Published in final edited form as:

J Pain. 2011 February ; 12(2): 228–235. doi:10.1016/j.jpain.2010.07.001.

Peak and end effects in patients' daily recall of pain and fatigue: A within-subjects analysis

Stefan Schneider, Arthur A. Stone, Joseph E. Schwartz, and Joan E. Broderick

Department of Psychiatry & Behavioral Science Stony Brook University

Abstract

Clinical research often relies on retrospective recall of symptom levels, but the information contained in these ratings is not well understood. The “peak-and-end rule” suggests that the most intense (peak) and final (end) moments of an experience disproportionately influence retrospective judgments, which may bias self-reports of somatic symptoms. This study examined the extent to which peak and end symptom levels systematically affect patients' day-to-day recall of pain and fatigue. Rheumatology patients ($N = 97$) completed 5-6 momentary ratings of pain and fatigue per day as well as a daily recall rating of these symptoms for 28 consecutive days. For pain, peak and end momentary ratings predicted daily recall of average pain beyond the actual average of momentary ratings. This effect was small, yet was confirmed both in between-person and in within-person (repeated measures) analyses. For fatigue, neither peak nor end momentary symptoms significantly contributed to daily recall. Of note, the evidence for peak- and end-effects in recall of pain and fatigue varied significantly between individual patients. These findings suggest that peak- and end-effects create a small bias in recall reports of pain, but not fatigue. However, there are considerable individual differences in susceptibility to peak and end heuristics.

Keywords

pain; fatigue; momentary assessment; electronic diaries; patient reported outcomes; heuristics

Introduction

Patient-reported outcomes (PROs) of somatic symptoms are indispensable for research and clinical practice, but the information contained in these self-reports is not well-understood. Over the last 15 years, biases in retrospective reports have become a concern. Theoretical and empirical work has suggested that retrospective ratings are affected by the peak- and end- heuristics: when people are asked to recall their “average” or “usual” experiences, such reports may be disproportionately influenced by the most intense (“peak”) and the part of the experience (“end”) closest to the reporting time.^{9,14,19} To the extent that these heuristics move a recall rating off the actual average of experiences, they can threaten the content validity of self-reports, that is, the degree to which the measure reflects the intended domain of content.^{6,22}

© 2010 The American Pain Society. Published by Elsevier Inc. All rights reserved.

Corresponding author: Stefan Schneider, PhD Department of Psychiatry and Behavioral Science Putnam Hall, South Campus Stony Brook University Stony Brook, NY 11794-8790 Phone: 631-632-3114 Fax: 631-632-3165 Stefan.Schneider@sunysb.edu .

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Evidence of the peak-end heuristic for pain was initially documented under laboratory conditions involving relatively short and circumscribed events, such as medical procedures or cold pressor tasks.^{9,14,15} Subsequent studies have also found these heuristics to operate in recall of chronic and ongoing symptoms, but less reliably and with small effect sizes.^{7,8,19} Research on chronic symptom recall has mostly examined longer recall periods such as a week.^{8,19,21} It is important to understand how strongly these heuristics operate in shorter recall of chronic symptoms, given that the use of short recall periods has been recommended for clinical trials.²⁴ Here, we examine peak-end effects in one-day recall.

With few exceptions,¹³ studies on heuristics have been limited to pain. For clinical research, it is important to know if peak-end effects influence the recall of other symptoms. Fatigue symptoms are common to almost every major chronic illness,⁵ and are especially prevalent in chronic pain disorders.^{25,26} In this study, we extend previous research by examining peak-end effects for both chronic pain and fatigue.

To date, virtually all of the work has been based on between-person data. In the typical study, real-time momentary reports are collected over a single time-period and compared with a recall assessment. As with cross-sectional data, findings from this approach must be interpreted with caution, given that any effects of third variables (e.g., demographics, personality factors) that similarly affect both momentary and recall ratings may confound the results. In the present study, we sought to address this problem by using a within-person design. Patients recorded their momentary pain and fatigue several times a day and also recalled their symptoms at the end of each day for 4 weeks. If people rely on peak-end heuristics, we would expect that the amount of bias in any patient's recall systematically changes from day to day in accordance with the peak and end symptoms experienced during each day, which can only be tested using a within-person approach.

The within-person perspective also gave us the opportunity to examine the extent to which some patients characteristically may be more prone to peak-end effects than others. Literature on social cognition suggests that people differ in the degree to which they generally rely on heuristics.⁴ It would be instructive to know if this has implications for the recall of symptoms in particular. Thus, we examined if patients significantly differ in the use of peak-end heuristics and whether such individual differences are stable over time, as well as consistent within patients, such that those patients evidencing stronger use of peak-end heuristics for pain would do the same for fatigue.

To summarize, we hypothesized 1) that peak-end effects would be evident in daily recall of pain and fatigue, 2) that peak-end effects would be evident both using a between-person and a within-person approach, and 3) that there would be reliable individual differences in the use of peak-end heuristics.

Materials and Methods

Participants

Patients were recruited from two offices of a community rheumatology practice. Participants were required to be available for 30 consecutive days and to meet the following eligibility criteria: ≥ 18 years of age; physician-confirmed diagnosis of a chronic rheumatological illness; experienced symptoms of pain or fatigue during the last week; no significant sight, hearing, or writing impairment; fluency in English; normal sleep-wake schedule; ability to come to the research office twice within a month; had not participated in another electronic diary study in the last 5 years. A total of 279 patients were telephone screened, and 86 (31%) were excluded due to one or more of the above eligibility criteria. Of the 193 eligible

patients, 76 (39%) declined participation, and 117 (61%) participated. Eleven participants dropped out, and 106 completed the study.

Procedure

The study protocol was approved by the Stony Brook University Institutional Review Board. Participants provided informed consent and were compensated \$100. Data were collected from September 2005 through June 2006. Eligible patients came to the research office to complete demographic and questionnaire measures and to be trained in the use of an electronic diary (ED). Momentary and daily recall ratings of pain and fatigue intensity were collected for 29-31 days on a hand-held computer (Palm Zire 31). The ED utilized a software program provided by invivodata, inc. (Pittsburgh, PA) that featured auditory tones to signal the participant to complete a set of momentary ratings. It was programmed to generate an average of 7 randomly-scheduled (within intervals) prompts spread across the participant's waking hours (an average of one every 2 hours and 20 minutes, constrained to ensure a minimum of 30 minutes between prompts) determined by when the participant informed the ED that she was going to bed at night and set the wake up alarm the next morning. In addition to the random signals, the ED prompted the participant to complete a daily recall assessment at the time the ED was put to sleep at night, the "End of Day" assessment. A research assistant telephoned the patient 24 hours after the initial research office visit to answer any questions and troubleshoot potential problems with using the ED. A follow-up call was made once per week for the following three weeks to ensure the ED was working properly and to answer any questions. At the end of the month, patients returned the ED to the research office.

Measures

Items for this study were drawn from the Brief Pain Inventory (BPI)³ and the Brief Fatigue Inventory (BFI),¹² with wordings modified to correspond to the different reporting periods. For pain intensity, the item for momentary ratings was "Before the prompt: how intense was your bodily pain?", and the item for the daily recall was "What was the average level of your pain today?" For fatigue intensity, the item for momentary ratings was "Before the prompt: how fatigued (weary, tired) did you feel?", and the item for the daily recall was "What was the usual level of your fatigue today?" All ratings were made on a 100-point horizontal VAS with anchors "not at all" to "extremely".

Analytic strategy

This investigation is a secondary analysis of a study reported elsewhere.^{1,2} End-of-day (EOD) recall ratings of symptom intensity will be viewed as accurate to the extent that they correspond with the average of Ecological Momentary Assessments (mean-EMA) for that day. EOD reports will be viewed as biased by peak- and end-effects to the extent that they systematically reflect the highest (peak-EMA) and last (end-EMA) daily momentary assessment when statistically controlling for mean-EMA ratings. Accordingly, we used multiple regression analyses with EOD recall ratings as the outcome variable and with daily mean-, peak-, and end-EMA as predictor variables. Given that peak- and end-EMA tend to be highly correlated with mean-EMA, we centered them around the daily mean-EMA rating in order to reduce multicollinearity of predictors due to nonessential ill conditioning (i.e., as an artifact of scaling, see ¹⁰). This is consistent with the theoretical concept that peak- and end-EMA can only introduce bias in recall to the extent that they differ from the actual average of symptoms experienced.

Analysis of average peak-end effects

In order to conduct multiple regression analyses for between-person and within-person data, it is necessary to differentiate between-person and within-person (day-to-day) relationships among the study variables. We used multilevel modeling analyses (using the MIXED procedure in SAS, version 9.1) to achieve this. Specifically, we estimated a three-level multilevel model in which pain or fatigue intensity ratings (Level 1) were nested within days (Level 2), which, in turn, were nested within individuals (Level 3, see ¹¹). Mean-EMA, peak-EMA, end-EMA, and EOD recall ratings were treated as simultaneous (i.e., correlated) outcomes. To accomplish this in PROC MIXED, we structured the data such that the observations pertaining to all four outcomes were “stacked” on top of each other in a single variable. Dummy indicators were used to distinguish among the outcomes and to simultaneously estimate four intercept terms, one for each outcome. The NOINT option was used to prevent inclusion of the traditional, unnecessary, intercept term.^{17,23} In this model, the fixed effects represent the estimated means across all patients and days for each outcome variable. Level 3 random effects represent “trait-like” between-person variances and covariances among the outcomes. Level 2 random effects represent the within-person (day-to-day) variances and covariances among the outcomes. Consistent with analyses previously reported on this data,¹ Level 1 residual variances were estimated only for mean-EMA and represent “within-day” random error of measurement associated with the fallibility of momentary ratings in capturing the latent “true” mean-EMA score of each day (given that peak-EMA, end-EMA, and EOD recall ratings were represented by a single measurement at each day, these variables could not be adjusted for “within-day” measurement error). Using the full-information maximum likelihood method in mixed models allowed inclusion of all 97 subjects in the analyses despite some missing observations. This method has been referred to as “state of the art” for handling missing data.¹⁶ The between-person and within-person random effects matrices were then used in regular regression analyses to estimate between-person peak-end effects in daily EOD recall (pooled across all days) and within-person peak-end effects in day-to-day changes in EOD recall.

Analysis of individual differences in peak-end effects

Multilevel modeling was also used to explore individual differences in peak-end effects. Specifically, the daily peak-EMA, end-EMA, and (observed) mean-EMA were entered as predictors of daily EOD recall, and their effects (regression slopes) were allowed to vary randomly between individual patients (see ¹⁷). To examine whether individual differences in peak-end effects would be correlated across pain and fatigue --such that patients showing stronger effects for pain would do the same for fatigue -- a multivariate model was fit in which the variation (and covariation) of peak- and end-effects between patients was estimated simultaneously for pain and fatigue. To examine test-retest correlations of individual differences in peak-end effects, the 4-week assessment period was divided into two 2-week periods, and a multivariate model was fit in which the variation (and covariation) of peak- and end-effects between patients was estimated simultaneously for the two time periods.¹⁷ For all analyses, statistical significance was determined using a *p*-level of .05.

Results

Compliance with protocol rating schedule

The analysis sample included patients who met compliance criteria for momentary and recall assessments. In terms of momentary assessment compliance criteria, a minimum of 3 out of the 5 to 7 scheduled momentary reports per day were required, given that an insufficient number of momentary reports could fail to accurately capture a patient’s symptom experience for the day. We also required EOD reports and ≥ 3 momentary reports

for 24 out of 28 consecutive days for inclusion in the analyses. Out of the 106 study completers, 9 did not meet these criteria and were excluded, yielding an analysis sample of $n = 97$.

Out of these, 73 (75%) provided acceptable data for all 28 days, 12 (12%) had data for 27 days, 6 (6%) had data for 26 days, 2 (2%) had data for 25 days, and 4 (4%) had data for 24 days. Across all patients and days (2716 reporting days), 35 days (1.3%) were missing because less than 3 momentary reports were completed, and 20 (0.7%) EOD assessments were missing. Patients provided a mean of 5.5 ($SD = 1.23$) momentary assessments each day with 80% of days having 5 or more momentary reports. The EOD assessment was completed on average at 10:35 pm ($SD = 1.43$ hrs); the last momentary report was completed on average 1.35 hours ($SD = 1.16$) before the EOD recall and within 2 hours before the EOD recall on 76% of all days.

Sample characteristics

The most prevalent diagnoses in the sample were osteoarthritis (50%), rheumatoid arthritis (28%), lupus (17%) and fibromyalgia (10%). Participants had a mean age of 56 years (range 28 – 88, $SD = 11.1$), and were predominantly female (87%), married (67%), and White (91%). Most were high school graduates (97%), with 70% having completed some college. One-fourth (25%) of the sample were currently on disability.

Means and variance components of ratings

Mean ratings of the study variables across all participants and days are shown in Table 1. EOD recall ratings exceeded corresponding daily mean-EMA ratings by 5 points ($p < .001$) for pain, and by 4 points ($p < .001$) for fatigue on the 100-point scale. Peak-EMA ratings exceeded the daily mean-EMA on average by 14 points ($p < .001$) for pain, and by 18 points ($p < .001$) for fatigue. End-EMA ratings also exceeded the daily mean on average by 2 points ($p < .001$) for pain and by 8 ($p < .001$) points for fatigue, evidencing a diurnal cycle for ratings of fatigue (and to a lesser extent for pain), with higher symptom levels in the evening.

Examination of linear trends over the 4-week study period showed no significant changes on average for EOD recall, mean-EMA, or end-EMA. However, peak-EMA ratings on average declined for pain by 1.22 points per week ($SE = .22$; $p < .001$) and for fatigue by 1.53 points per week ($SE = .22$; $p < .001$), indicating possible effects of repeated testing reflected in peak-EMA ratings. Day of measurement (as a categorical variable) was entered as a covariate in the further analyses to control for this.

Inspection of between-person and within-person sources of variance showed that there was substantial within-person, day-to-day variability ($p < .0001$) for all study variables (see Table 1). For EOD recall and mean-EMA ratings, the proportion of within-person variance ranged between 20% and 43%, indicating a moderate amount of day-to-day changes for these variables. For peak- and end-EMA scores, the proportion of within-person variance consistently exceeded 60%; changes in peak and end symptoms from day-to-day were at least as pronounced as differences between people. Thus, both between- and within-person regression analyses were warranted.

Between-person estimates of average peak-end effects in EOD recall

Results for the between-person regression analyses predicting EOD reports are shown in Table 2 (upper panel). For pain, mean-EMA ratings entered at step 1 explained 84% of the variance in EOD recall. Peak- and end-EMA entered together at step 2 explained a small but

significant 2% variance increment, with both peak- and end-EMA showing significant unique contributions to the prediction of EOD recall (see Table 2).

In the between-person regression for fatigue, mean-EMA ratings entered at step 1 explained 76% of the variance in EOD recall. Neither peak- nor end-EMA ratings entered at step 2 significantly added to the prediction of fatigue EOD recall.

Within-person estimates of average peak-end effects in EOD recall

Results for the within-person regression analyses are shown in Table 2 (lower panel). For pain, mean-EMA ratings entered at step 1 explained 63% of the variance in EOD recall. Paralleling the between-person regression results, peak- and end-EMA, entered together at step 2, explained a significant 3% variance increment, with both peak- and end-EMA showing significant unique contributions to the prediction of EOD recall.

In the within-person regression for fatigue, mean-EMA ratings entered at step 1 explained 52% of the variance in EOD recall. Peak- and end-EMA ratings entered together at step 2 did not significantly add to the prediction of fatigue EOD recall.

Evidence for individual differences in peak-end effects in EOD recall

Whereas the previous within-person analyses tested evidence for peak-end effects in general (i.e., for the “average patient”), we subsequently examined whether the regression coefficients of peak- and end-EMA predicting EOD recall reliably differed across patients.

Individual differences in peak effects

The estimated regression coefficients of peak-EMA predicting EOD recall evidenced highly significant variance across patients for pain ($SD = .25$; $z = 5.03$, $p < .0001$) and for fatigue ($SD = .26$; $z = 4.17$, $p < .0001$). Figure 1 shows the predicted regression slopes for the average patient, for patients with a high effect of peak-EMA (with a regression coefficient 1 SD above the average coefficient), and for patients with a low effect of peak-EMA (i.e., with a regression coefficient 1 SD below the average coefficient) to illustrate the magnitude of these individual differences of peak-effects in EOD recall.

Individual differences in the effects of peak-EMA on recall were significantly correlated between pain and fatigue, $r = .32$ ($z = 2.04$, $p = .04$), suggesting that patients evidencing stronger use of the peak heuristic in the recall of pain also tended to do so for fatigue. Moreover, the test-retest correlations of the effects of peak-EMA on EOD recall were highly significant at $r = .59$ ($z = 3.33$, $p < .001$) for pain and $r = .71$ ($z = 3.05$, $p < .01$) for fatigue, suggesting that individual differences in the use of the peak-heuristic were relatively stable from one 2-week assessment period to the next.

Individual differences in end effects

We also found significant individual differences in the effects of end-EMA predicting EOD reports of both pain ($SD = .13$, $z = 2.57$, $p < .01$) and fatigue ($SD = .14$, $z = 2.13$, $p = .02$), but the magnitude of these individual differences was somewhat less pronounced than that for peak-effects (see Figure 1).

Individual differences in the effects of end-EMA on recall were not significantly correlated between pain and fatigue, $r = .06$ ($z = .19$, $p = .85$). Moreover, the test-retest correlations of the effects of end-EMA on EOD recall were not significant for pain ($r = .11$, $z = 0.39$, $p = .70$) or for fatigue ($r = .05$, $z = .17$, $p = .86$). This suggested that individual differences in the use of the end heuristic were not consistent across pain and fatigue and not stable over time.

Differences of peak-end effects by diagnosis

Finally, we explored whether individual differences in peak-end effects could be explained by different diagnoses of patients. If this was the case, then the noncorresponding results of average peak-end effects for pain and for fatigue could potentially be attributable to the particular composition of diagnoses in the study sample. Type of medical diagnosis did not significantly predict individual differences in peak-effects ($p = .52$) or end-effects ($p = .38$) of pain, nor did it predict individual differences in peak-effects ($p = .70$) or end-effects ($p = .14$) of fatigue.

Discussion

The peak-and-end heuristic was first observed under controlled short-term laboratory conditions.^{9,14} The extent to which it introduces bias when people are asked to recall the average intensity of chronic and ongoing symptoms is less well understood, and two previous studies of patients with chronic pain conditions yielded modest effects,^{7,19} whereas another showed no influence of peaks.²¹ These studies have all evaluated the peak-and-end heuristic from a between-person perspective. Between-person analyses cannot answer the question posed in this paper: are changes in peaks (or ends) from one period to another systematically linked with changes in recall *within* a given person? Demonstrating this effect would increase our confidence that the results are not explained by confounding third variables, such as person characteristics similarly affecting both momentary and recall reports. We found considerable within-person variation in peaks and ends of pain and fatigue, which allowed us to test the hypothesis that the peak-end heuristic operates within individuals.

For recall of pain, we found evidence for both peak- and end-effects in between- and within-person analyses. Patients commonly rated average pain for the day higher on those days that were characterized by more intense peak pain and that ended on a more painful note than on other days, after accounting for the actual average of pain intensity. This finding confirms that peak and end pain levels play a role in recall of chronic pain not only in laboratory studies, but also in reports of clinical symptoms with a recall period of a day. The results are in accord with Jensen's (2008) finding that 2-4% of the variance of end-of-day recall of pain was explained by peaks and ends; however, those analyses were between-person.⁷ Our within-person analyses showed a statistically reliable effect of 3% of explained within-person variance, comparable to Jensen's results. The effect is considerably lower in magnitude than one study of the peak-end effect for 7-day recall of pain,¹⁹ where 8 to 11% of the variance was explained, but is higher than a recent study of the peak effect,²¹ where peaks were not significantly associated with one-week recall.

Contrary to our hypothesis, we found no evidence that peak-end effects on average affect recall for fatigue using either between- or within-person analyses. This is surprising given that the correspondence between mean-EMA ratings and EOD ratings was somewhat less pronounced for fatigue than for pain, suggesting that heuristics might potentially play a larger role in fatigue recall than in pain recall. Also, the finding cannot be explained by a lack of variability in peak and end ratings of fatigue, which would limit the maximum association possible. Variances between- and within-subjects were pronounced for fatigue and even exceeded those for pain. In addition, we found no evidence to suggest that the different findings for pain and fatigue could be explained by the constellation of medical diagnoses in the sample. It is particularly interesting that the "end" heuristic did not impact the EOD rating, since fatigue has a predictable cycle across the day with higher levels in the evening.²⁰ In view of the pronounced diurnal cycle of fatigue, it may be argued that daily peak and end instances of fatigue were largely redundant, such that their influences canceled each other out in multiple regression analyses. However, even though the correlations

between peak and end symptoms were higher for fatigue (r s of .49 in between- and of .41 in within-person analyses) than for pain (r s of .15 in between- and of .17 in within-person analyses), they were still sufficiently distinct from each other to potentially allow unique contributions to EOD recall. Symptoms of fatigue may be more diffuse and less discernable than pain, which is often localized and salient. It is also possible that recall ratings of fatigue evoke more complex considerations that go beyond symptom experiences at single moments. For example, patients may rate fatigue on an EOD report as higher if they find that they were unable to accomplish all of the tasks for the day, rather than basing their ratings purely on discrete symptom intensities.

Although our results suggest that peak-end effects operates in most patients for daily ratings of pain (but not for fatigue), these results can also be viewed from a pragmatic perspective. Specifically, do peak-end effects threaten the validity of patient reported outcomes based upon end-of-day ratings for pain? We find ourselves in agreement with Jensen (2008); even with slightly stronger between-person effects of peak-end than we observed, he concluded that "... if the target of treatment is a patient's usual or average pain over a period of time, and that time period is 24 h, then these findings suggest that single ratings of recalled pain intensity appear to have adequate validity for this purpose (p. 425)."⁷ This conclusion also reinforces a recently published finding from the dataset used in this paper showing that the average of only 4-5 EOD pain ratings was highly correlated with 42 momentary pain rating taken over a week, suggesting this as a technique for assessing weekly pain experience.¹

A novel feature of this study was determining if patients differed reliably and consistently in usage of heuristics, which was made possible by the within-person approach. Our hypothesis was partially confirmed. Individual differences in the peak-effect were pronounced, consistent across pain and fatigue recall, and relatively stable over time. In comparison, individual differences in end-effects were less pronounced, not consistent across symptoms, and not reliable/stable. Thus, even if the influence of the peak heuristic is limited for patients in general, it may be more pronounced or only exist for a selected subset of patients. For those patients evidencing a comparatively high peak-effect, each 10-point increase in peak symptoms over the day's mean was associated with an inflation in recall of about 4 points – more than twice as much as what was found for the average study participant. A person's tendency to rely on the peak heuristic may be grounded in a relatively stable and general disposition, and future research may wish to identify factors underlying this tendency. For example, it may be promising to examine if patients who characteristically process information in an intuitive mode of thinking are more susceptible to the peak-effect than those who operate primarily in an analytical mode, factors that have previously been shown to influence people's proneness to fallacies in decision making.^{4,18}

This study has several limitations. First, the study focused on daily recall. Thus, the results have potentially narrow applicability, and the generalizability to other recall periods is questionable. It is possible that effects of heuristics become more pronounced for extended periods such as a week or longer, especially considering that the accuracy of recall has been found to be reduced for longer reporting periods.² However, the few studies looking at peak and end heuristics with 7-day recall also found small to moderate effects.^{8,19,21} Other research has found that a composite of usual and least pain recall ratings is more closely associated with actual average pain intensity than recalled usual pain, a possibility that was not considered here.⁸

Second, this observational study was based on naturally occurring day-to-day changes in pain and fatigue intensity. The results may be different in settings involving systematic changes in symptom levels over time, such as in the context of a clinical trial.

Third, these data may not generalize to symptoms other than pain and fatigue or to other populations. All patients in this study suffered from chronic rheumatological conditions characterized by pain and fatigue. It is unclear whether the same results would be obtained in populations reporting other patient reported outcomes.

Finally, features of the study design may have limited the potential impact of peak- and end-effects. Assessing moments only 5-6 times per day may have missed some actual peaks in symptom intensity and many of the last daily momentary assessments were not obtained at the exact end of the day and immediately before EOD reports. However, when more stringent daily inclusion criteria were applied (i.e., limiting the analyses to days with 4 or more assessments or to days at which the last EMA rating was taken no less than 2 hours before the EOD report; data not shown), the results and conclusions did not change.

In summary, the results of this study suggest that peak and end experiences generally have a small, yet significant, influence on daily recall of pain, but not for fatigue. From a basic science perspective, these results add to the literature by demonstrating that fluctuations in peak and end pain from one day to the next may change retrospective impressions of daily pain experience. From an applied perspective the magnitude of these effects does not seem to threaten the validity of end-of-day pain and fatigue assessments. Furthermore, we found evidence that these heuristics do not operate uniformly across all individuals, and are reliably observed only in a subset of patients. It may be important to identify for which types of patients and under which conditions these heuristics do and do not pose a threat to the validity of symptom reports.

Perspective

The peak-end cognitive heuristic could bias end-of-day recall of pain and fatigue. An effect was shown for pain, but not for fatigue. The effects were small and were unlikely to substantially bias end-of-day assessments. Individuals were shown to differ in the degree that the heuristic was associated with recall.

Acknowledgments

This research was supported by grants from the National Institutes of Health (1 U01-AR052170-01; Arthur A. Stone, principal investigator), (R01 AR054626; Joan E. Broderick, principal investigator) and by GCRC Grant M01-RR10710 from the National Center for Research Resources. We would like to thank Pamela Calvanese, Doerte Junghaenel, and Leighann Litcher-Kelly for their assistance in collecting data. Software and data management services for the electronic diary assessments were provided by invivodata, inc (Pittsburgh, PA). JEB and AAS have a financial interest in invivodata, inc. and AAS is a senior scientist for the Gallup Organization.

References

1. Broderick JE, Schwartz JE, Schneider S, Stone AA. Can end-of-day reports replace momentary assessment of pain and fatigue? *J Pain* 2009;10:274–281. [PubMed: 19070550]
2. Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain* 2008;139:146–157. [PubMed: 18455312]
3. Daut RL, Cleeland CS. The prevalence and severity of pain in cancer. *Cancer* 1982;50:1913–1918. [PubMed: 7116316]
4. Epstein S, Pacini R, Denes-Raj V, Heier H. Individual differences in intuitive-experiential and analytical-rational thinking styles. *J Pers Soc Psychol* 1996;71:390–405. [PubMed: 8765488]
5. Evans EJ, Wickstrom B. Subjective fatigue and self-care in individuals with chronic illness. *Medsurg Nurs* 1999;8:363–369. [PubMed: 11000775]

6. Gendreau M, Hufford MR, Stone AA. Measuring clinical pain in chronic widespread pain: selected methodological issues. *Best Pract Res Clin Rheumatol* 2003;17:575–592. [PubMed: 12849713]
7. Jensen MP, Mardekian J, Lakshminarayanan M, Boye ME. Validity of 24-h recall ratings of pain severity: biasing effects of “peak” and “end” pain. *Pain* 2008;137:422–427. [PubMed: 18035495]
8. Jensen MP, Turner LR, Turner JA, Romano JM. The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain* 1996;67:35–40. [PubMed: 8895229]
9. Kahneman D, Fredrickson BL, Schreiber CA, Redelmeier DA. When more pain is preferred to less: Adding a better end. *Psychological Science* 1993;4:401–405.
10. Kreft IGG, de Leeuw J, Aiken LS. The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research* 1995;30:1–21.
11. Lambert PC, Burton PR, Abrams KR, Brooke AM. The analysis of peak expiratory flow data using a three-level hierarchical model. *Stat Med* 2004;23:3821–3839. [PubMed: 15580603]
12. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, Huber SL. The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer* 1999;85:1186–1196. [PubMed: 10091805]
13. Miron-Shatz T. Evaluating multiepisodic events: Boundary conditions for the peak-end rule. *Emotion* 2009;9:206–213. [PubMed: 19348533]
14. Redelmeier DA, Kahneman D. Patients’ memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain* 1996;66:3–8. [PubMed: 8857625]
15. Redelmeier DA, Katz J, Kahneman D. Memories of colonoscopy: A randomized trial. *Pain* 2003;104:187–194. [PubMed: 12855328]
16. Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychol Methods* 2002;7:147–177. [PubMed: 12090408]
17. Schwartz, JE.; Stone, AA. The analysis of real-time momentary data: a practical guide. In: Stone, AA.; Shiffman, S.; Atienza, AA.; Nebeling, L., editors. *The science of real-time data capture: self-reports in health research*. Oxford University Press; New York: 2007. p. 76-116.
18. Smith SM, Levin IP. Need for cognition and choice framing effects. *Journal of Behavioral Decision Making* 1996;9:283–290.
19. Stone AA, Broderick JE, Kaell AT, DelesPaul PA, Porter LE. Does the peak-end phenomenon observed in laboratory pain studies apply to real-world pain in rheumatoid arthritis? *J Pain* 2000;1:212–217. [PubMed: 14622620]
20. Stone AA, Broderick JE, Porter LS, Kaell AT. The experience of rheumatoid arthritis pain and fatigue: Examining momentary reports and correlates over one week. *Arthritis Care Res* 1997;10:185–193. [PubMed: 9335630]
21. Stone AA, Schwartz JE, Broderick JE, Shiffman SS. Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. *Pers Soc Psychol Bull* 2005;31:1340–1346. [PubMed: 16143666]
22. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. *Ann Behav Med* 1994;16:199–202.
23. Thiébaud R, Jacqmin-Gadda H, Chêne G, Leport C, Commenges D. Bivariate linear mixed models using SAS proc MIXED. *Comput Methods Programs Biomed* 2002;69:249–256. [PubMed: 12204452]
24. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims. [Accessed March 22]. 2010 Available at <http://www.fda.gov>
25. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. *J Rheumatol* 1996;23:1407–1417. [PubMed: 8856621]
26. Zautra AJ, Fasman R, Parish BP, Davis MC. Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. *Pain* 2007;128:128–135. [PubMed: 17055648]

Figure 1.

Regression estimates of peak and end symptoms predicting of end of day recall for patients 1 *SD* below the mean, at the mean, and 1 *SD* above the mean of the distribution of individual regression slopes, representing patients with low, average, and high peak/end effects in recall. EOD = End of day recall; EMA = Ecological momentary assessment. Values for peak-EMA and end-EMA are based on deviations from daily mean-EMA. Predicted EOD recall ratings are based on partial regression estimates in multiple regressions including daily mean-, peak-, and end-EMA ratings as independent variables, obtained from random effects coefficients in multi-level models.

Table 1

Means (standard errors) and variance components of study variables

	EOD	Mean-EMA	Peak-EMA^a	End-EMA^a
Pain				
Mean (SE)	46.5 (2.2)	41.6 (2.1)	13.9 (0.6)	2.3 (0.5)
Total variance	700.0	530.2	109.1	161.3
Within-person variance	238.6	108.1	79.2	141.2
Percent within-person	34.1 %	20.4 %	72.6 %	87.6 %
Fatigue				
Mean (SE)	50.1 (2.1)	46.2 (2.0)	17.5 (0.8)	8.1 (0.9)
Total variance	702.6	466.8	152.1	240.7
Within-person variance	299.8	98.6	96.8	174.7
Percent within-person	42.7 %	21.1 %	63.6 %	72.6 %

Note. EOD = End of day recall, EMA = Ecological momentary assessment, both assessed on 100-point visual analogue scales. Values were obtained from multilevel models. Within-person variance components are statistically significant ($p < .0001$) for all variables.

^aThe peak-EMA and end-EMA variables represent deviations from daily mean-EMA.

Table 2

Hierarchical multiple regressions predicting end-of-day recall from daily mean, peak and end momentary assessments

Between-subject		b	SE(b)	R ²	Δ R ²
Pain					
	Step 1			.84	.84***
		Mean-EMA	1.07***	.01	
	Step 2			.86	.02***
		Peak-EMA	.30***	.03	
		End-EMA	.10***	.02	
Fatigue					
	Step 1			.76	.76***
		Mean-EMA	1.07***	.02	
	Step 2			.76	.00
		Peak-EMA	.05	.03	
		End-EMA	.05	.03	
Within-subject		b	SE(b)	R ²	Δ R ²
Pain					
	Step 1			.63	.63***
		Mean-EMA	1.19***	.02	
	Step 2			.66	.03***
		Peak-EMA	.19***	.03	
		End-EMA	.14***	.02	
Fatigue					
	Step 1			.52	.52***
		Mean-EMA	1.28***	.03	
	Step 2			.52	.00
		Peak-EMA	.07	.04	
		End-EMA	.05	.03	

Note. EOD = End of day recall; EMA = Ecological momentary assessment.

 $p < .001$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript