier parts of the experience.

EVALUATION TASKS IN SEQUENCES OF OUTCOMES

Although research on order effects in information integration commonly found primacy effects, research on retrospective evaluations of experiences found strong influences of the end and the trend of the experience. A simple explanation for these differences may be the design of outcome sequences used in these different streams. Experiments studying order effects predominantly manipulated information and serial position orthogonally (e.g., Anderson, 1981), masking any interactions between the valence and position of information throughout the sequence. Thus, in contrast to the research on retrospective evaluations of experiences that focuses on the Gestalt of the sequence, research on information integration assumes independence between pieces information (c.f., Jones et al., 1968). Jones and Goethals (1972) argued that the methodology used by Anderson (e.g., Anderson, 1981) actually ignores the impact of the context in which events unfold; they suggested that researchers should focus on Asch's (1946) original interpretation of order effects as a function of the relationship between events. These differences in experimental procedures might be responsible for the inconsistent findings across research streams.

In addition to the differences in the tasks, the stimuli employed by these different research streams also varied greatly. Given the significant differences in methodologies, our experiments hold constant the stimuli as well as the experience respondents go through, ensuring that differences are truly systematic and not methodological artifacts. We suggest, that above and beyond methodological differences, differences in *evaluation tasks* may explain why in some instances the beginning of a series is weighted more heavily, while in others the end and final change have greater impact.

Evaluation tasks: evaluative judgments versus ratings of feelings

Our review of evaluation tasks used in the research on retrospective evaluations of experiences and on information integration revealed differences on a number of dimensions. However, one especially prominent dimension seems to be whether the task asked for evaluative judgments or ratings of feelings, a distinction that has shown to be important in other contexts (e.g., Edell & Burke, 1987, Millar & Tesser, 1986). Of particular relevance, Abelson, Kinder, Peters, & Fiske (1982) established that overall preferences for a person are independently predicted by affective ratings of the feelings that person evokes and by evaluative judgments of the person's personality traits. These authors point out that research in information integration theory has mostly asked for ''evaluative implications of trait combinations'' (p. 619). As our review indicates, such information integration tasks asked participants to form an *evaluative judgment* of a person (e.g., assessing intelligence, friendliness, etc.), at times even providing incentives for being accurate (e.g., Experiment 1, Jones et al., 1968). Research on retrospective evaluations of extended experiences, however, predominantly asked participants to *rate the feelings* generated by a given experience (e.g., rating happiness, displeasure, etc.). As such, the two research streams we examined seem to have tapped into distinct evaluation tasks. We therefore suggest that whether the task purpose is *hedonic* assessment of feelings or *informational* evaluative judgment could change whether earlier or later parts of an experience are perceived to be more relevant to the evaluation task and are consequently weighted more heavily.

When forming evaluative judgments in information integration tasks, decision makers are said to form a hypothesis and then integrate the information in the context of that hypothesis (e.g., Hogarth & Einhorn, 1992; Jones & Goethals, 1972). This anchoring and (insufficient) adjustment type of processing causes the initial part of the experience to have greater impact, because it is the basis for understanding subsequent information. In forming such informational evaluative judgments, people may therefore weight the task-consistent initial information more heavily and will rely less on seemingly unrelated aspects, such as the information contained in the trend and end. As a result, decreasing rather than increasing experiences should lead to more positive evaluations (primacy effect).

In contrast, ratings of the feelings evoked by an experience are more likely to be affected by anticipation and thus place greater focus on how experiences evolve over time (Ariely & Carmon, 2000; Ariely & Zauberman, 2003; Loewenstein, 1987; Loewenstein & Prelec, 1993). Indeed, work by Hsee and colleagues (Hsee & Abelson, 1991, Hsee, Abelson & Salovey, 1991) suggests that the experience's ''velocity'' (rate of change) rather than its ''position'' (outcome level) has a greater impact on overall evaluations when people focus on the affective (''consummatory'') aspects of an experience. Therefore, hedonic evaluations that ask for ratings of feelings should be relatively less sensitive to the initial information and more sensitive to the sequence's trend and the final information (recency effect).

H1: Improving experiences will be evaluated as more positive compared to declining experiences when individuals engage in *hedonic* rather than *informational* evaluation tasks.

H2a: The *trend* (rate of change) and the *end* of the experience will have greater impact on final evaluations when engaging in hedonic evaluation tasks rather than informational evaluation tasks.

H2b: The *start* of the experience will have greater impact on final evaluations when engaging in informational evaluation tasks rather than hedonic evaluation tasks.

Descriptive versus predictive tasks

We suggest that in informational evaluation tasks, the earlier parts of a sequence of information are weighted more heavily, while hedonic tasks place greater weight on the end of a sequence. This raises the question of whether there are other evaluation tasks that would change the sensitivity to the beginning versus the end of a sequence and thus would moderate the effect of hedonic versus informational tasks. One task that would render the latter parts of a sequence more relevant would be to ask for predictions rather than for descriptions of the past. We expect that individuals making predictions will shift their focus towards changes in the experience and end intensity, since these aspects of an experience seem more relevant to the immediate future. In hedonic tasks, changes in the experience, particularly at the end, are attended to even when evaluating the past; therefore, a predictive task is not expected to significantly alter preferences for improving experiences (Novemsky & Ratner, 2003). In informational tasks, however, making predictions should significantly alter evaluations of declining versus improving experiences by rendering the latter parts of the experience more relevant. We will test the moderating effects of predictive versus descriptive evaluation tasks on the relationship between hedonic and informational evaluations in experiments 1 and 2.

EXPERIMENT 1

Experiment 1 examines whether hedonic compared to informational evaluation tasks lead to different evaluations of improving versus declining experiences (H1). We also test the extent to which these evaluations change as a function of whether they focus on what did happen in the past (descriptive) versus what will happen in the future (predictive). Note that we will not simply show that different elicitation methods lead to different responses. Rather we want to test whether different evaluation tasks change respondents' weight for the beginning versus the end of a sequence, reflected in differential evaluations of improving over declining experiences.

Method

Participants and design

Forty undergraduate students at an East Coast university were recruited through advertising and paid $10 for their participation. The experimental design was a 2 (hedonic/informational) $\times$ 2 (predictive/descriptive) $\times$ 2 (increasing/decreasing) $\times$ 5 (replicates) mixed design. Whether participants engaged in a hedonic or an informational evaluation task was manipulated between-subjects. Whether the evaluation task was predictive or descriptive and whether patterns were increasing or decreasing were manipulated within-subjects. All participants were presented with an identical set of patterns across conditions. Participants evaluated five increasing and five decreasing pattern, in one of eight predetermined random orders.

Stimuli

As in the Jones et al. (1968) study, the stimulus domain was test performance and patterns were manipulated by varying the frequency of correct and incorrect responses. For increasing patterns, correct answers appeared more frequently over time; decreasing patterns were a perfect reflection where correct answers

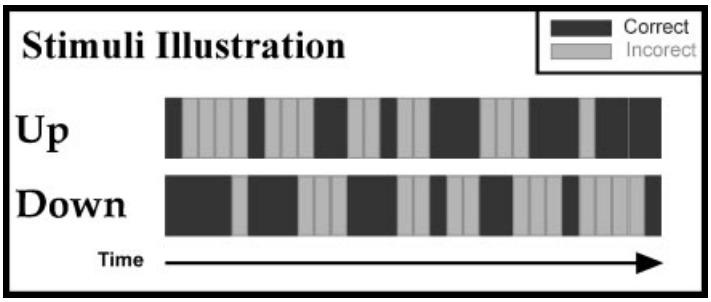

Figure 1. Experiment 1. Graphical illustration of the stimuli, demonstrating increasing (Up) and decreasing (Down) frequencies of correct answers.

occurred less frequently over time (see Figure 1). As each trial progressed, correct and incorrect responses were plotted cumulatively.

Procedure

Participants were introduced to the computer environment and were given task instructions. Participants responded to 10 trials, 5 trials in each of the two pattern conditions. On each of the 10 trials, participants were presented with the performance of a job candidate on 30 questions said to reflect their performance on the first half of the ''M.A.Y.A. Placement Test.'' A blue bar indicated a correct response; a red bar signified an error. At the end of each trial, all participants responded to two dependent measures (descriptive and predictive evaluations).[1] The experiment took approximately 30 minutes to complete.

Evaluation tasks and dependent measures

Participants had to respond to either a hedonic or an informational evaluation task. In the informational condition, participants were asked to evaluate the performance of potential job candidates on the placement test. In the hedonic condition, participants were asked to rate the job candidates' satisfaction with their performance on this placement test. In each condition, participants provided two types of evaluations (descriptive and predictive) on a scale anchored at 0 and 100, with higher numbers indicating more positive evaluations. For the exact wording see Table 1.

RESULTS

A repeated-measures ANOVA revealed a significant main effect of pattern, with increasing patterns being evaluated more positively than decreasing patterns ($M_{IN} = 56.42$ vs. $M_{DE} = 53.05$), $F(1, 38) = 6.9$, $p < 0.02$. Whether the task was predictive compared to descriptive had a small positive and marginally significant main effect ($M_{PRE} = 55.57$ vs. $M_{DES} = 53.90$), $F(1, 38) = 3.46$, $p < 0.08$. More importantly, the interaction between hedonic/informational evaluation tasks and increasing/decreasing patterns was significant ($F(1, 38) = 19.56$, $p < 0.0001$). In addition, the interaction between the hedonic/informational

[1]Following Jones et al. (1968), they also provided an ability evaluation that was common across conditions. This measure was not affected by the manipulations and will not be discussed further.

Table 1. Experiment 1: dependent measures for each evaluation task combination

	Evaluation task	
	Hedonic	Informational
Evaluation task		
Descriptive	Given the information you have, how satisfied was this person after completing the first part of the test?	Given the information you have, how many questions (out of 60) did this job candidate answer correctly in the first part of the test?
Predictive	Given the information you have, how satisfied would you predict that this job candidate would be with their performance on the next series of 60 questions?	Given the information you have, how many questions would you predict that this job candidate would answer correctly out of the next series of 60 questions?

and descriptive/predictive evaluations tasks was also significant ($F(1, 38) = 6.69$, $p < 0.02$). Although the three-way interaction was not significant ($F(1, 38) < 1$), in line with our prediction, hedonic and informational tasks differed significantly depending on whether descriptive or predictive evaluations were solicited. We will report findings for the descriptive evaluation task first and then examine how predictive evaluation tasks moderate the observed effects.

Descriptive evaluations

Whether increasing or decreasing patterns were evaluated had a small positive and marginally significant main effect on descriptive evaluations ($M_{IN} = 55.02$ vs. $M_{DE} = 52.79$), $F(1, 38) = 2.9$, $p < 0.1$. In addition, the interaction of increasing/decreasing pattern by hedonic/informational evaluation task was significant ($F(1, 38) = 23.13$, $p < 0.0001$). Participants engaged in hedonic evaluation tasks rated increasing patterns significantly better ($M = 57.69$) than decreasing patterns ($M = 49.24$), $F(1, 19) = 21.57$, $p < 0.001$. The reverse was true for participants engaging in an informational evaluation task, with respondents rating the decreasing pattern ($M = 56.34$) more positive than the increasing pattern ($M = 52.34$), $F(1, 19) = 4.72$, $p < 0.05$; see Figure 2a.

Predictive evaluations

For the predictive evaluation task the main effect of increasing versus decreasing patterns was positive and significant ($M_{IN} = 57.83$, $M_{DE} = 53.31$), $F(1, 38) = 6.86$, $p < 0.02$. In addition, the interaction of increasing/decreasing patterns by hedonic/informational evaluation task was also significant ($F(1, 38) = 9.43$, $p < 0.01$). As before, participants engaging in a hedonic evaluation task rated the increasing patterns significantly more positively ($M = 62.35$) than the decreasing pattern ($M = 52.53$), $F(1, 19) = 14.03$, $p < 0.002$. However, unlike our findings for descriptive tasks, those engaging in an informational evaluation task did not show any difference in their predictions following increasing compared to decreasing patterns ($M_{IN} = 53.30$, vs. $M_{DE} = 54.08$), $F(1, 38) = 0.12$, $p > 0.7$; see Figure 2b.

Discussion

Experiment 1 demonstrates that when individuals engage in *descriptive hedonic* evaluations, improving experiences lead to more positive evaluations than declining experiences. When engaging in *descriptive informational* tasks, however, declining experiences lead to more positive evaluations. These findings

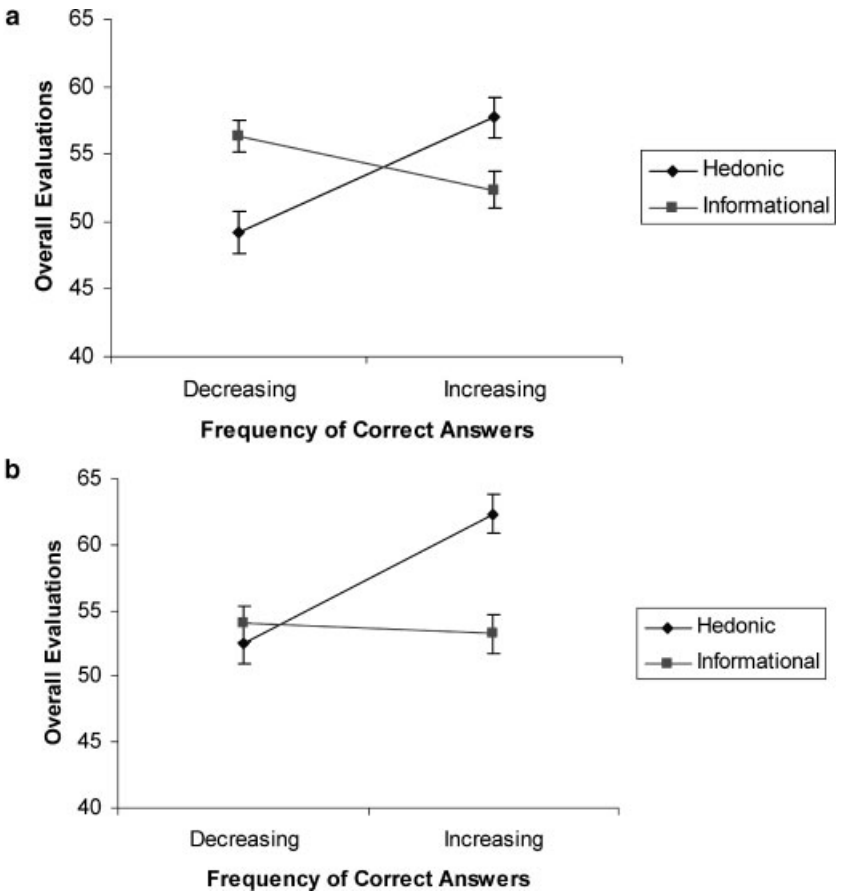

Figure 2. Experiment 1. Overall evaluations as a function of decreasing versus increasing frequency of correct answers and evaluation task for descriptive (panel a) a predictive (panel b) evaluations.

suggest that under hedonic evaluation tasks the rate of change and end of the experience are weighted more heavily, while informational tasks place relatively greater emphasis on the start of an experience, favoring declining experiences (H1). However, even when engaging in informational evaluation tasks, the emphasis on the change and end of the experience can be strengthened. If the task is to predict the future, informational evaluation tasks no longer lead to more positive evaluations of declining experiences. Consistent with our predictions, engaging in predictive compared to descriptive tasks strengthened the main effect of patterns and weakened the interaction of patterns by hedonic/informational evaluation tasks, that is, increasing patterns lead to more positive evaluations. These results demonstrate that for predictive evaluation tasks, individuals rely more on the change and the end of the experience across both hedonic and information tasks.

Note that in this mixed design, the three-way interaction was not significant. One reason may be that descriptive and predictive evaluation tasks were within-subject repeated measures, which may have weakened the observed differences between these tasks. Experiment 2 will isolate the hedonic/information and descriptive/predictive tasks in order to provide a stronger test of our hypotheses. In addition, a different set of stimuli will allow us to directly test whether different aspects of the experience are weighted more heavily under certain evaluation tasks (H2a and H2b).

EXPERIMENT 2

In Experiment 2, participants again evaluated identical sequences under different evaluation tasks. In order to separate descriptive and predictive evaluation tasks, this factor was manipulated between-subjects, as was whether the evaluation task was hedonic or informational in nature. To increase generalizability, we used a different stimulus domain and a different way to display the information (a line graph that presented the information noncumulatively). In this experiment, participants were asked to evaluate reject rates (i.e., the percentage of defective products of a production line in a computer chip factory). Note that, as such, increasing reject rates should lead to *less* positive hedonic evaluations. If our results replicate with this scenario, we can ensure that our previous findings were not due to participants mindlessly following the graphical depiction of the pattern, but rather that they encoded the resulting experience. Note that we will report the findings in a way that corresponds to the visual pattern of increasing or decreasing reject rates presented to participants. Thus, higher reject rates should decrease hedonic evaluations (i.e., satisfaction with situation), but should lead to more extreme informational evaluations (i.e., assessment of situation).

As before, we will examine the effect of hedonic versus informational evaluation tasks on overall evaluations of increasing and decreasing patterns (H1). However, in addition, a second level of analysis will allow us to examine the effects of evaluation tasks through the impact of specific experience characteristics (start, trend, and end) on overall evaluations (H2a and H2b). The latter analysis enables us to examine the underlying integration processes in more detail.

Method

Participants and design

One hundred undergraduate students at an East Coast university were recruited through advertising and paid $8 for their participation. Participants were randomly assigned to one of four experimental conditions. The experimental design was a 2 (hedonic/informational) $\times$ 2 (predictive/descriptive) $\times$ 2 (increasing/decreasing) $\times$ 5 (replicates) mixed design. Whether participants engaged in a hedonic or an informational evaluation task and whether they provided descriptive or predictive evaluations was manipulated between-subjects. Whether reject rates were increasing or decreasing was manipulated within-subjects. Participants evaluated five increasing and five decreasing patterns in a random order.

Stimuli and procedure

The general procedure was similar to that of Experiment 1. Participants observed the level of defective products (reject rate) at 10 different factories, where each factory is a separate trial, presented in a random order for each participant. The vertical axis of the display was marked 0% reject rate at the bottom and 10% reject rate on the top. The horizontal axis was marked January 1 on the left and December 31 on the right. At the beginning of each trial, participants saw an empty grid. Over the course of 45 seconds weekly performance was plotted for the period between January 1 and December 31. Unlike Experiment 1, each observation point (i.e., a week) appeared non-cumulative, one bit of information at a time. The experiment took approximately 20 minutes to complete.

Unlike in Experiment 1, increasing and decreasing patterns were not simple replicates of one another, but rather were designed with different starting points, end points, and slopes (see Table 2). Using information gathered across all 10 patterns allows us to determine how key aspects of these patterns are weighted across different evaluation tasks.

Table 2. Objective characteristics for each of the 10 patterns

Pattern parameters							
	Increasing				Decreasing		
Pattern	Start	End	Slope	Pattern	Start	End	Slope
A-B	0.5	3.5	1.5	B-A	3.5	0.5	−1.5
A-C	0.5	6.5	3.0	C-A	6.5	0.5	−3.0
B-C	3.5	6.5	1.5	C-B	6.5	3.5	−1.5
B-D	3.5	9.5	3.0	D-B	9.5	3.5	−3.0
C-D	6.5	9.5	1.5	D-C	9.5	6.5	−1.5

Dependent measures

Participants responded to one of four distinct measures corresponding to hedonic or informational evaluation tasks with either a descriptive or predictive focus (see Table 3 for exact wording). All responses were provided on a scale from 0 to 10. Under informational evaluation tasks, participants responded in terms of the reject rate (i.e., in percentages ranging from 0% to a maximum of 10%). Under hedonic evaluation tasks, they indicated their satisfaction ranging from 0 (''not satisfied'') to 10 (''very satisfied''). It is important to note again that, unlike Experiment 1, there was not necessarily a directional correspondence between the graphically depicted changes and the experiential evaluation. In the current experiment, increasing reject rates correspond to higher values on the informational response scale (i.e., higher reject rates), but lower values on the hedonic response scale (i.e., lower satisfaction). As such, hedonic evaluations of decreasing reject rates should lead to improving experiences if participants actually encoded the experience instead of merely responding to the graphical pattern. Note that H1 is predicated on individuals' *experiences*. Translating the proposition of H1 to the current set-up, we therefore predict that for descriptive evaluations, decreasing reject rates (i.e., improving experiences) compared to increasing reject rates (i.e., declining experiences) will have a stronger effect on overall evaluations when facing hedonic rather than informational tasks. Further, we expect that given predictive evaluation tasks, those engaging in informational evaluations will increase their reliance on later parts for the experience, leading to more extreme evaluations of increasing reject rates.

Table 3. Experiment 2: dependent measures for each evaluation task combination

	Evaluation task	
	Hedonic	Informational
Evaluation task		
Descriptive	Looking back at the information you just observed, how satisfied are you with the average reject rate of production line A over the past year?	Looking back at the information you just observed, what was the average reject rate (in %) of production line A over the past year?
Predictive	From the information you just observed, how satisfied do you predict that you will be with the reject rate of production line A over the next year?	From the information you just observed, what do you predict will be the average reject rate (in %) of production line A over the next year?

Results

A repeated-measures ANOVA revealed a significant three-way interaction between hedonic/informational evaluation tasks, descriptive/predictive evaluation tasks, and increasing/decreasing reject rates ($F(1, 96) = 36.78$, $p < 0.0001$). We first analyze the effects on descriptive evaluation tasks and then examine how predictive tasks moderate these findings.

Descriptive evaluations

For descriptive evaluation tasks, the interaction of hedonic/informational tasks and increasing/decreasing reject rates was significant ($F(1, 48) = 83.35$, $p < 0.0001$). Participants engaging in a hedonic evaluation task evaluated their experience of decreasing reject rates as significantly more positive ($M = 61.68$) than that of increasing reject rates ($M = 36.92$), $F(1, 24) = 84.67$, $p < 0.0001$. Those engaging in informational evaluations reported similar average reject rates for decreasing ($M = 49.88$) and increasing reject rates ($M = 50.92$), $F(1, 24) = 1.45$, $p = 0.24$; see Figure 3a.

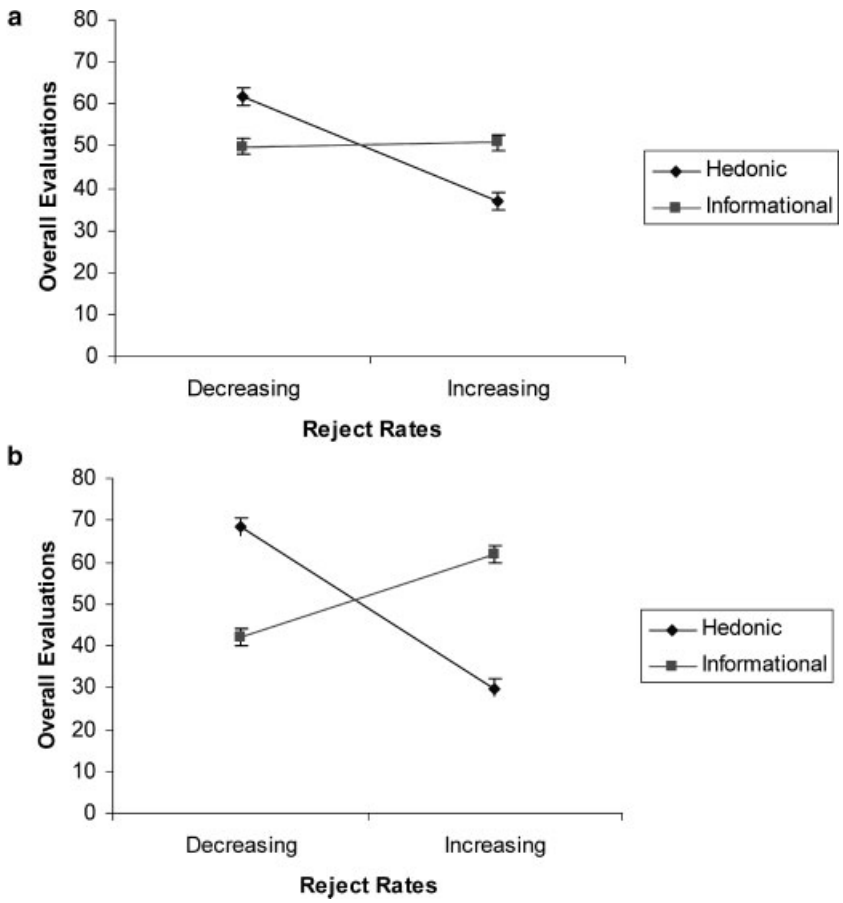

Figure 3. Experiment 2. Overall evaluations as a function of decreasing versus increasing reject rates and evaluation task for descriptive (panel a) a predictive (panel b) evaluations.

Predictive evaluations

For predictive evaluation tasks, the interaction of hedonic/informational evaluations by increasing/decreasing patterns was significant ($F(1,48) = 164.80$, $p < 0.0001$). Participants engaging in a hedonic evaluation task, predicted their future satisfaction to be significantly more positive when experiencing decreasing reject rates ($M = 68.48$) than increasing reject rates ($M = 29.88$), $F(1,24) = 124.41$, $p < 0.0001$. Those engaging in an informational evaluation task, predicted significantly higher future reject rate when being exposed to increasing ($M = 61.68$) compared to decreasing reject rates ($M = 42.08$), $F(1,24) = 44.79$, $p < 0.0001$; see Figure 3b.

Note that the difference between hedonic and informational tasks is smaller for predictive than descriptive evaluations, although the graphical representation seems to indicate the opposite. This appearance is due to the nature of the stimuli, where increasing reject rates corresponds decreasing satisfaction. As such, smaller experiential differences between hedonic and informational evaluations, as predicted by our framework, are actually manifested in higher informational evaluations but lower hedonic evaluations. Therefore, the larger gap between informational and hedonic evaluations for predictive versus descriptive evaluations found in the graphical illustration is fully consistent with our predictions.

Experience characteristics

An important contribution of Experiment 2 is its ability to explicitly test whether different characteristics of the experience (i.e., start, end, and trend) are differentially weighted as a function of the evaluation task. Considering our stimuli, with simple linear patterns, any one sequence possesses only two degrees of freedom. This makes it impossible to simultaneously estimate whether the evaluation task alters the weighting of all three key aspects of the experience (i.e., start, end and trend). However, in this experiment, the increasing and decreasing patterns were specifically designed with different starting points, end points, and slope (see Table 2). Using information gathered across these 10 patterns allows us to determine how the key aspects of these patterns are weighted across different evaluation tasks. To do this, we estimated three separate individual-level regression models for each subject, regressing final evaluations on the manifestations of one of the three characteristics of the experience across 10 trials. The resulting individual-level raw beta coefficients are used as within-subjects, dependent measures. Given the relationship between reject rates and hedonic as compared to informational response scales, beta coefficients for hedonic tasks will have a negative sign, while those for informational tasks will be positive. Since our predictions focus on the magnitude of the impact that experience characteristics exert, we present absolute rather than signed raw beta coefficients as a direct test of our hypotheses.

We test our predictions using a $2 \times 2 \times 3$ mixed-design ANOVA model. Hedonic/informational and descriptive/predictive evaluation tasks were between-subjects factors. Experience characteristic was a within-subject factor with three levels: start, end, and trend. As predicted, the three-way interaction between hedonic/informational tasks, descriptive/predictive tasks and experience characteristics was significant ($F(2,192) = 15.14$, $p < 0.001$), suggesting that aspects of the experience are weighted differently depending on the evaluation task. We will first analyze which aspects of the experience are weighted more heavily under descriptive evaluation tasks, and then analyze how predictive tasks moderate these findings.

Experience characteristics in descriptive evaluations

The interaction of hedonic/informational evaluation task by characteristics of the experience was significant ($F(2,96) = 105.85$, $p < 0.001$). As Figure 4a shows, and consistent with H2a, the end ($b_{hed} = 0.27$, $b_{info} = 0.21$, $F(1,48) = 41.83$, $p < 0.0001$) and the trend ($b_{hed} = 0.22$, $b_{info} = 0.03$, $F(1,48) = 113.75$, $p < 0.0001$) had a bigger impact on final evaluations for hedonic compared to informational evaluation tasks. The start of an experience, however had a stronger impact on final evaluations when the task was

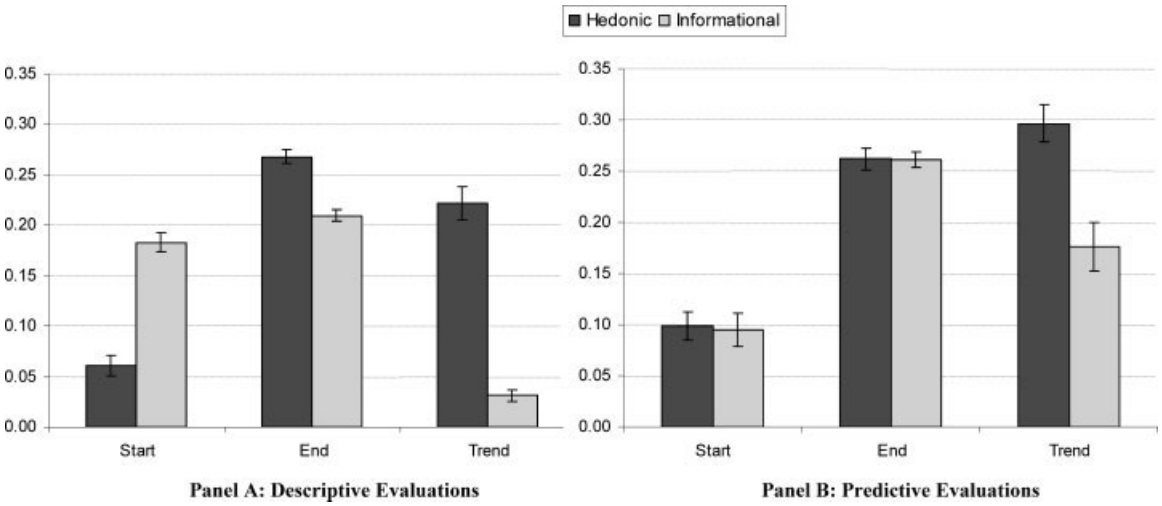

Figure 4. Experiment 2. Impact of experience characteristics on overall evaluations as a function of hedonic/informational evaluation tasks for descriptive (panel a) a predictive (panel b) evaluations.

informational rather than hedonic in nature ($b_{hed} = 0.06$, $b_{info} = 0.18$, $F(1, 48) = 78.40$, $p < 0.0001$), supporting H2b.

Experience characteristics in predictive evaluations

Recall that a predictive evaluation task should diminish differences between hedonic and informational tasks. The interaction of hedonic/informational evaluation task by experience characteristics was still significant ($F(1, 96) = 8.23$, $p < 0.001$), but was much smaller compared to the descriptive tasks. With a predictive task, we did not observe a differential effect of hedonic versus informational tasks on either the effect of the start ($b_{hed} = 0.10$, $b_{info} = 0.09$, $F(1, 48) = 0.03$, $p > 0.8$) or the end ($b_{hed} = 0.26$, $b_{info} = 0.26$, $F(1, 48) = 0.01$, $p > 0.9$); see Figure 4b. The trend of the experience remained more impactful for hedonic than informational tasks ($b_{hed} = 0.30$, $b_{info} = 0.18$, $F(1, 48) = 16.81$, $p < 0.001$), but this difference was smaller than what we observed for descriptive tasks.

Note that across all four evaluation task combinations, the end of an experience ($b = 0.25$) had a bigger impact on overall evaluations than the start ($b = 0.11$) or the trend ($b = 0.18$). Further, the effect of end on final evaluations ($b_{des\ hed} = 0.27$, $b_{des\ info} = 0.21$, $b_{pred\ hed} = 0.26$, $b_{pre\ info} = 0.26$) was less affected by the specific combination of evaluation tasks ($F(3, 96) = 12.09$, $\omega^2 = 0.25$) than were the trend ($F(3, 96) = 42.26$, $\omega^2 = 0.553$) or the start ($F(3, 96) = 16.88$, $\omega^2 = 0.323$).[2]

Discussion

Experiment 2 provides further evidence that improving experiences lead to more extreme overall evaluations when the evaluation task is hedonic rather than informational. We find support for this hypothesis even though in the case of hedonic evaluation tasks, the graphical representation of the pattern (increasing/decreasing reject rates) and the associated experience (declining/improving experiences) lack directional

[2]The ω^2 effect size values are partial ω^2 (Keren & Lewis, 1979), excluding variance due to ANOVA terms unrelated to the tested effect. Partial $\omega^2 = \sigma^2\text{Effect}/(\sigma^2\text{Effect} + \sigma^2\text{Error})$.

correspondence, which makes for a strong test of our predictions. We further replicate our findings that, for informational tasks, asking individuals for predictions attenuate differences between improving and declining experiences and, in the current situation, leads to more positive evaluations for improving experiences.

Because of the lack of directional correspondence between the graphically depicted profiles and the experiential evaluation, comparing results from Experiment 1 to those of Experiment 2 is not entirely straightforward. Focusing solely on the effects for the informational evaluation task, Experiment 1 showed that for descriptive tasks, decreasing frequencies of correct answers led to more positive evaluations, and predictive tasks shifted the weight toward later parts of the experience, attenuating these preferences. In Experiment 2, however, the descriptive informational task led to similar evaluations for decreasing and increasing reject rates (rather than a true advantage for the increasing pattern). For predictive informational tasks, we again found a shift in the weight toward later parts of the experience, leading to higher predicted assessments when reject rates were increasing. Thus, across these two studies we find that, for informational evaluation, predictive tasks shift the weight from the beginning to latter parts of the experience, thus changing the effect of increasing and decreasing patterns. However, the result of these changes in absolute terms, that is, whether decreasing or increasing patterns lead to higher overall evaluations, is a matter of calibration.

Experiment 2 also allows us to observe which aspects of the experience have greater influence on final evaluations. In line with our predictions, process level evidence from this study suggests that the trend and the end of an experience are relatively more influential under hedonic evaluation tasks, while the start has greater impact under informational evaluation tasks. In informational tasks we further observe that, when asked to evaluate what will happen, the trend and end of an experience gain greater influence on final evaluations.

These results support our idea that, although individuals rely less on the trend and the end of the experience when providing descriptive, informational evaluations, they are able to use these aspects of the experience in tasks where these aspects seem more relevant. Whether the task was descriptive or predictive had the biggest effect for informational evaluation tasks. Under hedonic evaluation tasks individuals rely on the trend and end of an experience regardless of whether the task is descriptive or predictive. This raises the question whether there are other evaluation tasks that would change the sensitivity to the beginning versus the end of a sequence even for hedonic tasks. Experiment 3 examines whether overall hedonic evaluations can be affected by the start of the experience if the start of the experience becomes more relevant to the task. Such a finding would further support our argument that greater reliance on the beginning (primacy) versus the end (recency) of an experience is indeed caused by the perceived relevance of that aspect to the task at hand.

EXPERIMENT 3

Experiments 1 and 2 examined the effects of hedonic versus informational tasks and the moderating effect of descriptive versus predictive tasks. Experiment 3 extends our investigation to an additional evaluation task. Research on information integration differed from research on retrospective evaluations not only in terms of the factual versus affective nature of the task, but also in terms of the focus of the assessment. In many information integration tasks the focus was on a specific object or person, while hedonic evaluations of extended experiences focused on the experience itself. This distinction would suggest that the focus of the evaluation may also affect whether earlier or later aspects of the experience are more influential. In line with this observation, we manipulate whether participants are asked about their satisfaction with a *service provider* (source) or a *service experience* (experience). The task of evaluating the source should heighten the impact of earlier parts of the experience even for hedonic evaluation tasks, similar to the effects found for person perception. The task of evaluating the service experience rather than the service provider, however, should rely more heavily on the trend and the end of the experience. In Experiment 3 we will focus solely on descriptive

hedonic tasks, asking participants to indicate their level of satisfaction given a certain evolution of service quality. Unlike Experiment 2, in this experiment the graphical representation of the pattern (increasing/decreasing service quality) corresponded to the development of the experience (improving/declining experience).

Method

Participants and design

Thirty members of an East Coast university community were recruited through advertising and paid $8 for their participation. Fifteen participants were randomly assigned to each of the two experimental conditions. The experiment followed a 2 (source/experience) × 2 (increasing/decreasing) × 5 (replicates) mixed design. Evaluation task was a two level between-subjects factor, manipulating whether participants were asked to indicate their satisfaction with the *service provider* or the *service experience*. Whether service quality was increasing or decreasing was a two-level within-subjects factor. Participants evaluated five increasing and five decreasing patterns for a total of 10 trials. The experiment took approximately 20 minutes.

Stimuli

Service quality was displayed as a line graph similarly to that used in experiment 2. The horizontal axis was marked January 1 on the left and December 31 on the right. The vertical axis was marked 0 at the bottom and 100 on top for the level of service quality. As in Experiment 2, participants started with an empty grid; performance was subsequently plotted non-cumulatively, one observation point (a week) at a time. Again, patterns of service quality were designed with varying start and end points, as well as different slopes (see Table 2), allowing us to extract the influence of start, end, and trend of the experience on final evaluations. Following exposure to each service experience, participants indicated their satisfaction with either the service provider or the service experience on an unnumbered scale anchored by ''not satisfied'' (0) and ''very satisfied'' (100).

Results

As expected, we found a significant two-way interaction of increasing/decreasing quality and the source/experience task ($F(1, 28) = 8.62$, $p < 0.01$). Results show that the difference between increasing and decreasing pattern was smaller for the source task ($M_{IN} = 55.27$, $M_{DE} = 39.53$, Diff = 15.73) than for the experience task ($M_{IN} = 64.67$, $M_{DE} = 31.20$, Diff = 33.45). These results indicated that, compared to focusing on the service experience, focusing on the source of the service reduces the extent to which improving experiences generate more positive ratings than declining experiences (Figure 5).

Experience characteristics

Using the same procedure as in Experiment 2, we regressed final evaluations on three characteristics of the experiences (i.e., start, end, and trend) across 10 trials. We estimated a 2 × 3 mixed ANOVA model with evaluation task as the between-subject factor and the individual-level raw beta coefficients of these pattern characteristic as the three level within-subject factor (start, end, and trend). Since there is no disconnect between the graphical representation and its corresponding experience or response scale, we simply compare the raw, signed regression coefficients across evaluation tasks (Figure 6).

The main effect of task ($b_{experience} = 0.45$, $b_{source} = 0.51$, $F(1, 28) = 5.28$, $p < 0.05$) and the main effect of experience characteristics ($b_{start} = 0.12$, $b_{trend} = 0.59$, $b_{end} = 0.73$, $F(2, 56) = 26.39$, $p < 0.0001$) were significant. More importantly, we found a two-way interaction between evaluation task and pattern characteristics

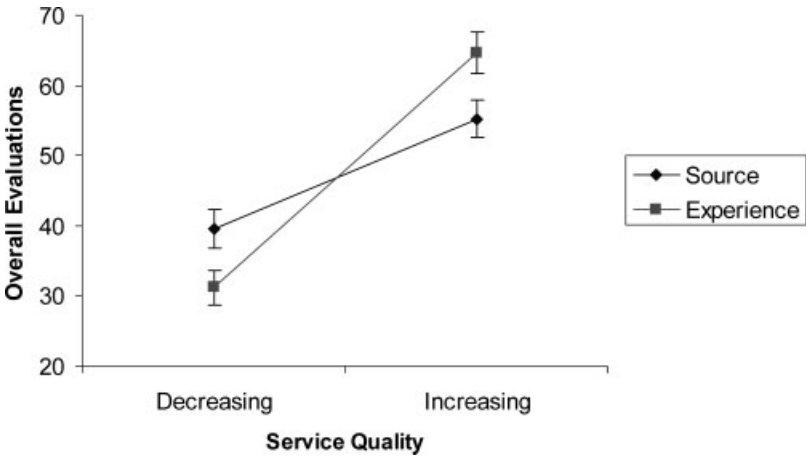

Figure 5. Experiment 3. Overall evaluations as a function of decreasing versus increasing service quality and evaluation task.

($F(2, 56) = 9.08$, $p < 0.001$), suggesting that aspects of the experience were weighted differently depending on the evaluation task. Recall that in this experiment all evaluations were hedonic, for which we generally found greater impact of the end. Still, when evaluating satisfaction with the service provider compared to satisfaction with the service experience we predicted greater impact of the start. Supporting our prediction, we found a significant effect of evaluation task on the impact the start has on overall evaluations ($F(1, 28) = 8.94$, $p < 0.01$) indicating that the beginning of an experience had significantly greater impact on final evaluations when the task was to evaluate satisfaction with the provider ($b = 0.29$, $t(14) = 6.29$, $p < 0.0001$) rather than satisfaction with the experience ($b = -0.04$, $t(14) = -0.47$, $p = 0.65$).

Further, for tasks focusing on the experience itself, the effect of trend and end on overall evaluations should be stronger compared to tasks focusing on the source of the experience. Indeed, we found that evaluation task has a significant effect on the impact trend had on final evaluations ($F(1, 28) = 9.07$, $p < 0.01$). Trend had a stronger impact when evaluating satisfaction with the service experience ($b = 0.8$, $t(14) = 6.56$,

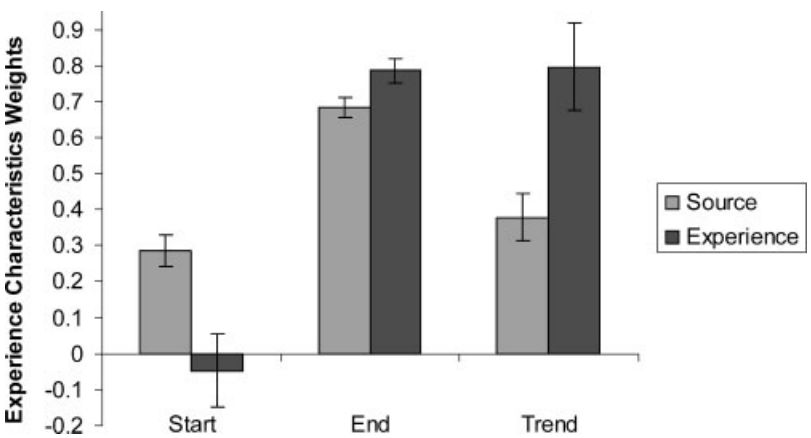

Figure 6. Experiment 3. Impact of experience characteristics on overall evaluations as a function of evaluation tasks.

$p < 0.0001$) compared to evaluating satisfaction with the service provider ($b = 0.39$, $t(14) = 5.66$, $p < 0.0001$). Similarly, we found a significant effect of task on the impact of end on final evaluations ($F(1, 28) = 5.67$, $p < 0.05$). The end had a stronger impact when evaluating the service experience ($b = 0.79$, $t(14) = 23.69$, $p < 0.0001$) as compared to the service provider ($b = 0.68$, $t(14) = 24.25$, $p < 0.0001$).

Note that in both evaluation tasks, the end of the experience had the largest effect on overall satisfaction. In addition, evaluation tasks had a smaller differential effect on the impact of end on overall evaluations ($F(1, 28) = 5.67$, $\omega^2 = 0.13$) than on the impact of start ($F(1, 28) = 8.94$, $\omega^2 = 0.21$) or trend ($F(1, 28) = 9.07$, $\omega^2 = 0.21$). We find that, compared to the start and the trend, the end of the experience has a more consistent effect on overall evaluations across different tasks.

Discussion

Experiment 3 included only hedonic evaluations and manipulated whether the evaluation task focused on the experience itself or the source of the experience. Consistent with the results of the first two experiments, this experiment demonstrates that overall evaluations of increasing versus decreasing patterns depend on the underlying evaluation task. The preference for improving compared to declining experiences is significantly reduced when evaluating satisfaction with the source of the experience compared to satisfaction with the experience itself. Analyses of the underlying impact of start, end and trend show again that the *differential* effects of evaluation tasks are driven mostly by the start and trend of the experience, while the end plays an important, yet more constant, role across all evaluation conditions.

GENERAL DISCUSSION

This paper investigates one mechanism that influences how people evaluate sequences that extend over time. We demonstrate that different evaluation tasks trigger distinct evaluation processes that then determine specific order effects. We employed multiple operationalizations of evaluation tasks: hedonic versus information tasks and descriptive versus predictive evaluations (Experiments 1 and 2), as well as evaluation of the source versus the experience (Experiment 3). Our experiments also allowed for two types of analyses, the first examining effects of evaluation tasks on the relationship between increasing and decreasing patterns and overall evaluations of (i.e., order effects). The second analysis enabled us to examine the underlying integration processes in more detail, uncovering the relative impact of experience characteristics (start, trend, and end) on overall evaluations.

Our results show that hedonic evaluation tasks lead to stronger preferences for improving experiences than do informational tasks, and that under hedonic evaluation tasks, descriptive and predictive evaluations have similar effects. In contrast, under informational tasks, the predictive task shifts relative preferences towards improving experiences. In addition, we show that for hedonic evaluations, evaluating the source of the experience rather than the experience itself can attenuate preferences for improving experiences.

Analyzing the effect of pattern characteristics on overall evaluations allows us to more closely examine the effect of evaluation tasks on the underlying integration process. In direct support of our hypotheses, we show that the differences in order effects observed for overall evaluations are driven by specific changes in the weighting of key characteristics of the experience. Our results show that the most important difference between evaluation tasks is the relative strength of the initial information compared to the rate of change. The initial information had significantly stronger impact given informational or source evaluation tasks, while the rate of change had a stronger impact under hedonic or predictive evaluation tasks. It is interesting to note that, across all evaluation conditions, the end seems to have a relatively strong constant effect, a finding consistent with interpreting recency as a memory phenomenon.

Lastly, our findings suggest that caution is warranted when interpreting order effects as primacy or recency. First, there appears to be a constant hidden recency effect even when the overt order effects do not indicate that. Second, what appear to be recency effects could be caused by the impact of the sequence's trend rather than the end *per se*. More attention should be given to the impact of the multiple factors, including the sequence's trend, in theories of information integration and order effects.

A task-goal perspective

We conceptualize our effects using a task-goal perspective (e.g., Fischer et al., 1999). This perspective states that any task activates a specific evaluation goal. Fischer et al. (1999) introduced the concept of task goals in the context of preference formation. We extend the concept of task goals to the context of evaluating experiences over time. Under this perspective, requests to form different judgments trigger distinct evaluation task goals, subsequently influencing the type of processes individuals use to generate a response. Consequently, the information processing of identical external stimuli can differ depending on the specific evaluation goal triggered.

The peak–end rule

In the current paper, we examine the effect that the start, end, and trend have on overall evaluations of extended experiences. Others have argued that the peak-end rule can parsimoniously describe the hedonic integration process (e.g., Kahneman et al., 1993; Schreiber & Kahneman, 2000). Our design does not allow us to isolate the influence of the peak, since it is confounded with either start or end. Yet, although our results are silent about the effects of peak, they clearly demonstrate that the end of the pattern is important, even when overall evaluations might not reveal that. Our findings demonstrate explicitly that the end is an important driver of overall evaluations regardless of the specific evaluation task. While we argue that the trend of the experience is also a key driver of hedonic evaluations (Ariely, 1998; Ariely & Zauberman, 2000, 2003; Diehl & Zauberman, 2005), the design of our stimuli renders trend a particularly salient feature. We expect that people are more likely to rely on the peak and the end the less detectable the trend of the experience is.

Summary and conclusion

Our goal in this paper was to illustrate how the evaluation task people engage in can change which aspects of an experience become more influential when evaluating an experience and can thus lead to differential order effect. We show that characteristics of the experiences (specifically, the start and trend) are weighted differently as a function of the evaluation task and that these differences drive overall evaluations.

In conclusion, how people combine information over time is of central importance in the study of judgment and decision making. The objective of the current work was to highlight the role of the task-goal in understanding information integration processes. However, this is only an initial investigation. Additional work is needed to develop a unifying framework that helps us understand more precisely the effects of different patterns, different modes of evaluation, and the inferences that individuals make about sequences of outcomes.

ACKNOWLEDGMENTS

The current research originated in the behavioral lab at Duke University's Fuqua School of Business, where all three authors used to play. We thank Jim Bettman and John Lynch for their many suggestions. For comments on an earlier draft we thank Jonathan Levav, Cait Poynor, and Rebecca Ratner. We also thank the editor and reviewers for helpful comments.

REFERENCES

Abele, A., & Petzold, P. (1998). Pragmatic use of categorical information in impression formation. *Journal of Personality & Social Psychology, 75*(2), 347–358.

Abelson, R. P., Kinder, D. P., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology, 42*(4), 619–630.

Anderson, N. H. (1981). *Foundations of Information Integration Theory.* New York: Academic Press.

Ariely, D. (1998). Combining experiences over time: the effects of duration, intensity changes & on-line measurements on descriptive pain evaluations. *Journal of Behavioral Decision Making, 11*, 19–45.

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., & Zauberman, G. (2000). Breaking and combining experiences over time: the effect of experience patterns and its segmentation on overall experiences evaluation. *Journal of Behavioral Decision Making, 13*, 219–232.

Ariely, D., & Zauberman, G. (2003). Differential partitioning of extended experiences. *Organizational Behavior and Human Decision Processes, 91*(2), 128–139.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258–290.

Diehl, K., & Zauberman, G. (2005). Searching ordered sets: evaluations from sequences under search. *Journal of Consumer Research, 31*(4), 824–832.

Edell, J. A., & Burke, M. C. (1987). The power of feelings in understanding advertising effects. *The Journal of Consumer Research, 14*(3), 421–433.

Fischer, G. F., Carmon, Z., Ariely, D., & Zauberman, G. (1999). Goal-based construction of preferences: task goals and the prominence effect. *Management Science, 45*(8), 1057–1075.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65*(1), 45–55.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*(3), 258–268.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: the belief-adjustment model. *Cognitive Psychology, 24*(1), 1–55.

Hsee, C. K., & Abelson, R. P. (1991). Velocity relation: satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology, 60*, 341–347.

Hsee, C. K., Abelson, R. P., & Salovey, P. (1991). The relative weighting of position and velocity in satisfaction. *Psychological Science, 2*, 263–266.

Jones, E. E., Rock, L., Shaver, K. G., Goethals, G. R., & Ward, L. M. (1968). Pattern of performance and ability attribution: an unexpected primacy effect. *Journal of Personality and Social Psychology, 10*(4), 317–340.

Jones, E. E., & Goethals, G. R. (1972). Order effects in impression formation: attribution context and the nature of the entity. In E. E. Jones, D. E. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior.* New Jersey: General Learning Press.

Jones, E. E., Goethals, G. R., Kennington, G. E., & Severance, L. J. (1972). Primacy and assimilation in the attribution process: the stable entity proposition. *Journal of Personality, 40*(2), 250–274.

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.

Keren, G., & Lewis, C. (1979). Partial omega squared for ANOVA designs. *Educational and Psychological Measurement, 39*, 119–128.

Lichtenstein, M., & Srull, T. K. (1987). Processing objectives as a determinant of the relationship between recall and judgment. *Journal of Experimental Social Psychology, 23*(2), 93–118.

Loewenstein, G. F. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal, 97*, 666–684.

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

Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude-behavior relation. *Journal of Personality and Social Psychology, 51*(2), 270–276.

Miller, N., & Campbell, D. (1959). Recency and primacy in persuasion as a function of the timing of speeches and measurements. *Journal of Abnormal & Social Psychology, 59*(July), 1–9.

Mussweiler, T. (2003). Comparison processes in social judgment: mechanisms and consequences. *Psychological Review, 110*(3), 472–489.

Novemsky, N., & Ratner, R. K. (2003). The time course and impact of consumers' erroneous beliefs about hedonic contrast effects. *Journal of Consumer Research, 29*, 507–516.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and descriptive evaluations of two minimally invasive procedures. *Pain, 66*(1), 3–8.

Ross, W. T., & Simonson, I. (1991). Evaluations of pairs of experiences: a preference for happy endings. *Journal of Behavioral Decision Making, 4*, 273–282.

Schreiber, C. A., & Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General, 129*(1), 27–42.

Trope, Y., & Liberman, N. (2000). Temporal construal and time-dependent changes in preference. *Journal of Personality and Social Psychology, 79*(6), 876–889.

Authors' biographies:

Gal Zauberman is an assistant professor of Marketing at the University of North Carolina at Chapel Hill. He received a PhD in Marketing from Duke University in 2000 and a BA degree in Psychology and Economics from The University of North Carolina, Chapel-Hill in 1994.

Kristin Diehl is an assistant professor of Marketing at the University of Southern California. She received a PhD in Marketing from Duke University in 2002 and a Diplom-Kauffrau from Johannes Gutenberg Universität in Mainz, Germany in 1997.

Dan Ariely is the Luis Alvarez Renta Professor at the Sloan School of Management and at the Media Laboratory, both at MIT. He received a PhD in Marketing from Duke University in 1998, and a PhD in Psychology from the University of North Carolina, Chapel-Hill in 1998. Professor Ariely directs the Center for Advanced Hindsight.

Authors' addresses:

Gal Zauberman, McColl Building, CB# 3490, The University of North Carolina, Chapel-Hill, NC 27599-3490, USA.

Kristin Diehl, Marshall School of Business, University of Southern California, 3660 Trousdale Parkway, Room ACC 306E, Los Angeles, CA 90089-0443, USA.

Dan Ariely, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.

