Journal of Experimental Psychology: General
2000, Vol. 129, No. 4, 524–529

0096-3445/00/$5.00 DOI: 10.1037//0096-3445.129.4.524

Joint Comment on "When Does Duration Matter in Judgment and Decision Making?" (Ariely & Loewenstein, 2000)

Dan Ariely
Massachusetts Institute of Technology

Daniel Kahneman
Princeton University

George Loewenstein
Carnegie Mellon University

Recent research has demonstrated that people care about the temporal relationships within a sequence of experiences. There is considerable evidence that people pay particular attention to the way experiences improve or deteriorate over time and to their maximum (peak) and final values. D. Kahneman and coauthors suggested in earlier articles that people ignore or severely underweight duration (which they referred to as *duration neglect*). In the preceding article, D. Ariely and G. Loewenstein (2000) challenged the generalizability of these findings and their normative implications. In the current commentary, D. Ariely, D. Kahneman, and G. Loewenstein jointly examine the issue to provide a better understanding of what they feel they have learned from this literature and to discuss the remaining open questions.

Areas of Complete Agreement

We are in agreement that when people evaluate experiences retrospectively, they do not play back the equivalent of a movie but instead tend to recall specific salient features of the experience—for example, the peak (or trough), ending value, and slope. We also agree—although none of us has discussed it previously in print—that for many experiences, the most important feature may be the aspects of the experience that gave it meaning; for example, someone may recall, "that was the trip when I proposed to my wife" or "that was the colonoscopy when they discovered the polyp."

We are also in agreement that duration is an attribute of extended experiences that people are often likely to neglect. For many types of experiences, people may automatically make judgments about the quality of the experience in which duration is not considered. This can be the case for vacations, hospital stays, spells of unemployment, and so forth. The same is true for shorter episodes, such as brief aversive films or colonoscopies. In many of these situations, an evaluation of the pleasure or displeasure of the experience is immediately available, whereas a global judgment that incorporates duration must be constructed more laboriously. Duration is not unique in this respect, nor is it always neglected. Indeed, in some cases such as labor pain or commuting, it may be the most important feature that is stored in memory. Which features of an experience are recalled depends on which are salient during the experience and which are rehearsed in memory and discussed with others. Duration is often not salient, rehearsed, or discussed.

We agree that people consider duration relevant to many decisions and will rely on this variable when it is made sufficiently salient. For example, we agree that people do care about duration when it comes to prospective choices in which duration is explicit or well known. When faced with prospective choices between alternatives that differ in duration and other attributes—such as a 5-day ocean cruise in a first class cabin versus a 2-week cruise in tourist class—people will pay attention to duration and trade it off against other attributes. We also agree that the manipulation of duration in the context of a within-subject factorial design increases the effect of this variable on rating responses and on choices. This pattern is not isolated: Within-subject designs typically increase the salience of manipulated variables (Poulton, 1989). In summary, none of us expects to observe complete

Editor's Note. This article uses an unusual format to highlight issues regarding the article "When Does Duration Matter in Judgment and Decision Making?" (Ariely & Loewenstein, 2000). Originally, in response to the above-mentioned article, Daniel Kahneman wrote a comment and Dan Ariely and George Loewenstein wrote a response to that comment. However, these authors were afraid that the two separate articles exaggerated their differences and failed to represent their considerable areas of agreement. Therefore, they have written a joint comment to underline their agreements and to identify their disagreements so that interested readers can read their points in sum and arrive at their own conclusions.—NSN

Dan Ariely, Sloan School of Management, Massachusetts Institute of Technology; Daniel Kahneman, Woodrow Wilson School of Public and International Affairs, Princeton University; George Loewenstein, Department of Social and Decision Sciences, Carnegie Mellon University.

Correspondence concerning this article should be addressed to Dan Ariely, Massachusetts Institute of Technology, 38 Memorial Drive, E56-329, Cambridge, Massachusetts 02142; Daniel Kahneman, Woodrow Wilson School of Public and International Affairs, 434 Robertson Hall, Princeton University, Princeton, New Jersey 08544-1013; or George Loewenstein, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890. Electronic mail may be sent to ariely@mit.edu, kahneman@princeton.edu, or gl20@andrew.cmu.edu.

duration neglect when attention is explicitly or implicitly drawn to this variable.

We agree that the violations of dominance in choice are one of the most striking results in this domain of research (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Schreiber & Kahneman, 2000). The particular form of these violations was predicted by a conception that incorporated duration neglect, but we agree that, as Ariely and Loewenstein (2000) point out, the violations of dominance by themselves do not demonstrate duration neglect.

We also agree that duration is properly neglected in many evaluations and in many conversational contexts (see Ariely & Loewenstein, 2000). The question, "how painful was your colonoscopy?" is appropriately answered by describing the worst moment of the procedure, without reference to duration. Indeed, conversational norms generally call for the exclusion of duration from the overall evaluation of an episode except in situations, such as commutes or labor pain, in which duration is a highly salient attribute.

Areas of Mixed Agreement

We agree that people normally form a duration-free representation of episodes, which they often rely upon in various judgments and choices. Kahneman and his colleagues have described this representation as a composite snapshot (Fredrickson & Kahneman, 1993) or as a prototypical moment (Kahneman, 2000; Schreiber & Kahneman, 2000). Other work has shown that such prototypical perceptions also underlie visual perception of multiple items (Ariely, in press).

Ariely and Loewenstein believe that it is typically normative for people to encode a unitary evaluation of sequences that neglects duration. Most often, people know the duration of extended experiences that they choose between, or they can decide when to terminate an experience as it is occurring. If duration were encoded into the stored representation of desirability and the decision maker was deciding whether to experience a new episode of different duration from the one already experienced, judging the new sequence would require the decision maker to partial out the effect of duration from his or her judgment of the initial evaluation and then combine the new duration into it. As Ariely and Loewenstein (2000) discussed, the circumstances in which duration neglect can lead to problems in prospective choice are when (a) people do not know the duration of future experiences they make decisions about, (b) they cannot decide upon their duration, and (c) there is a high degree of correlation between the duration of past and future episodes. Ariely and Loewenstein believe this combination of circumstances is rare.

Kahneman believes that people normally form a representation of a prototypical moment of an episode and store information about its duration separately. Affect is associated with the duration-free representation and will dominate preferences unless attention is specifically called to duration. As is illustrated by violations of dominance—adding diminishing pain to an episode makes it less aversive in memory—choosing by one's stored evaluation of episodes will sometimes lead people to choices they would not want to make.

Kahneman and colleagues (Kahneman, 2000; Kahneman, Ritov, & Schkade, 1999; Schreiber & Kahneman, 2000) have drawn an analogy between the low salience of duration in the judgment of episodes and the low salience of base rate information in individual prediction, sample size in intuitive statistical inferences, and the number of animals that die in ecological disasters. In all of these cases, the "extensional" information (duration, base rate, sample size, extent of damage to a species) has lower salience than other aspects of the situation, and in all of these cases, complete extension neglect is observed but only in between-subjects designs. However, the manipulation of an extensional variable in a within-subject factorial design always yields a significant effect of that variable on judgments. Kahneman views duration neglect as a special case of a broader family in which categories or sets (e.g., a set of moments) are evaluated by judging a prototype. Ariely and Loewenstein agree that the analogies are intriguing but believe that more research needs to be done to establish that this single concept underpins findings in such apparently disparate literatures such as duration neglect, contingent valuation, and base rate utilization.

We agree that instructions to rate expressions such as "total pain" are difficult to write and difficult to interpret and follow. We also agree that some respondents in previous studies may have interpreted the task as reporting some measure of the representative pain or affect experienced during the episode, although we disagree about the importance of such interpretations for past findings of duration neglect. Ariely and Loewenstein believe that their study (Ariely & Loewenstein, 2000) provides evidence that this is a serious problem when it comes to studies that involve ratings. Kahneman believes that two experiments discussed in the next section of this article show that duration neglect is not an artifact of using ratings as a dependent measure because it has been found in choices in the same experiments in which it was found in ratings. Ariely and Loewenstein question the conclusion that can be drawn from these two studies.

Disagreements

We have identified a few issues on which we disagree, and we discuss these in turn. The issues concern (a) whether duration neglect has been observed in choices and (b) the ecological validity of results observed in within-subject factorial experiments.

Is There Evidence of Complete Duration Neglect in Choice? (Part A)

Kahneman believes that there is no evidence for any qualitative difference between measures of choice and ratings with respect to duration effects and that this conclusion is not altered by the experiment that Ariely and Loewenstein (2000) reported. Ariely and Loewenstein, of course, disagree. We discuss several points of contention in sequence.

Kahneman. The possible ambiguity of ratings was a major concern in the design of the two studies in which essentially complete duration neglect was observed (Fredrickson & Kahneman, 1993; Redelmeier & Kahneman, 1996). Both studies therefore included choice questions, which also showed complete duration neglect.

In Experiment 2 of Fredrickson and Kahneman (1993), 96 participants saw film clips (pleasant or unpleasant) that varied in affective intensity and in duration. Each participant saw short versions of half the films and long versions of the others (coun-

terbalanced across participants). Unlike Experiment 1, in which ratings had been used, rankings of preference (a choice task) were used as a measure of global evaluation. The ranking task required participants to provide information that would be used to select a subset of the films for a subsequent experiment. They were instructed that their objective was "to MINIMIZE the overall unpleasantness of the experience of viewing the unpleasant videotape . . ." I interpret these results as meaningful preferences in a choice to which duration was certainly relevant. The results were clear: The pattern of duration effects (and noneffects) was exactly the same for the ranking task of Experiment 2 as for the ratings of Experiment 1. In particular, there was no sign of greater sensitivity to duration in choices than in ratings.

Ariely and Loewenstein. We do not believe this study (Fredrickson & Kahneman, 1993) provides evidence of duration neglect in choice. First, participants did not, in fact, make choices (but gave rankings that would be used by the experimenter to create stimuli for a future experiment). Second, based on the instructions they were given, they may well have believed that it would be most useful for the experimenters to have a measure of average pleasantness or unpleasantness that ignored duration due to the fact that the researchers would make their own decisions about the duration of the films used in the final experiment.

We do, however, believe that it would be possible to design an experiment in which one would obtain duration neglect in choice. Such a study could involve exposing people to stimuli with two attributes—one that varies and is salient and the other, duration, that has low saliency. In such a design, the more we distract peoples' attention from duration, the bigger its neglect will be. The case of duration is by no means unique. Consider an experiment in which we show participants many unpleasant, qualitatively different pictures and also state money amounts that they will receive for viewing each of them. In such an experiment, the content of the pictures will be the highly salient feature and money (payment for viewing the unpleasant pictures) will be the low-salience feature. If we were to ask participants to make choices from memory about pairs of stimuli, we expect that we would obtain "money neglect."

Kahneman. Ariely and Loewenstein (2000) found the results of Experiment 2 of Fredrickson and Kahneman (1993) unpersuasive on the grounds that participants may have thought that neglect of duration was appropriate since the researchers would make their own decisions about the duration of each clip that would be included in the final experiment (Ariely & Loewenstein, 2000). This is a strained reading of instructions that stated, "The unpleasant videotape will consist of several of the unrelated unpleasant clips that you just saw" and instructions that were addressed to participants who had no idea that the film clips they were to rank could also exist in another version. In my view, the similarity of results observed with the ranking task and with a rating task in this experiment refutes the claim that the evidence for duration neglect in ratings can be rejected as an artifact of incomprehensible rating instructions. I believe that we all agree that the existence of a duration-free representation and the low salience of duration could also produce duration neglect—and I believe that this shared interpretation is likely to be the correct one.

Is There Evidence of Complete Duration Neglect in Choice? (Part B)

Kahneman. The second study in which essentially complete duration neglect was observed with ratings involved evaluations of a colonoscopy procedure. Patients were prompted to report their current level of pain every 60 s. They were also interviewed immediately after the procedure, and 89 of the patients were also interviewed after about a month (see Redelmeier & Kahneman, 1996, for details). The second interview included the following question: "If in the future you had to choose between a colonoscopy or a stomach X ray, which medical test would you select?" The results were clear: The biserial correlation between patients, preferences, and the duration of their colonoscopy was .01. In the same data, the biserial correlation between choices and the average of peak and end of their immediate ratings of pain was marginally significant ($p < .03$, one-tailed). Because of an unfortunate oversight, these results were not included in the published report.[1]

I conclude that there is no evidence of a substantial difference between duration effects in choice tasks and in ratings. When complete duration neglect was found in ratings, it was also found in choices. When a factorial design produced an effect of duration on ratings, it also produced effects, perhaps slightly larger, in choices.

Ariely and Loewenstein. Colonoscopy is carried out under the influence of anesthesia. Versed, a popular "hypnotic drug," was given to some of the patients (the researchers were unable to tell us exactly how many) who participated in Redelmeier and Kahneman's (1996) study; only 17% of the patients did not get any drugs. One of the well-established effects of such drugs is that they undermine time perception (as well as virtually eliminating pain). This, then, scarcely seems like an ideal situation in which to examine participants' encoding of the duration of aversive sequences. The fact that the hypothetical choices were made one month later and were between colonoscopy and a procedure that must have been unfamiliar to most or all participants further increases the difficulty of reaching strong conclusions from these findings. Furthermore, the choice question followed a rating question, which could have caused dependency between the answers. Finally, it is always problematic to draw conclusions from null results. Whereas choices were uncorrelated with duration, peak and end were only marginally correlated ($p < .06$, two-tailed) and explained less than 4% of the variance in choice.

In addition, there is a question of whether such evidence should be taken as indicating that people are wrong when choosing between colonoscopy and barium enema when they fail to take duration into account. To make a normative claim of this sort, one has to assume that the duration of the previous colonoscopy is a good predictor of its duration in the future. In their article, Redelmeier and Kahneman (1996) reported that the duration of colonoscopy varied from 4 to 67 min. If this variance is explained by individual difference, then people should take duration into account, and their failure to do so is a mistake (although they may

[1] The procedure of the study is described in detail by Redelmeier and Kahneman (1996). Questions may be addressed to Donald A. Redelmeier, MD, University of Toronto, Sunnybrook Hospital G-151, 2075 Bayview Avenue, Toronto, Ontario, Canada M4N 3M5.

not *know* that the variance in durations is caused by individual differences). However, if the duration of colonoscopies is uncorrelated for an individual across time (on the basis of four experts we asked, this seems to be the case), a normative claim cannot be made, and hence this study does not provide evidence that duration neglect in retrospective evaluations leads to nonnormative duration neglect in prospective decisions.

We should note that we do not reject the idea that duration neglect can be demonstrated in choice. Indeed, above, we outline a design that we believe would obtain duration neglect in choice. However, we do not believe that it has been demonstrated so far in a compelling fashion, and we believe that the conditions that will produce duration neglect in prospective choices in the lab are likely to be exceedingly rare in the real world.

Kahneman. The anesthesia was certainly not sufficient to eliminate pain in the colonoscopy study, in which the average of peak ratings of pain was 7.7 on a 0–10 scale (*no discomfort at all* to *awful discomfort*). If the sense of time was reduced to the same extent that pain was, considerable sensitivity to time must have remained. Temporal disorientation was particularly unlikely in this situation because patients were probed every 60 s to report the intensity of their pain. Time does appear to stretch and shrink in some circumstances, but forgetting regular probes is another matter. There was no indication of duration effect in a small subset of patients who had no anesthesia at all. Furthermore, the same pattern of duration neglect was observed in judgments of total pain by the patients and by the unanesthetized physicians who administered the procedure.

Ariely and Loewenstein. What effects the drugs have in these situations is still an open question worthy of further investigation. In addition, if the physicians knew what effects the drugs would have on their patients (and presumably they did), their answers should reflect their understanding of patients' experience with the drugs in question. If the drugs are such that they cause misperception of time and influence memory, the physicians should have included these effects in their evaluations. Finally, this is a type of experience that is unique, with unpredictable duration, no control over the duration, and a choice between colonoscopy and an experience of unknown duration—all of which could produce duration neglect but could not easily be generalized beyond this limited set of conditions.

What Can Be Learned From Ariely and Loewenstein's (2000) Experiments?

Kahneman. Ariely and Loewenstein (2000) consistently described duration neglect as the topic of their research, but their experiments were conducted in a design in which duration neglect is effectively guaranteed not to occur. It was well-established by prior research that a factorial design consistently yields an effect of duration and other extensional variables on ratings (Ariely, 1998; Kahneman, Ritov, & Schkade, 1999; Schreiber & Kahneman, 2000). Clearly, a factorial design is not optimal for the study of duration neglect, and duration neglect is not what Ariely and Loewenstein studied. The hypothesis they did examine—that a choice task reliably enhances duration effects—was not unambiguously supported by their own data and is not compatible with the results that my collaborators and I have obtained in between-subjects designs.

More important, the issue of ecological validity should be raised. Factorial designs are known to create some of the effects that they are intended to measure by reminding participants of a consideration that they might otherwise have neglected (e.g., Fischhoff & Bar-Hillel, 1984; Poulton, 1989). Life, however, is not structured as a factorial design. One normally evaluates events one at a time, and then forms memories of these evaluations. If someone chooses to be exposed to experiences that he or she remembers having liked, variables that did not influence the evaluative memory will not have much effect on his or her preferences. Duration is therefore likely to be neglected in many of the evaluations and choices that people make every day.

Ariely and Loewenstein. Within-subject designs have been used in prior research (e.g., Schreiber & Kahneman, 2000; Ariely, 1998) that has obtained much weaker effects of duration than we (Ariely & Loewenstein, 2000) found. Nevertheless, we went to some lengths to make our design conservative in the sense of not exaggerating participants' concern for duration.

First, we deliberately designed the study to reduce the demand characteristics associated with within-subject comparisons as much as possible. We did this by having participants rate sequences that differed not only in their duration but also in the way they progressed over time (starting intensity, slope, mean intensity, and final intensity). In this manner, participants' attention was not directed to only one of the features of these sequences.

Second, participants were exposed to the different sequences one at a time or compared each sequence to a reference sequence that had generally not just been experienced but was stored in memory. Numerous studies, including those conducted by Kahneman (Hsee, Loewenstein, Blount, & Bazerman, 1999; Kahneman & Ritov, 1994), have shown that the most critical design element that determines the salience of what is varied within subject is not whether subjects are exposed to a variety of attribute values (as in our study; Ariely & Loewenstein, 2000) but whether they compare pairs of stimuli explicitly (in which case salience is high for attributes that differ across comparison pairs) or evaluate them one at a time (in which case salience of such differences is low). The willingness-to-accept-pain-in-exchange-for-payment experiment in our study (Ariely & Loewenstein, 2000), in which we observed robust concern for duration, was completely noncomparative, that is, one at a time.

Third, we used the same general procedure in all four studies. Whether we found complete duration neglect was not the main point of our work. Instead, we wanted to examine whether the elimination of problems associated with conversational norms and with norms of comparison increases the role of duration in decision making. The results show that at least some of neglect of duration can be caused by these factors.

On the issue of ecological validity, we disagree with Kahneman's view that daily life does not resemble a factorial design. We believe that in day-to-day life, decision makers often face repeated trade-offs between duration and other attributes in choices for which they have considerable experience. People decide which route to take home, whether to cook at home or take out, what time to go to the supermarket to avoid long lines, and so forth. In cases of this type, decision makers have extensive experience with the experiences and their duration. If, as we believe, this type of decision is common, a within-subject experimental design might be more suitable than a between-subjects design.

Kahneman. The inference from the results of the Ariely–Loewenstein (2000) experiment to the role of conversational norms in earlier studies (Fredrickson & Kahneman, 1993; Kahneman et al., 1993; Redelmeier & Kahneman, 1996) that were conducted in completely different designs seems quite remote. The within-subject factorial design has powerful effects on attention and powerful demand characteristics. It draws participants' attention to the manipulated variable, and it also provides them with a strong clue that the experimenter expects each manipulated variable to affect judgments. The demand characteristics of life are not quite the same.

Final Words

Kahneman

We never disagreed much on issues that are amenable to experimental test, and much of what we agree on today was explicit or implicit in my prior articles. In the first discussion of duration neglect, Fredrickson and Kahneman (1993) pointed out that duration neglect does not imply that duration information is lost, nor does it imply that people believe that duration is unimportant.

> Rather, we view duration neglect as an attentional phenomenon: Although people may be aware of duration and consider it important in the abstract, we suggest that what comes most readily to mind in evaluating episodes are the salient moments of those episodes and the affect associated with those moments. Duration neglect might be overcome, we suppose, by drawing attention more explicitly to the attribute of time. (Fredrickson & Kahneman, 1993, p. 54)

In my view, this early statement captures the essence of what the three of us are saying now, but the existence of this level of agreement could not be guessed from reading Ariely and Loewenstein's (2000) article. Specifically, the above statement narrows the scope of duration neglect to situations in which attention is not drawn explicitly to the duration variable.

The three of us also agree that within-subject factorial designs enhance attention to the duration variable. The observation of significant duration effects in a within-subject factorial design was reported and discussed in detail by Schreiber and Kahneman (2000), and a theoretical analysis of it was also offered in Kahneman (2000) and in Kahneman, Ritov, and Schkade (1999). The most robust result of Ariely and Loewenstein's (2000) work—the finding of duration effects in a factorial design—was therefore anticipated in detail. The more specific hypothesis that attracting more attention to duration will increase its effects is new in Ariely and Loewenstein (2000) but certainly congenial with my views. Thus, the main hypothesis that was tested in their study is one on which there was never any disagreement, although we do disagree on the merits of the study itself. Finally, the idea that duration neglect is normatively appropriate in some judgments was explicit in Kahneman (2000), although Ariely and Loewenstein emphasize different examples.

Ariely and Loewenstein

We are delighted that Kahneman agrees with us to the extent that he does. However, we feel that the positions expressed here (as well as the position he expressed in Kahneman, 2000) were not reflected in his earlier writing. We believe that our article accurately represented the views that Kahneman has expressed in his published work. In addition, we believe that our article (Ariely & Loewenstein, 2000) made a number of new points, several of which are discussed above in the Areas of Complete Agreement section. Readers who are sufficiently motivated can read the relevant articles and arrive at their own conclusions about what was said by whom and when. Most importantly, we feel that Kahneman has revised and clarified his opinions in this process and now generally agrees with us regarding the following issues: (a) a more limited scope of circumstances in which duration neglect will be observed, (b) about the differences between using duration for judgments and for decision making, (c) on the normative implications of relying on duration in encoding and judgments, (d) on the difficulty of interpreting previous results of dominance violation as providing unique support for duration neglect, and (e) on the problems of conversational norms when providing overall ratings of experiences.

Kahneman

The conversation that we have had in preparing this article has clarified my opinions in several ways that reduce the gap between our positions—and social psychology suggests that I have probably changed more than I think. The issue of attention certainly looms larger for me than it did before, as does the issue of conversational norms. Neither of these ideas is new to me, but I will be more aware of them in future and more aware of the need to make the boundary conditions for duration neglect explicit.

Ariely and Loewenstein

Our own position has also been modified since we started this project. First, despite our disagreements with Kahneman's interpretation of duration neglect, we have strengthened our belief in the overall theory that preceded this work—the representation by moments theory.

Summary

In sum, our disagreements revolve mainly around the issues of ecological validity and normative claims. Kahneman finds the hypothesis offered by Ariely and Loewenstein (2000) plausible but believes that results based on factorial designs are not representative of judgments and decisions in the real world. Ariely and Loewenstein accept the idea that there likely may be some situations in which people underweight duration in prospective decisions, on the basis of their failure to encode duration retrospectively, but believe that such situations are exceedingly rare. This is not an issue that can easily be settled by running further studies. Instead, we encourage readers to make their own judgments about the degree to which different research findings generalize to the types of decisions which are most common in the world outside of the laboratory.

From this joint comment, it should be evident that we agree upon far more than we disagree on. Our final point of agreement is that further work is needed by researchers not involved in this debate on the circumstances in which duration is or is not neglected and on the factors that moderate the influence of duration in decision making.

References

Ariely, D. (1998). Combining experiences over time: The effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *Journal of Behavior Decision Making, 11,* 19–45.

Ariely, D. (in press). Seeing sets: Representation by statistical properties. *Psychological Science.*

Ariely, D., & Loewenstein, G. (2000). The importance of duration in judgment and decision making. *Journal of Experimental Psychology: General, 129,* 4, 508–523.

Fischhoff, B., & Bar-Hillel, M. (1984). Diagnosticity and the base rate effect. *Memory and Cognition, 12,* 402–410.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65,* 44–55.

Hsee, C. K., Loewenstein, G., Blount, S., & Bazerman, M. (1999). Preference reversals between joint and separate evaluations of options: A theoretical analysis. *Psychological Bulletin, 125,* 576–590.

Kahneman, D. (2000). Evaluation by moments: Past and future. In D. Kahneman & A. Tversky (Eds.), *Choices, values and frames* (pp. 699–708). New York: Cambridge University Press and Russell Sage Foundation.

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

Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty, 9,* 5–38.

Kahneman, D., Ritov, I., & Schkade, D. (1999). Economic preferences or attitude expressions? An analysis of dollar responses to public issues. *Journal of Risk and Uncertainty, 19,* 220–242.

Poulton, E. C. (1989). *Bias in quantifying judgment.* Hillsdale, NJ: Erlbaum.

Redelmeier, D. A., & Kahneman, D. (1996). Patient's memories of painful medical treatments—real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66,* 3–8.

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

Received July 21, 2000
Accepted August 1, 2000 ■