Pain, 66 (1996) 3–8

PAIN 2994

Research papers

Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures

Donald A. Redelmeier*[a,b] and Daniel Kahneman[c]

[a]*Department of Medicine, University of Toronto, Toronto (Canada M4Y 1J3);* [b]*Division of Clinical Epidemiology, Wellesley Hospital Research Institute, Jones Bldg. #123, Toronto (Canada M4Y 1J3) and* [c]*Department of Psychology, Princeton University, Princeton (USA)*

(Received 10 April 1995; revision received 12 September 1995; accepted 20 September 1995)

Summary Patients' memories of painful medical procedures may influence their decisions about future treatments, yet memories are imperfect and susceptible to bias. We recorded in real-time the intensity of pain experienced by patients undergoing colonoscopy (n = 154) and lithotripsy (n = 133). We subsequently examined patients' retrospective evaluations of the total pain of the procedure, and related these evaluations to the real-time recording obtained during the experience. We found that individuals varied substantially in the total amount of pain they remembered. Patients' judgments of total pain were strongly correlated with the peak intensity of pain ($P < 0.005$) and with the intensity of pain recorded during the last 3 min of the procedure ($P < 0.005$). Despite substantial variation in the duration of the experience, lengthy procedures were not remembered as particularly aversive. We suggest that patients' memories of painful medical procedures largely reflect the intensity of pain at the worst part and at the final part of the experience.

Key words: Medical treatment and pain; Memory; Colonoscopy; Lithotripsy; Pain intensity

Introduction

The experience of pain is too complex to be fully described by a single number, but reports of pain intensity provide one useful measure of a person's experience at a particular moment. Using such reports of current experience, a painful medical procedure could be represented by a profile of intensity over time (Lewis et al. 1995). Consider, for example, the pain profiles of two patients undergoing colonoscopy who report the intensity of their pain throughout the procedure (Fig. 1). A natural question arises; namely, which procedure was more painful? One way to answer this question is to ask each patient for a retrospective evaluation of the overall episode. A more analytic approach is to assess the patients' overall experience from their real-time reports. The present article compares these two approaches for evaluating painful medical procedures.

Several aspects of a pain profile could be used to assess a painful episode. *Peak Pain* is the intensity of pain at the worst moment of the episode. *Initial Pain* and *End Pain*, respectively, refer to the intensity of pain at the initial and final moments of the episode. If the intensity measure satisfies the requirements of an interval scale, an estimate of *Average Pain* can also be calculated over the entire episode. Similarly, an estimate of *Total Pain* can be calculated as the area under the curve. Observers comparing the profiles in Fig. 1 generally agree that the experience of Patient B was worse than Patient A because of the greater total pain. This judgment assumes that both patients used the scale similarly. This judgment does not assume that the patients, themselves, would accurately recall the episode and integrate diverse moments of the experience.

How do people retrospectively evaluate unpleasant episodes? Previous research indicates that judgments about painful episodes are potentially inaccurate (Corli et al. 1986; Varey and Kahneman 1992; Algom and Lubel 1994). In particular, two counter-intuitive results have been found. *Peak and End Evaluation*: total remembered discomfort is largely determined by the in-

* *Corresponding author:* at Wellesley Hospital address. Tel.: +1 416 926-7696; fax: +1 416 926-4910.

PII 0304-3959(96)02994-6

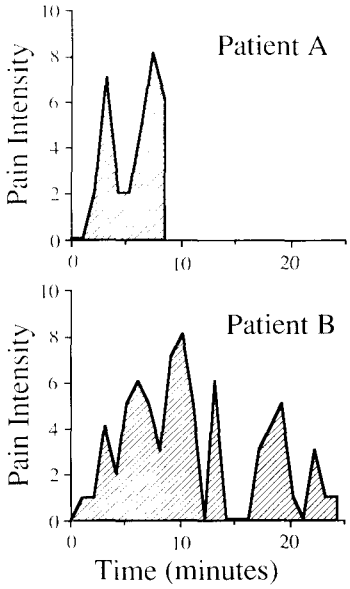

Fig. 1. Real-time recordings from two patients. Each graph displays the intensity of pain recorded each minute by a patient undergoing colonoscopy. The experiences of two individuals are shown (Patient A and Patient B). The *x*-axis represents time in minutes from the start of the procedure. The *y*-axis represents the intensity of pain recorded in real-time on a visual analogue scale with ends denoted as 'no pain' and 'extreme pain'. The procedure lasted 8 min for Patient A and 24 min for Patient B.

tensity of discomfort at the worst part and at the final part of the episode (Fredrickson and Kahneman 1993). Other moments of the experience, such as the beginning, count much less in people's overall evaluation. *Duration Neglect*: the duration of the episode has little direct effect on total remembered discomfort (Kahneman et al. 1993). Although long episodes are sometimes remembered as quite aversive, this typically occurs when discomfort escalates over time so that the final parts of the episode are relatively intolerable. In this study we explored whether Peak and End Evaluation and Duration Neglect occur in patients' memories of painful medical procedures.

Methods

Patients and setting

We identified consecutive outpatients having elective medical procedures at the Wellesley Hospital in Toronto, Canada. The first study involved patients undergoing colonoscopy ($n = 154$) and the second study involved patients undergoing lithotripsy ($n = 133$). These two procedures were selected because they provided unambiguous starting and ending points. The colonoscopy procedure was defined as the interval between insertion and removal of the colonoscope, and the lithotripsy procedure was defined as the interval between the first and last shock. In both studies we excluded individuals who did not speak English, had neurologic limitations that precluded participation, or were suffering severe comorbid conditions. No colonoscopy or lithotripsy was performed solely for research purposes and all patients gave informed consent.

Real-time measures of pain intensity

Patients' real-time experiences were recorded using the Gottman–Levenson method for measuring emotional responses (Gottman and Levenson 1985). A hand-held device was used to control the position of a marker displayed on a computer screen. The computer screen presented a 19-cm visual analogue scale, with ends denoted as 'no pain' and 'extreme pain', and prompted the patient every 60 sec to 'rate the current intensity of pain'. The position of the marker was then converted into a score that ranged between 0 and 10, where higher numbers indicated more severe pain (Price et al. 1982). To check whether the act of providing real-time recordings might alter the nature of an individual's experience, a group of consecutive patients in the colonoscopy study were not required to provide ratings ($n = 53$); instead, a research assistant observed the patient throughout the procedure and made real-time ratings every 60 sec.

Patients' memories of the total amount of pain experienced

Patients provided retrospective evaluations within an hour after the procedure by judging the 'total amount of pain experienced'. They used a 10-point rating scale with ends denoted as 'no discomfort' and 'awful discomfort' (Chapman et al. 1985). We used this scale to emphasize the distinction between real-time reports of pain intensity (which assessed a single moment) and retrospective evaluations (which assessed the entire experience). To eliminate possible effects of shared question format (Linton and Gotestam 1983), patients also retrospectively judged the unpleasantness of the procedure relative to seven other bad experiences, such as 'an average visit to the dentist'. For clarity, we refer to the former set of retrospective evaluations as ratings, and the latter as rankings. As a check on the reliability of retrospective evaluations, we asked patients 1 month after colonoscopy and 1 year after lithotripsy to recall the experience and again rate the total amount of pain from the procedure.

Physicians' memories of the total amount of pain experienced

The attending physician provided a convenient additional method for analyzing people's retrospective evaluations of painful medical procedures. They were asked immediately after the procedure to estimate how the patient would subsequently rate the total amount of pain experienced. Additionally, they were asked to judge whether more anesthetic should have been used if they had the procedure to do over again. Most patients received some medication prior to conducting the procedure (typically a short-acting benzodiazapene or short acting opiate) which had been ordered according to the physician's clinical judgment. Physicians were unaware of the hypothesis of the study, had been present throughout the medical procedure, and made all judgments without knowledge of the patient's real-time reports.

Statistical analyses

All available data from all patients were included in the analysis. Missing data on real-time recordings were rare, accounting for 2% of all data values, and primarily due to patients failing to respond within the time prompted. Missing values were replaced by carrying forward the value that preceded the missing value for a given patient. Retrospective evaluations in the recovery room were incomplete for seven colonoscopy patients (one patient who provided neither a rating nor a ranking and six other patients who did not provide a ranking). Retrospective evaluations in the recovery room were complete for all lithotripsy patients. Retrospective evaluations at follow-up were not collected for the 53 patients who did not provide real-time reports, and were missing for 12 (12%) colonoscopy patients and 28 (21%) lithotripsy patients. We did not attempt to model missing retrospective evaluations.

We examined the relation between different retrospective ratings by Pearson correlation statistics and by the *t*-test. We examined the relation between retrospective evaluations and real-time measures using the Pearson correlation coefficient. To assess the adequacy of the Peak and End rule for explaining retrospective evaluations we developed two multivariable models based on patients' real-time reports. The simple model included only a linear combination of Peak Pain and End Pain as predictors. The comprehensive model included four additional predictors: Total Pain, Average Pain, Initial Pain, and Duration. We measured the adequacy of each model for explaining retrospective evaluations using the multivariable correlation coefficient and we tested the improved accuracy of the comprehensive model by using forward stepwise regression with the *F*-test.

Results

Descriptive

Real-time recordings indicated that both procedures caused moments of considerable pain (Table I). Overall, 38% of colonoscopy patients and 22% of lithotripsy patients reported a pain score of 10 (the maximum possible) at least once during the procedure. Pain intensity during the procedure was generally lower during the final 3 min than the initial 3 min for colonoscopy (mean score: 2.0 vs. 2.6, $P < 0.001$). In contrast, pain intensity was generally higher during the final 3 min than the initial 3 min for lithotripsy (4.4 vs. 1.6, $P < 0.001$). The correlation between Peak Pain and End Pain was smaller for colonoscopy than for lithotripsy ($r = 0.34$ and $r = 0.67$ for colonoscopy and lithotripsy, respectively).

The procedures varied considerably in total duration: 4–67 min for colonoscopy and 18–51 min for lithotripsy. There was no significant correlation in either case between the duration of the procedure and the patient's average intensity of pain ($r = 0.03$ and $r = 0.08$, respectively). Additionally, there were no significant correlations between Duration and Initial Pain ($r = 0.09$ and $r = 0.02$, respectively) or End Pain ($r = 0.04$ and $r = 0.01$, respectively). The correlation between Duration and Peak Pain was statistically significant but small ($r = 0.21$ and $r = 0.12$, respectively). A small positive correlation between duration and Peak Pain was expected because prolonging a procedure can only bring about an increase of Peak Pain, never a decrease.

Patients varied in the total amount of pain they remembered from the procedure. Retrospective ratings obtained in follow-up were highly correlated with the ratings obtained in the recovery room ($r = 0.77$ and $r = 0.54$ for colonoscopy and lithotripsy, respectively). For colonoscopy patients, the average retrospective rating did not change over time (score: 4.6 vs. 4.6, $P > 0.20$). For lithotripsy patients, delayed retrospective ratings were more aversive than the immediate retrospective ratings (score: 5.5 vs. 4.8, $P = 0.002$). The two retrospective evaluations obtained in the recovery room, ratings and rankings, yielded generally similar assessments ($r = 0.68$ and $r = 0.51$ for colonoscopy and lithotripsy, respectively).

Physicians' and patients' retrospective evaluations were generally similar. The correlation between physicians' ratings and patients' ratings was significant for both colonoscopy ($r = 0.67$, $P < 0.001$) and lithotripsy ($r = 0.46$, $P < 0.001$). Additionally, physicians' mean rating was comparable to patients' mean rating for both colonoscopy (score: 4.3 vs. 4.6, $P = 0.12$), and lithotripsy (score: 4.6 vs. 4.8, $P > 0.20$). For colonoscopy, physicians' judgments about the desirability of using more anesthetic were correlated with physicians' judgments of the overall pain ($r = 0.48$, $P < 0.001$) and patients' ratings of overall pain ($r = 0.44$, $P < 0.001$). For lithotripsy, physicians' judgments about the desirability of using more anesthetic were correlated with physi-

Table I
CHARACTERISTICS OF PARTICIPANTS AND PROCEDURES

	Colonoscopy	Lithotripsy
Number of patients	154	133
Mean age (years)	56 ± 14	47 ± 14
% Female	55	35
% Previous experience with procedure	42	46
% Received intravenous analgesia	85	68
Real-time assessments during procedure		
Duration (min)	23 ± 13	33 ± 6
Peak Pain	7.7 ± 2.7	6.4 ± 3.1
Initial Pain	2.6 ± 2.5	1.6 ± 2.0
End Pain	2.0 ± 2.7	4.4 ± 3.1
Average Pain	3.1 ± 2.0	3.8 ± 2.5
Total Pain (area under the curve)	72 ± 61	126 ± 84
Retrospective assessments of procedure		
Patient's rating (immediate)	4.6 ± 2.6	4.8 ± 2.4
Patient's rating (follow-up)	4.6 ± 2.7	5.5 ± 2.4
Patient's relative ranking	4.3 ± 2.2	3.7 ± 1.9
Physician's rating	4.3 ± 2.6	4.6 ± 2.6
Physician's judgment (% should use more anaesthetic)	16	12

Values are mean ± standard deviation, or percentage of each group.

cians' judgments of the overall pain ($r = 0.29$, $P < 0.007$) but not with patients' ratings of overall pain ($r = 0.12$, $P > 0.20$).

Analytic

We found no significant correlation between the duration of the procedure and retrospective evaluations — a striking illustration of Duration Neglect (Table II). Despite the substantial variability in the duration of pain, longer procedures were not judged more aversive than shorter procedures. The low correlation of retrospective evaluations with duration were also evident for those patients who received no intravenous analgesia. For the 42 lithotripsy patients who received no intravenous analgesia, for example, the correlation between the procedure's duration and patients' retrospective ratings in the recovery room was not statistically significant ($r = 0.08$, $P > 0.20$). Moreover, Duration Neglect was also observed in physicians' retrospective ratings of patients' pain and physicians' judgments about the desirability of using more anesthetic (Table II).

Patients' retrospective evaluations were strongly correlated with Peak Pain and End Pain. This pattern was found in patients' immediate retrospective ratings, delayed retrospective ratings, and relative rankings of the procedure (Table II). The same pattern was evident for those patients who did not provide real-time reports. Specifically, for the 53 colonoscopy patients who had real-time ratings estimated by a research assistant, the correlation between Peak Pain and patients' retrospective ratings in the recovery room was substantial ($r = 0.62$, $P < 0.001$), as was the correlation between the End Pain and patients' retrospective ratings in the recovery room ($r = 0.50$, $P < 0.001$). Moreover, physicians' retrospective ratings and physicians' judgments about the desirability of using more anesthetic were also both significantly related to pain at the worst moment of the procedure and at the end of the procedure (Table II).

The adequacy of the Peak and End rule was examined by comparing two multivariable models for explaining peoples' retrospective evaluations. The simple model included only Peak Pain and End Pain as predictors. The comprehensive model included four additional predictors: Total Pain, Average Pain, Initial Pain, and Duration. The improvement in accuracy obtained by including the additional variables was small, and in several cases not statistically significant (Table III). In no case was the marginal improvement in accuracy related to measurement of Total Pain. The largest single difference between the simple model and the comprehensive model occurred for physicians' judgments about the desirability of using more anesthetic in lithotripsy patients, and was related to measurement of the procedure's Average Pain.

Table II
RELATION BETWEEN RETROSPECTIVE ASSESSMENTS AND SELECTED REAL-TIME MEASURES

	Duration	Peak pain	End pain
Colonoscopy			
Patient's rating (immediate)	0.03	0.64*	0.43*
Patient's rating (follow-up)	0.12	0.61*	0.44*
Patient's relative ranking	0.14	0.51*	0.42*
Physician's rating	0.15	0.64*	0.44*
Anaesthetic judgment	0.05	0.35*	0.32*
Lithotripsy			
Patient's rating (immediate)	0.11	0.63*	0.56*
Patient's rating (follow-up)	0.04	0.46*	0.45*
Patient's relative ranking	0.02	0.36*	0.40*
Physician's rating	0.10	0.42*	0.33*
Anaesthetic judgment	0.02	0.23*	0.30*

Values are Pearson correlation coefficients.
*Values which are statistically significantly different from zero ($P < 0.05$).

Discussion

Previous research has suggested that people remember pain with fair accuracy. However, studies have tended to compare immediate retrospective evalua-

Table III
PREDICTING PEOPLES' MEMORIES FROM COMBINATIONS OF REAL-TIME MEASURES

	Simple model	Comprehensive model
Colonoscopy		
Patient's rating (immediate)	0.67	0.69*
Patient's rating (follow-up)	0.65	0.65
Patient's relative ranking	0.56	0.59*
Physician's rating	0.68	0.73*
Anaesthetic judgment	0.40	0.42*
Lithotripsy		
Patient's rating (immediate)	0.65	0.67*
Patient's rating (follow-up)	0.48	0.48
Patient's relative ranking	0.41	0.41
Physician's rating	0.42	0.42
Anaesthetic judgment	0.27	0.35*

Values are Pearson correlation coefficients.
*Values which are statistically significantly different between the two models ($P < 0.05$).

tions with delayed retrospective evaluations — which we also found were correlated (Hunter et al. 1979; Erskine et al. 1990). In this study we compared real-time assessments to retrospective evaluations and found systematic discrepancies. In accord with laboratory research, patients' memories of the overall pain of both colonoscopy and lithotripsy were characterized by Peak and End Evaluation and Duration Neglect. Furthermore, we found the same pattern in physicians' judgments, indicating that the failures of memory were not merely an effect of analgesia or real-time reporting. Together, these observations suggest that distortions in peoples' retrospective evaluations of painful medical procedures occur when judgments are first constructed rather than as a result of gradual forgetting.

Our study focused on people's judgments of the total pain of a completed episode, not people's memory of selected moments. The accuracy of patient's recall of pain at selected moments remains debatable (Babul and Darke 1994; Morley 1994). Some investigators claim that peoples' memory for selected moments is fairly accurate (Rofé and Algom 1985; Salovey et al. 1993). Others suggest that post-hoc ratings differ from ratings at the time of the experience (Linton and Melin 1982; Eich et al. 1985; Kent 1985; Beese and Morley 1993; Bryant 1993), and a few report that inaccuracies are related to pain and emotion at the time of the elicitation (Teasdale and Fogarty 1979; Pearce et al. 1990; Smith and Safer 1993). We did not test for state dependency or mood congruity in memory for pain, in that almost all individuals were pain-free and calm at time of follow-up assessment. And we make no claims about the accuracy of patients' memories of pain at selected moments. Our study suggests, however, that even perfect recall of selected moments would not imply that patients accurately remember an entire episode of pain.

The discrepancy between people's real-time and retrospective evaluations is not surprising given the limitations of human memory and judgment (Simon 1959; Redelmeier et al. 1993). Episodes of pain are extremely complex and storing all the details might be overwhelming. Peak Pain and End Pain are distinct moments that occur in all episodes and provide convenient measures of comparison. Other summary measures, such as Average Pain or Total Pain, are much more difficult to construct and would require the individual to integrate the whole experience over time. Whereas the discrepancy between real-time and retrospective evaluations may yield conflicting comparisons in select circumstances, in most situations Peak and End Evaluation should result in a reasonable judgment of past painful episodes. In particular, the neglect of Duration would not be a serious memory bias if mild pain was usually brief and severe pain was usually long-lasting.

Peak and End Evaluation and Duration Neglect have significant implications for how clinicians conduct painful medical procedures. If the objective is to reduce patients' memory of pain, for example, lowering the peak intensity of pain could be more important than minimizing the duration of the procedure. By the same reasoning, gradual relief may be preferable to abrupt relief if patients retain a less aversive memory when the intense pain does not occur near the end of the procedure. In contrast, if the objective is to reduce the amount of pain actually experienced, conducting the procedure swiftly may be appropriate even if doing so increases the peak pain intensity and leaves patients with a particularly aversive memory. Regardless of specific objectives, both clinicians and researchers should recognize that retrospective evaluations may not be an adequate substitute for real-time reports for assessing patients' pain (such as when comparing the effectiveness of different analgesics).

Our findings also raise a complex ethical issue for clinicians who perform painful medical procedures on awake patients. Suppose retrospective evaluations conflict with real-time reports, as in our patients undergoing colonoscopy and lithotripsy. Which perspective should be taken more seriously? If the patient had a cognitive limitation which impaired memory (such as Alzheimer's disease), most clinicians would concentrate on the real-time reports. Given the natural limitations of human memory, however, perhaps real-time reports should also be given priority for patients who are neurologically intact. Yet doing so is problematic because memories — not experiences — form the basis of patients' future decisions about treatment (including their compliance with recommendations for follow-up). For procedures where some pain is unavoidable, clinicians may need to decide whether it is more important to optimize patients' experiences or memories.

Acknowledgement

Support from the National Science Foundation, the Sloan Foundation, and the National Cancer Institute of Canada is gratefully acknowledged. Dr Redelmeier is supported by a career scientist award from the Ontario Ministry of Health. We wish to thank the attending staff, fellows, and nurses of the Endoscopy Unit of the Wellesley Hospital. We are also indebted to the participating patients for their thoughtful effort.

References

Algom, D. and Lubel, S., Psychophysics in the field: perception and memory of labor pain, Percept. Psychophys., 55 (1994) 133–141.

Babul, N. and Darke, A.C., Reliability and accuracy of memory for acute pain, Pain, 57 (1994) 131–132.

Beese, A. and Morley, S., Memory for acute pain experience is specifically inaccurate but generally reliable, Pain, 53 (1993) 183–189.

Bryant, R.A., Memory for pain and affect in chronic pain patients, Pain, 54 (1993) 347–351.

Chapman, C.R., Casey, K.L., Dubner, R., Foley, K.M., Gracely, R.H. and Reading, A.E., Pain measurement: an overview, Pain, 22 (1985) 1–31.

Corli, O., Grossi, E., Roma, G. and Battagliarin, G., Correlation between subjective labor pain and uterine contractions: a clinical study, Pain, 26 (1986) 53–60.

Eich, E., Reeves, J.L., Jaeger, B. and Graff-Radford, S.B., Memory for pain: relation between past and present intensity, Pain, 23 (1985) 375–379.

Erskine, A., Morley, S. and Pearce, S., Memory for pain: a review, Pain, 42 (1990) 255–265.

Fredrickson, B.L. and Kahneman, D., Duration neglect in retrospective evaluations of affective episodes, J. Pers. Soc. Psychol., 65 (1993) 45–55.

Gottman, J.M. and Levenson, R.W., A valid procedure for obtaining self-report of affect in marital interaction, J. Consult. Clin. Psychol., 53 (1985) 151–160.

Hunter, M., Philips, C. and Rachman, S., Memory for pain, Pain, 6 (1979) 35–46.

Kahneman, D., Fredrickson, B.L., Schreiber, C. and Redelmeier, D.A., When more pain is preferred to less: adding a better end, Psychol. Sci., 4 (1993) 401–405.

Kent, G., Memory of dental pain, Pain, 21 (1985) 187–194.

Lewis, B., Lewis, D. and Cumming, G., Frequent measurement of chronic pain: an electronic diary and empirical findings, Pain, 60 (1995) 341–347.

Linton, S.J. and Melin, L., The accuracy of remembering chronic pain, Pain, 13 (1982) 281–285.

Linton, S.J. and Gotestam, K.G., A clinical comparison of two pain scales: correlation, remembered chronic pain, and a measure of compliance, Pain, 17 (1983) 57–65.

Morley, S.J., Reliability and accuracy of memory for pain, Pain, 57 (1994) 132.

Pearce, S.A., Isherwood, S., Hrouda, D., Richardson, P.H., Erskine A. and Skinner, J., Memory and pain: tests of mood congruity and state dependent learning in experimentally induced clinical pain, Pain, 43 (1990) 187–193.

Price, D.D., McGrath, P.A., Rafii, A. and Buckingham, B., The validation of visual analogue scales as ratio scale measures for chronic and experimental pain, Pain, 17 (1982) 45–56.

Redelmeier, D.A., Rozin, P. and Kahneman, D., Understanding patients' decisions: cognitive and emotional perspectives, J. Am. Med. Assoc., 270 (1993) 72–76.

Rofé, Y. and Algom, D., Accuracy of remembering postdelivery pain, Percept. Motor Skills, 60 (1985) 99–105.

Salovey, P., Smith, A., Turk, D.C., Jobe, J.B. and Willis, G.B., The accuracy of memories for pain: not so bad most of the time, Am. Pain Soc. J., 2 (1993) 184–191.

Simon, H., Models of Man: Social and Rational, Wiley, New York, 1959.

Smith, W.B. and Safer, M.A., Effects of present pain level on recall of chronic pain and medication use, Pain, 55 (1993) 355–361.

Teasdale, J.D. and Fogarty, F.J., Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory, J. Abnorm. Psychol., 88 (1979) 248–257.

Varey, C. and Kahneman, D., Experiences extended across time: Evaluation of moments and episodes, J. Behav. Decision Making, 5 (1992) 169–185.