

# Lay intuitions about overall evaluations of experiences

Irina Cojuharenco\*

Department of Economics and Business  
Universitat Pompeu Fabra

## Abstract

Previous research has identified important determinants of overall evaluations for experiences lived across time. By means of a novel guessing task, I study what decision-makers themselves consider important. As *Informants*, some participants live and evaluate an experience. As *Guessers*, others have to infer its overall evaluation by asking Informants questions. I rewarded accurate inferences, and analyzed and classified the questions in four experiments involving auditory, gustatory and viewing experiences. Results show that Guessers thought of overall evaluations as reflecting average momentary impressions. Moreover and alternatively, they tended to consider the personality and attitudes of the experiencing person, experience-specific holistic judgments and behavioral intentions regarding the experience. Thus, according to lay intuitions, overall evaluations are more than a reflection of the experience's momentary impressions.

Keywords: lay intuitions, extended experiences, remembered utility.

## 1 Introduction

People often report experiences by expressing a number on a scale. Someone might say, “7 out of 10 for this concert”, or “In terms of painfulness, I rate this medical procedure as 90 out of 100.” Such overall evaluations of experiences have been shown to be important decision inputs (Wirtz et al., 2003; Oishi & Sullivan, 2006), and studied extensively.

Kahneman, Wakker, and Sarin (1997) suggested that experiences can be represented as intensity profiles of pleasure (or discomfort) over bounded intervals of time, i.e., time profiles of “experienced utility.” Experiments and field studies have shown that people evaluate more positively experiences with increasing, rather than decreasing time profiles at equivalent levels of total pleasure experienced (Ariely & Carmon, 2003). There is a preference for steeper rates of improvement (Hsee & Abelson, 1991), as well as variability in experience (Read, Loewenstein, & Rabin, 1999). Finally, the “Peak-End rule” finding suggests that overall evaluations are best predicted by only two moments of the experience: the most pleasant/unpleasant and final (Kahneman, 2000). Kahneman, Wakker and Sarin (1997) present a set of assumptions about experiences explaining why integration/summation of all moments would be correct from

a normative point of view. Life satisfaction researchers and psychologists, on the other hand, explore alternative paradigms and study the role of personality and the beliefs of the evaluating person for overall evaluations (Updegraff, Gable, & Taylor, 2004; Robinson, 2002; Trope & Liberman, 2003; Brendl & Higgins, 1995).

In contrast to previous research, the present work aims to reveal what decision makers themselves draw on as they think about overall evaluations of experiences. I will compare lay intuitions to what researchers have considered. This comparison may further enrich theories of overall evaluations, and suggest ways of testing them.

I employ a novel method, the guessing task, in order to elicit lay intuitions. The philosophy of the method is that of Active Information Search, a method of naturalistic decision-making that Huber, Wider, and Huber (1997) proposed for the study of risky choice. The method consists of giving participants a minimal description of a decision problem and allowing them to seek information. I report experiments in which participants had to guess the overall evaluation of an experience lived by another person. Active information search was allowed prior to making the guess. The information participants sought was taken to reveal lay intuitions about the target overall evaluation.

## 2 General Method

### 2.1 The Guessing Task

Participants were assigned randomly to be *Informants* or *Guessers*.

\*I am most grateful to Robin Hogarth for encouragement and continuous advice about this project. I am also grateful to colleagues and the anonymous reviewers who have read earlier drafts and provided valuable comments. Special thanks to Petya Platikanova and Aniol Llorente for help with experiments. Address: Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain, E-mail: irina.cojuharenco@upf.edu.

**Informant's task.** Informants lived and evaluated a certain experience. Their evaluations were unknown to Guessers. These were ratings on 0 to 100 point scale anchored by statements about experienced pleasure or discomfort. Ratings were real-time and overall. Real-time ratings ranged from 0, "Not pleasant at all", to 100, "Very pleasant", and overall from 0, "I experienced no pleasure at all" to 100, "I experienced a great deal of pleasure". For example, if an Informant listened to several musical performances, he/she evaluated each performance immediately after hearing it and the musical sequence overall. If an Informant tasted pieces of chocolate, he/she rated each piece and then rated the whole tasting session. If an Informant viewed affective images, he/she rated each image and then the experience of viewing the whole series.<sup>1</sup> Informants wrote on evaluation sheets distributed to them prior to the experience.

**Guesser's task.** A Guesser was a participant who faced the task of guessing the overall rating that an Informant gave to his/her experience. Guessers could not communicate with each other. They knew the class of stimuli experienced by the Informant, but not the duration of the experience (e.g., that the Informant had listened to musical performances, had tasted chocolate samples, etc.); and they knew that the Informant rated the experience in real-time and overall using 0 to 100 point scales with anchoring statements "Not pleasant at all"/"I experienced no pleasure at all" and "Very pleasant"/"I experienced a great deal of pleasure".

Guessers could ask Informants questions. They were instructed to refrain from judging the appropriateness of questions.<sup>2</sup> Questions had to be written down and could be asked simultaneously or sequentially.

## 2.2 Closed-Format Questionnaires

Closed-format questionnaires complemented the guessing task. Questionnaire A was designed prior to Experiment 1, and contained questions about the experience of the Informant inspired by the "time profile" perspective; that is, the items were questions about real-time ratings and statistics of these. Questionnaire B was designed after Experiment 1 for participants of subsequent experiments. The items were exemplars of question categories observed in Experiment 1. Each Guesser faced a differ-

ent order of items in each questionnaire. Guessers were asked to pick three questions among questionnaire items that would be most useful in the guessing task, and underline the most informative one.

## 3 Experiment 1

### 3.1 Method

#### 3.1.1 Stimuli

Informants were exposed to either a short or a long auditory experience, consisting, respectively, of two or six Moldovan folk music performances. Each performance was 2-4 minutes long. Informants had earphones. Windows Media player was used to reproduce the music.

#### 3.1.2 Participants

There were 54 participants in Experiment 1; 22 were male, average age was 22. All were undergraduates, and the vast majority were students in economics. 18 participants acted as Guessers, of which 6 were male, average age was 21.

#### 3.1.3 Procedure

Informants were assigned randomly to 2 versus 4 performances. The number of questions that a Guesser could ask was limited to 3 when the Informant had listened to 2 performances, and to 5, when the Informant had listened to 4 performances.

Informants and Guessers were paid 2 euros for participating. For each guess where the error was less than 5 points the Guesser received 5 euros.

There were 18 sessions in total, each involving a Guesser and two Informants (one Informant was involved in a separate task, and his role is not discussed in this article). In each session, the Informant had evaluated his/her experience prior to the arrival of the Guesser. The Guesser wrote down his/her questions. The author passed the questions of the Guesser to the Informant and delivered answers back to the Guesser. Once the Informant had answered all questions, the Guesser made his/her guess. Guessers wrote down a comment explaining the guessing strategy used in a short post-guess questionnaire. They wrote down additionally the question they would have asked had they been constrained to a single question, and completed questionnaire A. Finally, they received performance feedback.

#### 3.1.4 Classification and coding of questions

I content analyzed the questions in view of the perspectives on experiences that the questions reflected. For ex-

<sup>1</sup>In the experiment involving the viewing of aversive images, the anchoring statements for real-time and overall ratings referred to discomfort instead of pleasure

<sup>2</sup>Direct questions of the type "What was your overall evaluation?" or "Was your overall rating below 50?" were not transmitted to the Informants, and Guessers were asked to formulate a different question. If attempted repeatedly, such questions were allowed finally, but a special classification category was created for them. Questions of this category did not exceed 9% of total questions in any of the experiments/conditions.

ample, questions involving real-time ratings were consistent with the “time profile” perspective on experiences discussed earlier. Other perspectives involved the Informant’s personality, the category of the experience, other holistic attributes and judgments, as well as perceived behavioral implications. Classification codes for question types within each perspective, and examples, were formulated for use by independent coders (there were 14 codes). Finally, one female graduate student in clinical psychology and one male graduate student in economics coded the questions. There were no disagreements. In what follows, I present the resulting classification of questions, discussing the codes within 5 broad categories: “time profile” perspective, holistic attributes and judgments [holistic A/J], Informant’s personality, decision rule, and behavioral implications.

**Category 1: “time profile” perspective.** The first category comprised all questions inquiring for real-time ratings and any statistics of these [ratings stats]. These were questions of the type: “How did you rate the musical performances of this sequence?” or “How did you rate the performance you liked the best?”. Importantly, there was not a prevalence of questions about the maximum/minimum or final ratings. Guessers also asked for the average and modal ratings, the trend, the slope, and the variance of ratings.

**Category 2: holistic attributes and judgments [holistic A/J].** The second category included questions such as, “What was, or, how much did you like the rhythm of the music you listened to?”, “Was the music you heard classical?”, “Was your experience with music similar to your experience in a philosophy lecture?”, or “Did you feel tender emotion as you listened to this music?”. As the latter questions suggest, specific emotions or the category of the experience indicated the experience’s overall evaluation. Notable are questions that referred to holistic attributes of the experience identifiable only in retrospect (i.e., the overall rhythm of the music).

**Category 3: personality.** The third category included questions related to the personality of the Informant. Guessers asked about social status, general knowledge and culture, as well as enduring psychological dispositions. For example, “Are you a person who likes variety?”, “Are you a generally depressed individual?”.

**Category 4: decision rule.** The fourth category involved inquiries about the decision rule underlying the overall rating, for example, “Did you rate the experience overall based on the fact that you are generally fond of music or rather based on your actual experience with these pieces of music (that is your overall rating was equal to the average of piece ratings)?”.

**Category 5: behavioral implications.** The fifth category comprised questions, which explored the implications of a given overall rating for the experience’s fu-

ture use, willingness-to-pay (WTP) for it, what purpose the experience could serve and how useful it would be. For example, guessers asked “Could you use this music as a background for a romantic dinner?”, or “How often would you listen to this music if you had it at home?”.

Table 1 reports both the proportion of a question category in total questions asked (TQ, an idea’s “persistence”) and the proportion of participants asking questions of a particular category (AP, an idea’s “spread”). A given Guesser often asked questions pertaining to different categories. Therefore, I calculated proportions of participants asking a particular combination of question categories, and report these additionally. Table 2 reports the structure of single questions that Guessers wrote down.

### 3.2 Results and discussion

Guessers formulated a total of 66 questions. Note, that the instructions could direct their thinking towards a “time profile” perspective on experiences (they were told explicitly that Informants had rated their experiences in real-time and overall, and could think that the two types of ratings had to be related). Importantly, research adopting a “time profile” perspective provides an indication of how to use real-time ratings for predicting overall evaluations. One strategy is to compute the average of the most extreme and final real-time rating, and another — to compute the over-all average (Kahneman, 2000; Ariely & Zauberman, 2000; Langer, Sarin, & Weber, 2005). The results of the guessing task show that the first strategy was not intuited by Guessers. The second was pursued in some cases. It was preferred to other strategies involving real-time ratings, as closed-format questionnaires also showed (see Table 3). Duration of the experience was rarely a matter of concern to the Guessers.

Although the largest proportion of questions in total questions asked revealed a “time profile” perspective, questions involving holistic A/J were equally important in spread, i.e. in terms of the proportion of participants asking at least one question of the kind. In addition, Guessers asked frequently about behavioral implications of the overall rating and the personality of the Informant. Table 2 describes the structure of single questions formulated by Guessers, and shows that most question categories remained represented in a similar order of importance.

The analysis of how participants combined frames of analysis reveals that most of them asked questions pertaining to at least two question categories (66%). Most frequent types of combinations involved the “time profile” perspective and holistic A/J, or the latter and behavioral implications. 22% maintained a “time profile” perspective.

Table 1: Content structure of guessers' multiple questions: proportion of total questions (TQ, %), and proportion of participants asking the question of a particular type (AP, %).

Category and codes	Experiment 1 (music)		Experiment 2 (chocolate)		Experiment 3 (pleasant images)		Experiment 4 (aversive images)	
	TQ	AP	TQ	AP	TQ	AP	TQ	AP
<b>“Time profile” persp.</b>	<b>39</b>	<b>61</b>	<b>18</b>	<b>44</b>	<b>15</b>	<b>33</b>	<b>16</b>	<b>27</b>
– Ratings stats	20	44	7	13	8	21	16	27
– Real-time ratings	14	28	4	13	0	0	0	0
– Duration	5	17	7	22	7	17	0	0
<b>Holistic A/J</b>	<b>29</b>	<b>61</b>	<b>38</b>	<b>74</b>	<b>39</b>	<b>67</b>	<b>47</b>	<b>100</b>
– Attributes	5	11	7	13	15	33	17	42
– Attribute evaluations	3	11	1	4	1	4	3	8
– Experience category	11	39	4	13	6	13	9	27
– Category liking	5	11	25	61	3	8	0	0
– Emotions	5	17	1	4	14	38	18	46
<b>Behavioral implications</b>	<b>19</b>	<b>39</b>	<b>20</b>	<b>44</b>	<b>26</b>	<b>63</b>	<b>20</b>	<b>54</b>
– Future use	8	17	15	30	10	29	8	23
– Instrumentality	8	17	4	13	13	29	7	15
– Approach/avoidance	0	0	1	4	3	8	4	12
– WTP	3	11	0	0	0	4	1	4
<b>Personality</b>	<b>14</b>	<b>22</b>	<b>6</b>	<b>17</b>	<b>10</b>	<b>25</b>	<b>7</b>	<b>19</b>
<b>Decision rule</b>	<b>2</b>	<b>6</b>	<b>6</b>	<b>9</b>	<b>1</b>	<b>4</b>	<b>1</b>	<b>4</b>
<i>Non-classifiable</i>	2	–	9	–	8	–	9	–
Total /All	N=66	N=18	N=68	N=23	N=72	N=24	N=76	N=26

Questions Guessers asked helped them make 10 successful guesses in 18 attempts. If they had pursued the strategy of averaging across particular, or all real-time ratings, as described above, the success rate would have been 13 in 18.

## 4 Experiments 2, 3 and 4

One could argue that a musical experience is different from other hedonic experiences, such as food tasting or pain. I report replications involving tasting chocolate, and image-viewing experiments, which allowed experimentation with both pleasant and aversive stimuli.

### 4.1 Method

#### 4.1.1 Stimuli

In Experiment 2 Informants tasted two or six pieces of chocolate. Pieces were small portions of white, black, milk, baking, liquor filling, and nuts and raisins chocolate. Each piece was presented to the taster on a napkin and covered by a napkin with the number indicating its order of tasting and no other information.

Two and 15 pleasant, and two and 15 aversive images were the stimuli for viewing experiences in Experiments 3 and 4 respectively (Lang et al., 2005). Images appeared for 7 seconds each in a PowerPoint presentation.

#### 4.1.2 Participants

There were 27 participants in Experiment 2, of which 23 acted as Guessers (11 were male, average age was 21). There were 28 participants in Experiment 3, of which 24 acted as Guessers (12 were male, average age was 21), and 30 participants in Experiment 4, of which 26 acted as Guessers (9 were male, average age was 20). Participants were undergraduate students in diverse disciplines.<sup>3</sup>

#### 4.1.3 Procedure

There were two conditions: in condition 1 Guessers could ask three questions; and in condition 2 only one. Guessers were assigned randomly to two orders of conditions. In each condition they addressed a different Informant and had one guessing attempt. Guessers were

<sup>3</sup>Law, political science, economics, humanities, biology, engineering, computer science, journalism, management, and several related disciplines.

Table 2: Content structure of guessers' single questions: proportion of total questions (%).

Category and codes	Experiment 1 (music)	Experiment 2 (chocolate)	Experiment 3 (pleasant images)	Experiment 4 (aversive images)
<b>“Time profile” persp.</b>	<b>39</b>	<b>13</b>	<b>30</b>	<b>15</b>
– Ratings stats	22	9	17	15
– Real-time ratings	17	4	0	0
– Duration	0	0	13	0
<b>Behavioral implications</b>	<b>39</b>	<b>52</b>	<b>29</b>	<b>20</b>
– Future use	33	35	8	8
– Instrumentality	0	9	21	4
– Approach/Avoidance	0	4	0	4
– WTP	6	4	17	4
<b>Holistic A/J</b>	<b>12</b>	<b>34</b>	<b>34</b>	<b>58</b>
– Attributes	0	0	17	23
– Experience category	6	17	4	12
– Category liking	6	17	0	0
– Emotions	0	0	13	23
<b>Personality</b>	<b>6</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>Decision rule</b>	<b>0</b>	<b>0</b>	<b>4</b>	<b>0</b>
<i>Non-classifiable</i>	6	0	4	8
Total questions	N=18	N=23	N=24	N=26

rewarded for accurate guesses in one condition chosen at random (an error of 7 points was allowed). If the guess had been made based on three questions, Guessers were paid 10 euros. If it had been made based on one question they were paid 15 euros. The show-up fee was 12 euros for Informants and 5 for Guessers.<sup>4</sup>

There were two new items added to questionnaire A: questions about the trend and the variance of real-time ratings. Questionnaire B contained 9 items representative of question categories observed in Experiment 1 (see Table 4).

Every experiment was conducted in two 50-minute sessions. Prior to the beginning of each session, 2 participants had to prepare for the role of Informants. One experienced the long version of the experience, and the other the short version. Both Informants had rated experiences lived and were ready to reply to the questions of Guessers by the time the session began. Experiments were run in spacious classrooms with Informants seated in the back rows and at a distance from Guessers. Performance feedback was given after Guessers completed both conditions and answered questionnaires A and B.

<sup>4</sup>A new recruitment system established higher minimum hourly pay rates.

## 4.2 Results and Discussion

Tables 1-2 report the structure of questions in Experiments 2-4.<sup>5</sup> It is consistent with previous findings.<sup>6</sup>

The analysis of how participants combined frames of analysis revealed that most of them combined at least two (65% of all participants in Experiment 2 (chocolate tasting), 58% in Experiment 3 (positive images) and 81% in Experiment 4 (aversive images)). Most frequent types of combinations in Experiment 2 were equivalent to those observed in Experiment 1. Most frequent types of combinations in Experiments 3-4 involved holistic A/J and behavioral implications. In Experiment 4 two other important combinations involved the “time profile” perspective and holistic A/J, as well as the “time profile” perspective and personality. The most frequent single frame of analysis involved holistic A/J.

In Experiment 2, there were 26 successful guesses in 48 attempts in two conditions. If participants had asked for real-time ratings only, and averaged them, or computed the average of maximum and final ratings, their

<sup>5</sup>Classification scheme developed in Experiment 1 was used to categorize the questions of Guessers. The inter-coder agreement was 94% in Experiment 2, 85% in Experiment 3, and 86% in Experiment 4.

<sup>6</sup>A proper statistical qualification of the differences would require a greater sample, perhaps, a different method, and lies outside the scope of the present article.



Table 3: Questionnaire A. Proportion of guessers choosing an item (%).

Questionnaire items*	Experiment 1 (music)	Experiment 2 (chocolate)	Experiment 3 (pleasant images)	Experiment 4 (aversive images)
Average rating	67	65	63	65
Maximum rating	56	35	42	15
Minimum rating	56	9	21	15
Modal rating	39	52	63	73
Sum of ratings	17	26	42	38
Trend	na	39	17	31
Variance	na	35	25	31
Middle rating	11	4	0	4
End rating	11	13	4	8
Duration	11	17	42	23
Location of maximum	11	4	0	15
First rating	6	4	0	0
Total participants	N=18	N=23	N=24	N=26

\* Questionnaire items were questions formulated so as to avoid the use of statistical terms. For example, the question about the modal rating read: “What was the most frequent rating you used to rate these performances/pieces/images?”, the question about the trend: “Was the experience increasingly pleasant or increasingly unpleasant?”, the question about the maximum rating: “What was the rating of your favorite performance/piece/image?”, and so on.

Preferences in favor of items chosen by at least 35% of participants are significant at 10% and higher levels of statistical significance.

performance would not have improved (24, or 12 guesses would have been made). In Experiments 3-4, instead of 15 successful guesses in 48 attempts and 13 in 52, participants would have attained higher success rates of 37 in 48 and 42 in 52 by averaging across real-time ratings, or, been correct half of the time if believing in the “Peak-End rule”.

Choices of items in questionnaire A resembled results obtained in Experiment 1. Guessers believed that average and modal real-time ratings were the most informative pieces of information about the experience’s overall evaluation. Choices of items in Questionnaire B showed that the importance of average real-time ratings withstood the comparison to other kinds of questions. In Experiments 2-3, Guessers thought that the willingness-to-pay for the experience and the knowledge of experience category is comparable in importance to the knowledge of average real-time rating.

## 5 General Discussion

Overall evaluations of experiences have been studied extensively by economists, psychologists, and philosophers. Opposing the “time profile” perspective on experiences, the latter have argued for the role of “reconstruc-

tion” in overall evaluations (e.g., Alexandrova, 2005). By means of a novel guessing task with Active Information Search, I identified the considerations that decision-makers themselves relate to overall evaluations. These may play an important role in the process of “reconstruction”, as well as in interpersonal communication on the subject.

Researchers may be interested to learn whether certain features of overall evaluations are intuited by people. For example, most theories and empirical findings suggest that the duration of experiences is not an important determinant of overall ratings. Importantly, lay theorists manifested similar beliefs by paying little attention to duration in the search of information about overall evaluations. However, while researchers distinguish between overall ratings and willingness-to-pay judgments (Ariely & Loewenstein, 2000), some lay theorists confounded the two.

Frames of analysis employed by lay theorists have been shown to parallel frames of analysis in academic theorizing on subjective satisfaction judgments used within separate research traditions. However, multiple frames were evoked simultaneously in the minds of lay theorists with respect to very simple experiences. This suggests the need to explore potential interactions. Moreover, different people have used different frames. Thus,

Table 4: Questionnaire B. Proportion of Guessers Choosing an Item (%).

Questionnaire items	Experiment 2 (chocolate)	Experiment 3 (pleasant images)	Experiment 4 (aversive images)
– What was the average rating you had given to these pieces/images?	65	67	77
– How much would you pay to repeat this experience?	52	25	19
– What was the rating of your favorite piece/image?	48	54	69
– How many times per month would you like to repeat this experience?	48	25	8
– What other experience would you describe as similar?	43	46	77
– How would you rate your life in general?	13	38	19
– Would this experience be useful to entertain friends?	13	17	15
– How many pieces/images did you sample/view?	9	33	15
– How would you rate the sweetness/color scheme of the samples/images?	4	8	12
Total participants	N=23	N=24	N=26

Preferences in favor of items chosen by at least 43% of participants are significant at 10% and higher levels of statistical significance.

future research can be aimed at exploring features of the communication context that allow people to coordinate on a given frame and a meaning for overall evaluations of experiences.

Methodologically, this work contributes to the study of human judgment by demonstrating what lay intuitions can add to laboratory findings and the assumptions of a research tradition. The guessing task can be used for the study of lay intuitions in a number of settings, from the forecasting of preferences for specific objects to the predictions of actions in situations of strategic interaction. Importantly, Active Information Search can then be easily made incentive-compatible, and allow the manipulation of stakes involved. The separation between the target of the prediction and the Guesser provides additionally the possibility of exploring self-other differences in evaluation criteria, beliefs and the framing of many decision situations. Even when participants are not able to articulate their intuitions perfectly, the researcher is able to document general frames of analysis employed and, therefore, the fundamentals likely to be used in further articulating a lay theory.

## References

- Alexandrova, A. (2005). Subjective well-being and Kahneman's "objective happiness", *Journal of Happiness Studies*, 6, 301–324.
- Ariely, D. & Carmon, Z. (2003). The sum reflects only some of its parts: A critical overview of research about summary assessment of experiences. In G. Loewenstein, D. Read, and R. Baumeister (Eds.), *Time and Decision: Economic and Psychological Perspectives on Inter-temporal Choice*, pp. 323–350. New York: Russel Sage Foundation.
- Ariely, D., & Loewenstein, G. (2000). The importance of duration in ratings of, and choices between, sequences of outcomes. *Journal of Experimental Psychology: General*, 129, 508–523.
- Ariely, D. & Zauberman, G. (2000). On the making of an experience: the effects of breaking and combining experiences on their overall evaluation. *Journal of Behavioral Decision Making*, 13, 219–232.

- Brendl, M. & Higgins, T. (1995). Principles of judging valence: what makes events positive or negative? In M. P. Zanna (Ed.) *Advances in experimental social psychology*, vol. 28, pp. 95–160. New York: Academic Press.
- Hsee, C., & Abelson, R. (1991). The velocity relation: satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology*, 60, 341–347.
- Huber, O., Wider, R. & Huber, O. W. (1997). Active information search and complete information presentation in naturalistic risky decision tasks. *Acta Psychologica*, 95, 15–29.
- Kahneman, D. (2000). Evaluation by moments: past and future. In D. Kahneman, and A. Tversky (Eds.), *Choices, Values and Frames*, pp. 693–708. New York: Cambridge University Press.
- Kahneman, D., Wakker, P. & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112, 375–405.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6. University of Florida, Gainesville, FL.
- Langer, T., Sarin, R. & Weber, M. (2005). The retrospective evaluation of payment sequences: duration neglect and peak-and-end effects. *Journal of Economic Behavior and Organization*, 58, 157–175.
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97, 666–684.
- Oishi, S. & Sullivan, H. (2006). The predictive value of daily vs. retrospective well-being judgments in relationship stability. *Journal of Experimental Social Psychology*, 42, 460–470.
- Read, D., Loewenstein, G. & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, 19, 171–197.
- Robinson, M., & Clore, G. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128, 934–960.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110, 403–421.
- Updegraff, J. & Gable, S. & Taylor, S. (2004). What makes experiences satisfying? The interaction of approach-avoidance motivations and emotions in well-being. *Journal of Personality and Social Psychology*, 86, 496–504.
- Wirtz, D. & Kruger, J., Napa Scollon, C. & Diener, E. (2003). What to do on spring break? The role of predicted, on-line and remembered experience in future choice. *Psychological Science*, 14, 520–524.