SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,
Entscheidungsverhalten und
ökonomische Modellierung

No. 00-45

The Retrospective Evaluation of Payment Sequences: Duration Neglect and Peak-and-End-Effects

Langer, Thomas*
and Sarin, Rakesh**
and Weber, Martin***

October 2000

Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim, is gratefully acknowledged.

*Lehrstuhl für ABWL, Finanzwirtschaft, insb. Bankbetriebslehre, email: langer@bank.bwl.uni-mannheim.de

**UCLA, Anderson School of Management, email: rsarin@anderson.ucla.edu

***Lehrstuhl für ABWL, Finanzwirtschaft, insb. Bankbetriebslehre, email: weber@bank.bwl.uni-mannheim.de

Universität Mannheim
L 13,15
68131 Mannheim

Abstract:

In this paper we present experimental research examining the ability of individuals to make good retrospective evaluations of payment sequences. Inspired by the evidence on systematic biases in the retrospective evaluation of affective episodes involving pain and pleasure we designed choice scenarios for payment sequences, in which the existence of peak and end effects as well as duration neglect could be examined. There are two main results: We do not observe a systematic impact of payment sequence features (other than its sum) on the choices if subjects merely get delivered the payments without any affection or effort. Subjects, by and large, choose the sequence with the highest total payment. In a second scenario, in which payments were linked to the subjects' effort and performance in strenuous tasks, we observe a strong effect of duration neglect and a weaker, but still significant end effect. We further find that the mere number of peak losses in a sequence strongly influences its attractiveness. In this scenario subjects do not often choose the sequence with the highest total payment.

Key words: retrospective evaluation, duration neglect,
temporal preferences, evaluation bias

I. Introduction

There are many real world situations in which individuals need to make retrospective evaluations of payment sequences. When choosing the brand for a new car, it could be helpful to recall the repair costs of formerly driven cars. The desision for this year´s vacation type and destination might require a retrospective comparison of the costs involved in different types of vacations in the past. Examples are not restricted to cost situations, but can also involve profits (dividend payment sequences for different stocks, daily tips for a waiter having worked in different places, monetary gains in different types of gambling, etc.).

If the complete sequence of payments is objectively available ex post, the task of retrospective evaluation is mostly trivial. If we restrict our attention to sequences within a rather short time interval and thus neglect issues of discounting, the normatively appropriate aggregation rule „adding up all payments“ is obvious and easy to apply. In this paper, we address the question of how well individuals make retrospective evaluations if they cannot access the objective data ex post, but have to rely on their memory of the payments. In two experimental studies, we analyze which features of a payment sequence (other than its sum) influence its retrospective evaluation.

Our research is closely related to the experimental work on the retrospective evaluation of temporally extended streams of hedonic experiences like pain or pleasure. In these studies, a systematic bias was observed, which was dubbed „peak-and-end rule“ by Fredrickson and Kahneman (1993). Regarding to this rule individuals base their global evaluation mainly on the final and the most intense utility levels in the episode. Consequences of this aggregation procedure, the neglect of episode duration and possible violations of temporal monotonicity, were confirmed in many further studies.

However, the transferibility of these insights to the payment context is questionable as payment sequences differ in many respects from pain and pleasure streams. First, the affective pain/pleasure episodes are *continuous* streams of instant utilities while payment sequences provide *discrete* stimuli. Second, payments are only indirectly related to consumption and hedonic experiences. Third, there exists an obvious and easy to apply normative aggregation rule for payments. The latter fact simplifies our experimental work as it eliminates the problem of defining what should be considered an evaluation or choice bias.

We conducted two experiments to test whether peak and end effects and duration neglect translate to the retrospective evaluation of payment sequences. In the first experiment, subjects were shown two sequences of payments and were asked to choose between the two. In the second experiment, each payment in the sequence was contingent upon the length of time it took to

complete a task. Thus, the task produced a distraction so that the subjects had to rely on their memory in their retrospective choice between the two sequences. For example, a subject could remember a running total of a sequence of payments 5, 1, 8, ..., but may find it difficult to do so if each payment is the result of a task that requires some effort.

The remainder of the paper is organized in the following way. In section II, we summarize some literature on retrospective evaluation of affective episodes and highlight and discuss the aspects that are most relevant for our research project. In section III, we present the design and the results of the first experiment, in which we examined the retrospective evaluation of payment sequences. In section IV, the design and the results of the second experiment are presented. In section V, we summarize the results and discuss possible extensions.

II. Literature and General Discussion

The question, which features of temporally extended experiences determine preferences, is object of an extensive literature. In empirical research, it was demonstrated that preferences are strongly influenced by the ordering of the utility levels within the profiles. It was shown, for example, that increasing sequences are considered more attractive than decreasing ones, implying negative time preference (Loewenstein and Prelec, 1991, 1993). The relevance of trend for a sequence‘s evaluation was confirmed for many different stimuli such as pain (Ariely, 1998; Ariely and Carmon, 2000; Ariely and Zauberman, 2000; Kahneman et al., 1993; Redelmeier and Kahneman, 1996), affective episodes (Fredrickson and Kahneman, 1993; Varey and Kahneman, 1992), TV advertisements (Baumgartnern et al., 1997) and wages (Loewenstein and Sicherman, 1991). It is another robust finding that the duration of an experience is often strongly underweighted in the overall evaluation (Ariely, 1998; Ariely and Loewenstein, forthcoming; Fredrickson and Kahneman, 1993; Redelmeier and Kahneman, 1996).

Our research is inspired by some empirical work on the retrospective evaluation of pain and pleasure streams, which we describe in more detail. Fredrickson and Kahneman (1993) examined the impact of the duration of an episode on individuals retrospective evaluation. As the stimulus, they used pleasant and aversive film clips of different lengths (like flying over African landscape or a medical film of amputation). They obtained moment-by-moment reports of instant utility from their subjects. Immediately after the end of each film subjects had to report, how much (dis)pleasure they experienced overall. Assuming the remembered utility to be constructed as a „weighted average of instant utilities“ Fredrickson and Kahneman concluded that „... *most moments of an episode are assigned zero weight in the evaluation and a few select snapshots receive larger weights* “. They identified the peak and end values of instant utility to receive

excessive weight and showed that an unweighted average of these two instant utility reports can well explain the data. A tendency to neglect the duration of an episode is a direct implication of this peak-and-end evaluation. An even more salient and easily testable consequence is the potential violation of temporal monotonicity. Adding some moderate pleasure to an episode with a high final pleasure value could result in a lower retrospective evaluation. By the same argument, extending a stream of displeasure by some moderate discomfort could make the experience look less aversive in retrospective evaluation. This prediction was tested by Kahneman et al. (1993) in a simple choice task. In their experiment, subjects had to place a hand in painfully cold water. There were two treatments: the one lasted 60 seconds with a water temperature of 14°C, the other treatment started with the same 60 second experience, but was extended by an extra 30 seconds with the water temperature gradually increasing to a still painful 15°C. The order of the two treatments was randomly chosen and different hands were used for the two experiences. Subjects were informed that they had to undergo a third treatment and were asked to choose one of the two previously experienced unknown temperature profiles. A significantly higher number of subjects (22 of 32) chose the dominated treatment.

The practical relevance of the violation of temporal monotonicity was demonstrated by Kahneman and Redelmeier (1996) in a clinical setting. Six hundred eigthy-two patients undergoing a colonoscopy had to report their level of pain every minute. After the treatment the participants answered different questions about their experience, in particular they were asked to evaluate the total amount of pain. To test the violation of temporal monotonictiy the treatment was extended for half of the subjects by one minute. This extra time was used to leave the colonoscope in place without performing any clinical examination. It could be inferred from the momentary pain reports that the extra minute was less painful than the main colonoscopy, but still unpleasant. Nevertheless, the retrospective overall pain evaluation was significantly lower in the group undergoing the extended treatment.

In this paper, we examine whether these biases translate to the retrospective evaluation of monetary payment sequences. This type of stimulus differs in some important ways from the pain and pleasure settings mentioned above. A central point is the existence of an obvious and easy to apply normative aggregation rule for monetary payments. As we can ignore issues of discounting (given the short time horizon of our experiments) the overall evaluation of a monetary sequence should be solely based on the sum of the payments. This normative rule is not only obvious, individuals are also familiar with this procedure as they are used to apply it in many real world situations. In contrast, it is much less obvious in the pain/pleasure setting, how an overall evaluation of a profile of instant utilities should normatively be obtained. From a set of axioms Kahneman, Wakker and Sarin (1997) derive some integration rule to be the normative aggregation procedure. Nevertheless it remains rather ambiguous, whether giving excessive

weight to the most intense moment of a pain profile violates basic principles of normative evaluation.[1]

As a second important difference, payments are not affective per se. They are only indirectly related to consumption and hedonic experiences. However, in many real world situations, payments are linked to effort and performance and are affective through the sensation of achievement. In one of our experiments, we use this linkage and make the payments more affective experiences.

Finally, payment sequences present discrete experiences while all the affective episodes described above were continuous streams. This could be particularly relevant for the occurrence of a duration neglect effect for payment sequences. It might well be that individuals have problems to judge the length of a continuous stream, but nevertheless are able to correctly process the number of payments in a sequence. Ariely and Zauberman (1998) report experimental results that are in line with this idea. They addressed the impact of a temporal spacing of the overall experience and found that breaking an episode into smaller pieces by the inclusion of short pauses moderated the observed bias. They hypothesized that the breaks cause the individuals to perform interim evaluations of the episode pieces, such that in retrospective consideration these evaluations rather than the instant utility levels are aggregated. They further argue that individuals are less biased when aggregating a number of discrete evaluations than when evaluating a profile of instant utilities.

Despite these insights and the other above-mentioned differences between payment sequences and pain/pleasure episodes we expected to find peak and end effects and duration neglect in the retrospective evaluation of monetary sequences. To be able to refer to the rules in the following exposition, we will conclude this section with a formal definition of the general effects, which we hypothesize to observe in the payment scenarios. They are all straightforward modifications from the pain and pleasure scenario. For example the term duration neglect is identified with a neglect of sequence length.

Rule π (peak effect): In retrospective evaluation of payment sequences subjects give excessive weight to the most extreme payments in the sequence.

Rule ε (end effect): In retrospective evaluation of payment sequences subjects give excessive weight to the last payments in a sequence.

1 Cf. also the discussion of Ariely and Loewenstein (forthcoming) on the question whether duration neglect should be considered an error.

Rule δ (duration neglect): In retrospective evaluation of payment sequences subjects neglect the overall length of the sequences.

Now we can turn to the description of our experimental designs, which test these predictions.

III. Design and Results of Experiment 1

The first experiment consisted of two parts (1a and 1b). In experiment 1a, we focused exclusively on peak and end effects. Thus all sequences that had to be compared were of the same length. Study 1a was conducted in class without monetary incentives. Overall, 179 advanced business students of the University of Mannheim took part in the experiment.

In the first session, we presented 6 pairs of number sequences to 145 students. The numbers were displayed on an overhead projector and in addition read aloud by the experimentator. These sequences (pairs A to F) are listed in exhibit 1. In the first three pairs (A, B, C), each number was displayed only while it was read, i.e. subjects saw just one number at each point of time. In the next three pairs (D, E, F), subjects saw all numbers in a pair at the same time, though they were read one after another. In both cases, a new number was read every two seconds. The end of the first sequence in each pair was particularly highlighted, followed by a short pause and the start of the second sequence. After each pair the subjects were asked to mark on a piece of paper, which of the two sequences in the pair they preferred. We did not give any interpretation of the meaning of „preference", in particular we did not instruct subjects to consider the numbers to be monetary payments.

Some of the subjects in this first session stated confusion with the meaning of „preference". In a second session with 34 different subjects, we used the same design as for the pairs A through C, but asked subjects to choose the „more attractive" sequence. The four additional pairs (G through J) used in this session are also presented in exhibit 1. Again, we did not make any further comments on the meaning of the numbers. Nevertheless, we hypothesized subjects' preferences to reflect the peak and end effects stated in the rules π and ε in section II. All sequence pairs were designed in a way such that explicit preference hypotheses could be derived from these rules. The indices of the sequence names reflect these hypotheses. In all cases the sequence with the index 2 (A_2, B_2,...) is hypothesized to be more preferred (or more attractive). The notation $*_2$ is used to generally refer to these sequences. The order of sequence presentation can be inferred from exhibit 1 (left sequence first).

	Pair A		Pair B		Pair C		Pair D		Pair E		Pair F		Pair G		Pair H		Pair I		Pair J	
7 Order of presentation	A_1	A_2	B_1	B_2	C_2	C_1	D_2	D_1	E_2	E_1	F_2	F_1	G_1	G_2	H_2	H_1	I_1	I_2	J_2	J_1
	5	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**3**	**4**	**4**	**-6**	**23**	**56**	**4**	**-6**	**14**	**13**	**2**	**-5**
	3	**2**	**3**	**2**	**3**	**3**	**-2**	**-2**	**-2**	**-2**	**6**	**-1**	**14**	**23**	**6**	**-1**	**-21**	**-21**	**4**	**-1**
	8	**3**	**8**	**3**	**2**	**8**	**4**	**8**	**8**	**8**	**-2**	**5**	**76**	**12**	**-2**	**5**	**18**	**18**	**-7**	**6**
	6	**2**	**6**	**7**	**6**	**6**	**8**	**7**	**7**	**7**	**3**	**3**	**54**	**34**	**-7**	**7**	**7**	**7**	**-6**	**7**
	4	**1**	**4**	**9**	**4**	**4**	**-4**	**-9**	**-9**	**-1**	**-7**	**7**	**68**	**32**	**2**	**-4**	**-11**	**-19**	**1**	**-2**
	7	**6**	**7**	**6**	**1**	**7**	**2**	**6**	**6**	**6**	**2**	**-4**	**85**	**53**	**-6**	**6**	**16**	**16**	**-5**	**7**
	4	**3**	**4**	**3**	**4**	**4**	**1**	**3**	**5**	**5**	**-6**	**6**	**67**	**61**	**-1**	**2**	**15**	**15**	**4**	**-4**
	5	**4**	**5**	**4**	**5**	**5**	**3**	**4**	**4**	**6**	**-1**	**2**	**75**	**45**	**-4**	**5**	**6**	**-11**	**-7**	**5**
	2	**7**	**2**	**2**	**7**	**2**	**-3**	**-4**	**-4**	**3**	**-4**	**5**	**42**	**97**	**5**	**-2**	**13**	**-14**	**2**	**-1**
	1	**9**	**1**	**1**	**8**	**1**	**-1**	**-1**	**-1**	**-4**	**5**	**-2**	**31**	**89**	**7**	**-7**	**-14**	**14**	**3**	**-3**
							6	**6**	**6**	**-9**	**7**	**-7**					**-19**	**6**	**9**	**-9**
sum	**45**	**42**	**45**	**42**	**45**	**45**	**19**	**23**	**23**	**23**	**7**	**8**	**535**	**502**	**4**	**5**	**24**	**24**	**0**	**0**
results		*hyp.*		*hyp.*	*hyp.*		*hyp.*		*hyp.*		*hyp.*			*hyp.*	*hyp.*			*hyp.*	*hyp.*	
no. of subj.	***106***	***39***	***104***	***41***	***60***	***85***	***62***	***83***	***56***	***89***	***70***	***75***	***29***	***5***	***10***	***24***	***17***	***17***	***26***	***8***
perc.	*73 %*	*27 %*	*72 %*	*28 %*	*41 %*	*59 %*	*45 %*	*55 %*	*39 %*	*61 %*	*48 %*	*52 %*	*85 %*	*15 %*	*29 %*	*71 %*	*50 %*	*50 %*	*76 %*	*24 %*

exhibit. 1: Sequences used in experiment 1a and preferences (attractiveness judgments) of subjects.

All sequences in study 1a were constructed such that the sequence $*_2$ was predicted to be preferred, though it had the lower sum at the same time. This fact allows discrimination between a peak and end driven preference and a decision, which is based on the sum of the numbers in the sequence.

- In the pair A the sequence A_2 has the higher end as well as the higher peak. The total of A_2 is lower, however.
- The pair B is just a rearrangement of the numbers in pair A. Thereby the end effect is eliminated while the peak effect remains unchanged.
- Pair C provides identical numbers in both sequences and thus identical totals. By the ordering, however, we have a higher end in sequence C_2.

In the next sequences, positive and negative numbers are combined, such that there could be two directions for „most extreme numbers“ and thus peak effects. We chose the numbers in a way, however, that rule π makes unambiguous predictions.

- In the pair D, we have no end effect, but a much stronger negative peak in sequence D_1.
- In pair E, the two sequences are identical except for order. The negative peak at the end of sequence E_1 should make subjects prefer E_2 according to rule ε.
- The sequence F_2 was predicted to be preferred by most subjects despite its lower total, since the negative start and end of sequence F_1 are especially salient.
- Pair G was constructed to additionally examine the effect of the number size (possibly influencing the willingness and/or ability to add up the numbers). G_2 has the higher peak and end, but the lower total.
- The pair H is basically a replication of pair F (except a missing 3 in both sequences).
- By a rearrangement of the numbers, I_2 has the higher end in pair I.
- In pair J, we particularly highlighted the end difference by placing the positive and negative peaks at the end of the sequences.

Results of study 1a:

The results of study 1a are presented at the bottom of exhibit 1. In all pairs with different sequence totals, the majority of subjects chose the sequence with the higher total. This result is contrary to our hypotheses and is significant for the pairs A, B, D, G and H. Only in pair F, in which the sequence totals are almost the same, the choices do not significantly deviate from randomness. Our explanation of these results is that in this simple setting subjects try to choose a sequence by the higher total. Mental arithmetic problems can account for the variability in the choice data. A specific impact of the peak and the end of the sequences is not observed.

The choices in the pairs with identical sequence totals (C, E, I, J) are harder to interpret. In fact, we do not have an explanation for the choice pattern in the pairs C and E. They are opposite to our hypotheses and significantly deviate from randomness, which would be in line with a „choice by higher total“ procedure. As an alternative explanation for the observed choice pattern, it could be argued that some subjects considered the implemented peak and end effects to be too obvious (this explanation would also work for the pairs A, B, D, F, G and H). Suspiciousness could have made them avoid the sequence, which was obviously designed to look more attractive. This explanation is not supported, however, by a within subject analysis. It turns out that there are surprisingly few choices in the way that subjects state preferences consistently in line with our hypotheses or against our hypotheses in different pairs. For the pairs A and C for example, which show a strong similarity in the number profiles, a negative correlation is observed. Less than 50% of the subjects stated preferences either in line or against our hypotheses in both pairs. After the second session of the experiment, we asked several students

about their suspiciouness in the experiment. We got the unanimous answer that subjects did not consider any sequences to be suspicious and thus did not make intentional choices in the opposite direction.

Finally, the result in the pair J should be highlighted. The choice pattern in this pair is surprisingly distinct from all other pairs and in particular it is in line with our hypotheses. This result is especially puzzling as the structure of pair J is very similar to the structure of pair H. The main difference between these pairs consists of different proportions of negative to positive numbers in the sequences. While in pair H there are five negative numbers in both sequences, sequence J_2 includes less negative numbers (four) than sequence J_1 (seven). One could conclude that the number of negative numbers has a stronger negative impact on the attractiveness of a sequence than the size of the negative numbers (the negative numbers add up to the same total in H as well as in J).

Design of study 1b:

Study 1b differed from study 1a in three respects. First, the experiment was computerized and number sequences were individually assigned to each subject. In contrast to study 1a, this allowed us to control for order effects (i.e., the order of the sequences within a pair as well as the order of the pairs within the experiment). Second, the stimuli as well as the incentive structure were modified. Subjects were told that every displayed number meant a direct payment to them (which it really was). After watching and receiving both payment sequences of a pair, the subjects were asked to choose which of the sequences they preferred to be displayed and thus to be paid to them again. This is an important difference to the vague preference and attractiveness statements in study 1a. Assuming that subjects were trying to maximize their profits, the choice of a sequence with a lower total can indubitably be considered to be a bias. Third, within a pair the sequences could differ in length. Hence not only the peak and end rule, but also the general effect of duration neglect came into play.

Thirty-six business students from the University of Mannheim took part in experiment 1b. After individually entering the lab, they read the instructions on a computer screen. They were informed that in the course of the experiment they would see sequences of numbers with each number displayed on the screen constituting an actual Pfennig payment to them.[2] They were further instructed about their task to choose one of the sequences in each pair to be displayed (and paid to them) again. Then they saw the eight prespecified sequence pairs in random order

2 At the time of the experiment, 170 Pfennig were about 1 US$.

(random within the pair and between pairs, but not within the sequence).[3] The first sequence was presented in the left half of the screen on a red background with the numbers changing every two seconds. After the last number of the first sequence subjects were asked to rate the attractiveness of the sequence on a 0-100 scale using a sliding bar on the screen. The second sequence in the pair was then displayed on a blue background in the right half of the screen with the same frequency. Subjects were asked again to provide an attractiveness rating. After viewing both sequences subjects were asked to choose one of the two sequences for payment.

There were two sequence pairs with rather huge differences in total payment in the experiment. They were included as we did not want to provide the impression of only close choices (not worth investing any mental effort for payment reasons). We will not present sequence details and results for these irrelevant pairs. The six remaining sequence pairs used in experiment 1b and their totals are listed in exhibit 2. They were all constructed in a way such that the sequence with the lower total was predicted to be chosen by most subjects.

	Pair K		Pair L		Pair M		Pair N		Pair O		Pair P	
7 Order of presentation	K_1	K_2	L_1	L_2	M_1	M_2	N_1	N_2	O_1	O_2	P_1	P_2
	9	**9**	**19**	**16**	**4**	**11**	**6**	**12**	**11**	**5**	**12**	**11**
	21	**21**	**5**	**11**	**2**	**4**	**13**	**6**	**14**	**20**	**14**	**6**
	25	**25**	**15**	**4**	**15**	**2**	**4**	**5**	**14**	**23**	**16**	**8**
	11	**11**	**16**	**15**	**10**	**6**	**23**	**8**	**12**	**6**	**9**	**12**
	15	**15**	**11**	**27**	**13**	**6**	**12**	**13**	**9**	**9**	**23**	**15**
	12	**12**	**4**	**19**	**17**	**10**	**9**	**4**	**9**	**12**	**20**	**22**
	20	**20**	**11**	**10**	**13**	**12**	**8**	**23**	**15**	**9**	**8**	**6**
	27	**27**	**16**	**13**	**15**	**9**	**5**		**13**	**11**	**12**	**3**
	12		**13**		**8**	**19**						**7**
	6		**10**		**6**	**17**						**14**
sum	**158**	**140**	**120**	**115**	**103**	**96**	**80**	**71**	**97**	**95**	**114**	**110**
results		*hyp.*		*hyp.*		*hyp.*		*hyp.*		*hyp.*		hyp.
chosen by .. subj.	***28***	***8***	***18***	***18***	***28***	***8***	***31***	***5***	***24***	***12***	***30***	***6***
percent.	*78%*	*22%*	*50%*	*50%*	*78%*	*22%*	*86%*	*14%*	*67%*	*33%*	*83%*	*17%*

exhibit 2: Sequences used in experiment 1b and choices.

3 Since there was no systematic order effect found in the data of study 1b, we do not pay attention to the order in the presentation of the results.

- In pair K, two extra numbers are added to the end of sequence K_2, thereby increasing the total, but providing a lower end at the same time.
- In pair L, both sequences consist of the same payments with two exceptions. The peak payment 27 in L_2 is split into two payments 16 and 11 in L_1, causing L_1 to have a distinctly lower peak. This should make L_2 look more attractive by rule π, even though a further number 5 is added to L_1.
- In pair M, a pure end effect is examined with no duration difference and no distinct peak effect. M_2 has the higher end payments, though M_1 has the higher total.
- In the pair N, both sequences consist basically of the same payments (in different order) with the only exception that an extra payment 9 is added to N_1. Concerning to rule ε subjects should nevertheless prefer N_2 since the high end payment is very salient.
- In the pair O, the sequence O_2 has the higher peak, but the lower sum. The identical length of both sequences excludes any duration effects.
- Finally, pair P was designed to test the alternative hypothesis that the pure length of a sequence makes it more attractive. P_2 has the lower total, but consists of more payments. Thus it should be preferred according to this hypothesis.

Results of study 1b:

The results of study 1b are presented at the bottom of exhibit 2. With the exception of pair L, where choices were equally split, we found a strong majority of subjects choosing opposite to our hypotheses in all other pairs. Obviously, the effect of higher total dominated any intended effect of peak, end or duration. It should be noted that the results for pair P are in line with the duration neglect rule δ. However, a duration neglect effect cannot be discriminated from the unbiased choice of the higher total, as the pair was constructed to test a different hypothesis.

It seems that the change in the payment scheme from study 1a to study 1b has even strengthened subjects‘ willingness to base their decision on the totals of the sequences. From post experimental discussions, we inferred that most subjects tried to perform a mental on-line aggregation of the payments. We conclude that for this type of simple, well specified and ex ante known choice task structural features of the payment sequences do not influence subjects choices. Instead they:

- are aware of the normative status of an aggregation procedure,
- try to avoid the need for a (possibly biased) retrospective sequence evaluation
- and thus try to perform a mental online aggregation.

IV. Design and Results of Experiment 2

From study 1, we learned that individuals are willing and able to apply the normatively appropriate aggregation problem if the experimental setting allows one to perform an online aggregation. However, there are economic situations in which individuals receive the payments for their effort and performance. We examined this kind of situation in study 2. There are two important consequences of this scenario change. First, the linkage of payments to effort and performance makes the monetary sequences more affective episodes (without changing the obvious normative appropriateness of a simple payment aggregation). Second, as the strenuous tasks that determine payments require mental effort, subjects are not (or hardly) able to perform an online aggregation. Thus they have to make indeed retrospective evaluations of the payment sequences.

In the experiment, subjects were confronted with sequences of letter tasks. Each single letter task consisted of a letter between A and Z, an operator from the set {+,-} and a number between 1 and 5. Typical tasks were of the form D+3, M-2 or T+5. Subjects were instructed to interpret each task as a question concerning the ordering of letters within the alphabet. The interpretation of the task D+3 was: „Which letter is the 3rd letter after D in the alphabet?“. Hence, the correct answer to the question D+3 is the letter G. Similarly the question M-2 with the interpretation: „Which letter is the 2nd letter before M in the alphabet?“ has the correct answer K.

The performance in the tasks was measured by the time it took a subject to type in the correct answer on the keyboard. Each tenth of a second of waiting time corresponded to a loss of one Pfennig. The maximum time to think about a question was five seconds. If subjects did not answer within this time limit they lost 50 Pfennig. Wrong answers could not be corrected afterwards. They resulted in an even higher loss of 60 Pfennig. All subjects were informed ex ante about this payment scheme.

Subjects got feedback about their performance in each task. They either received the information „Your answer D was correct - you lost 43 Pfennig“ or „The correct answer was D - you lost 50 Pfennig“ or „Wrong - the correct answer would have been D - you lost 60 Pfennig“ (for details of the information presentation cf. the screenshots in exhibit 3).

On a few introductory screens subjects were informed that they were going to work through sequences of letter tasks, starting each sequence with a budget of 800 Pfennig. They got detailed instructions about the interpretation of the letter problems,[4] the task to perform and the relation

4 Multiple examples were displayed and subjects were asked to approach the experimentator, if they did not understand the principle of the letter tasks. Subjects had no problem to understand the procedure.

between performance and monetary loss. They were not informed ex ante about the fact that sequences were displayed pairwise and that they had to make a payment choice after the completion of each pair. Subjects then started with the first task sequence.

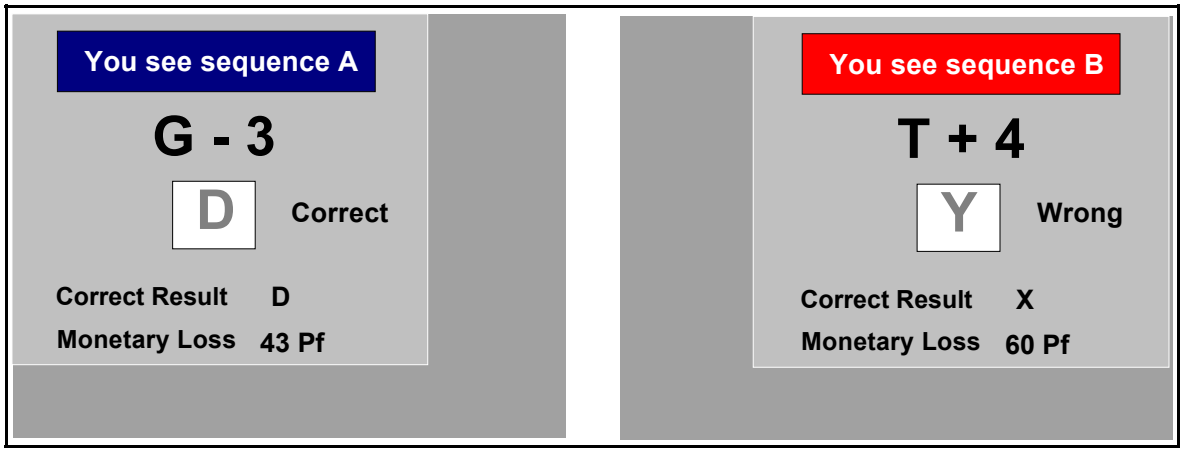

exhibit 3: Two screenshots from study 2 (translated version)

The background color of the computer screen was used to further highlight the connection between thinking time and money subtraction. Whenever money was deduced from the subject's account the background of the computer screen changed to a bright red color. As soon as the correct answer was provided or the time limit was reached the background turned dark again.[5] The tasks and the feedback of the first sequence in the pair appeared on the left side of the screen. The end of the first sequence was explicitly stated and after a short pause the second sequence started on the right side of the screen. After completion of the second sequence subjects were asked: „*Please choose one of the two sequences for your real payment*." We further asked subjects a question concerning the judged difficulty of the two task sequences. „*In your opinion, which sequence will on average result in the higher payment when presented to other subjects?*" Subjects could well give different answers to both questions (e.g., if they were especially

Questions usually referred to the treatment of „out of bounds"-situations like X+4 or A-3 (which we did not use) or the permission to write down the alphabet on a piece of paper (which we did not allow).

5 The linear relation between money loss and time of red background was even maintained in the case of a wrong answer. The immediate feedback „Wrong - the correct letter would have been G" appeared and the screen background remained red colored for a total time of six seconds (according to the size of the loss).

dissatisfied with their performance in an easy task sequence.)[6] At that point in the experiment, subjects did not get any feedback about their overall performance and the quality of their choice (i.e., they were neither informed about the size of their payment nor whether they picked the higher payment). Our impression was that most subjects did not realize that the choice task was the central issue of the experiment. After a short pause, the next sequence pair started. After completion of all pairs, a final screen appeared that listed the payments for all pairs. Subjects were paid by the experimentator according to their performance and choices.

The task sequence pairs were designed to result in payment profiles similar to the ones used in study 1. In exhibit 4 we display an example of a task sequence pair and the resulting losses for a specific subject in study 2. Note, that the sequence Q_2 caused the higher total loss (434 Pfennig compared to 414 Pfennig). The sequence Q_1, however, ends with higher losses. By rule ε we would thus hypothesize that Q_2 is chosen despite its lower total payout.[7]

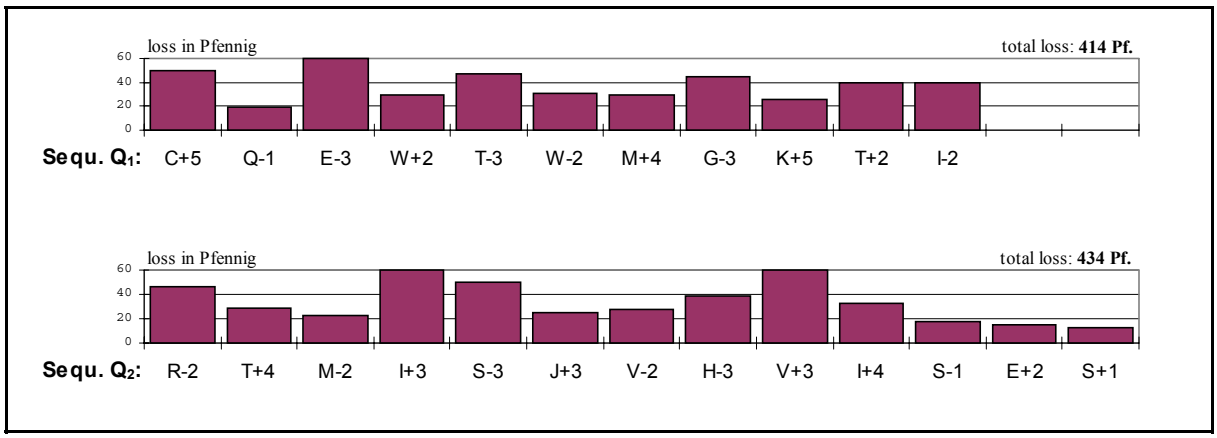

exhibit 4: Example for task sequences in study 2 and resulting losses for a specific individual.

Payment sequences as the one in exhibit 4 are well suited to test the end hypothesis, as we have a combination of a higher total loss, but lower final losses for Q_2. It should be clear, however, that the task sequence design of experiment 2 generally causes a control problem. By the endogenous determination of the payment sequences (dependent on the performance in the tasks) the experimentator no longer has perfect control over the payment profiles as in experiment 1. Even if the task sequences are constructed on the basis of pretest information about average response times, the variability in individual performances makes it difficult to control for payment profiles. The main problem is displayed in exhibit 5. To allow an examination of our hypotheses the difference between the accumulated losses in each sequence has to be in some

6 This second question was included to gain insights about the distinction between retrospective evaluation of performance and retrospective evaluation of task difficulty.

7 And in fact, this subject chose Q_2 for payment.

small range. If the variability in performance causes a much higher total loss in $*_2$ (the sequence, which is designed to look more attractive) as in region I of exhibit 5, we cannot expect payment profile features to outweight this obvious total loss difference. If the difference turns out to have the wrong sign (total loss in $*_2$ lower than in $*_1$ as in region III of exhibit 5), we cannot discriminate between different explanations for $*_2$ -choices. Only if the difference has the right sign and is of moderate size (as in region II of exhibit 5), we can expect to observe the predicted impact of peak, end or duration neglect on the choices.

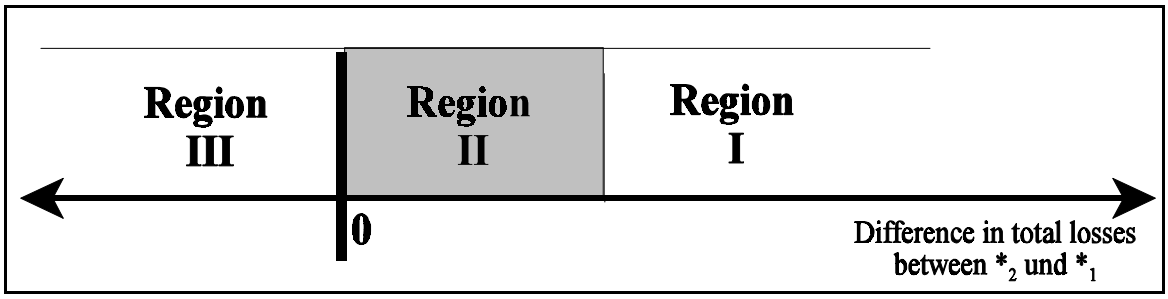

exhibit 5: Total loss differences suited for examination of hypotheses.

To ensure that the total loss differences are in region II, we used an interactive computer program that allowed an adjustment of the sequences to individual performance. While the first sequence in each pair was fixed and independent of performance, the second sequence was shortened or lengthened to guarantee an appropriate difference of total losses.

Four pairs of sequences were used in experiment 2. Two of them were constructed to examine an end-effect (pairs Q and R, referred to as“ end-pairs“ in the following), the other two (pairs S and T, the „duration-pairs“) addressed the phenomenon of duration neglect. The only difference between the pairs Q and R is that Q has a larger number of tasks. Similarly, S has a larger number of tasks than T. All four pairs are displayed in exhibit 6 together with data about the mean loss sizes for the tasks involved.

Notice that in the end-pairs Q and R the degree of difficulty of tasks, except for the end tasks, is similar for both sequences (Q_1 and Q_2, and R_1 and R_2). For the duration-pairs, sequence S_1 has tasks with higher degree of difficulty than S_2 and similarly T_1 has tasks with higher degrees of difficulty than T_2. To understand the underlying procedure consider the duration-pair T. In this pair, the sequence T_2 was intended to be a longer sequence of less difficult tasks with a higher total loss. We hypothesized that despite the higher total loss subjects will prefer sequence T_2, because of a neglect of sequence length (rule δ) in the retrospective evaluation. Both sequences T_1 and T_2 consist of a basic part and an extended part, separated by a dot in exhibit 6. For the first sequence in the pair, only the basic part was presented to subjects (i.e., the sequence ended with the letter task E-2 if T_1 was presented first and ended with A+1 otherwise). The second sequence had a variable length, which was determined by the computer in a way such

that T_2 had a slightly higher total loss. If T_2 was the second sequence, the computer stopped the presentation as soon as the total loss of T_2 exceeded the total loss accumulated in T_1 by more than 30 Pfennig. Thus the total loss difference ($*_2$-$*_1$) was always in the interval [31, 90]. If T_1 was presented as the second sequence the computer stopped the presentation as soon as the total loss of T_1 was less than 60 Pfennig lower than the total loss of T_2. Thus, the difference was always in the interval [0, 59]. We chose these stopping rules (with the different intervals for possible loss differences) as we wanted to have similar mean total loss differences for both orders of presentation to be able to examine unbiased order effects. In fact, the mean loss difference turned out to be almost identical (41) for both orders of presentation in pair T. The overall median was 40, more than 75% of the subjects had a loss difference within the interval [30,60]. For the sequence pair S, which is just a long version of T, the same procedure was used.

For the examination of the end-effect in the pairs Q and R, a more complicated algorithm was required. To understand the mechanism consider the pair R, which was designed such that R_2 ends with two very small losses, but accumulates the higher total loss. Except for the end tasks, the level of difficulty of tasks is about the same in the two sequences.

If R_1 was the first sequence in the pair, the computer stopped at the end of the basic part (with P-3 as the last task). Then the sequence R_2 was presented until the difference in total losses was less than 20 Pfennig. At this point the computer jumped to the end part and presented the two easy tasks B+2 and E+1. This stopping rule does not guarantee the total loss difference to be positive. However, it is hardly possible to cope with the final two tasks in less than 2 seconds and, in fact, we did not get any negative total loss differences in the pairs Q and R. The mean total loss difference for this case was 48. If R_2 was the first sequence in the pair, the computer presented the basic part (up to the task P+4) skipped the extended part and jumped to the end part, presenting B+2 and E+1. Then the sequence R_1 was presented until the total difference was less than 80. Thus all loss differences were in the interval [20, 79]. The mean total loss difference for this procedure was 55.

As in experiment 1, all of our hypotheses can be inferred from the indices of the sequences. In the end-pairs Q and R we hypothesize $*_2$ to be preferred by most subjects as the lower end losses make the sequence look more attractive. In the duration-pairs S and T we hypothesize $*_2$ to be preferred by most subjects as they neglect the overall length of the sequence and choose a sequence that contains somewhat easier tasks, but because of its length results in a higher total loss. We further hypothesize that the biases will be more pronounced for the longer sequence pairs Q and S than for the shorter pairs R and T. This is because the longer sequences force subjects to make more holistic evaluations while the explicit loss experiences might be recalled in the retrospective evaluation of the shorter pairs. Due to our experimental design, we have

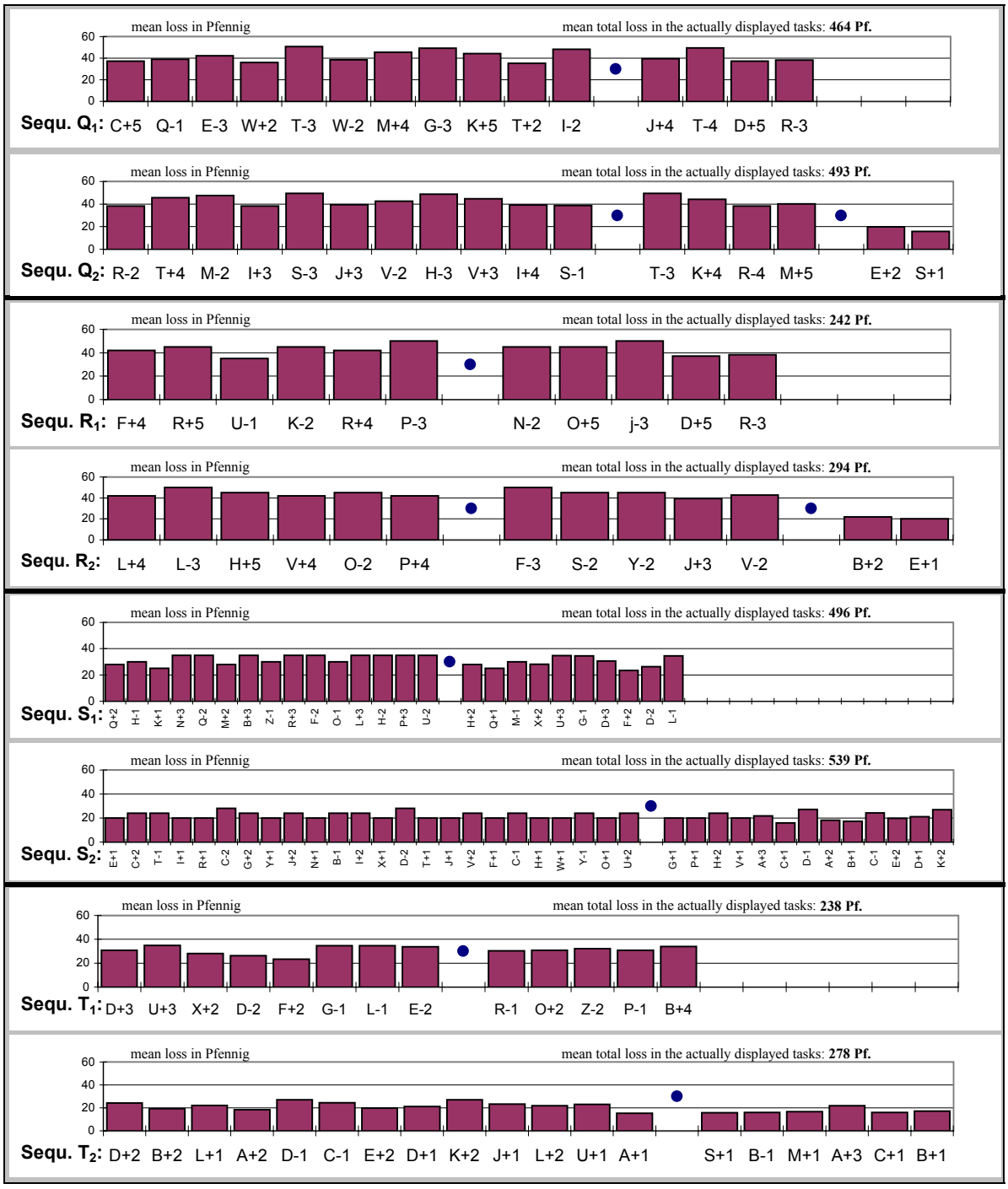

exhibit 6: Letter tasks and mean losses in experiment 2.

hardly any control over individual peak loss sizes.[8] Thus we could not include sequence pairs that were explicitly designed for an examination of peak effects. However, we will later provide some results on the impact of peaks on the choices in the given data.

Results:

Overall, 168 business students of the University of Mannheim took part in different sessions of experiment 2. For each subject, three of the four pairs were randomly selected and presented in random order with a random ordering of the sequences within each pair. From the 504 data points, 7 were eliminated because of computer problems. In these cases the computer unexpectedly ran out of extra tasks due to a very asymmetric performance in the sequences (very bad performance in the first sequence, good performance in the second sequence). On average subjects earned 12,20 DM, with individual payments varying from 6,90 DM to 17,60 DM. The experiment took about 25 minutes.

The main results of experiment 2 are displayed in exhibit 7. In all four pairs, the majority of subjects chose in line with our hypotheses. For the pairs Q, R and T an alternative explanation of random choices can be rejected on a 1% level. For the pair R the difference (53% vs. 47%) is not significant. In general the effect is much stronger for the duration-pairs S and T (79% in line with our hypothesis in both cases) than for the end-pairs Q and R (65% and 53%).

	Pair Q		Pair R		Pair S		Pair T	
	end - long		end-short		duration - long		duration - short	
	Q_1	Q_2	R_1	R_2	S_1	S_2	T_1	T_2
		hyp.		**hyp.**		**hyp.**		**hyp.**
No. of choices (total):	**43**	**80**	**58**	**65**	**27**	**103**	**25**	**96**
Percentage	*35%*	*65%*	*47%*	*53%*	*21%*	*79%*	*21%*	*79%*
Percentage if $*_2$ presented first	*41%*	*59%*	*54%*	*46%*	*18%*	*82%*	*31%*	*69%*
Percentage if $*_2$ presented second	*30%*	*70%*	*40%*	*60%*	*23%*	*77%*	*9%*	*91%*
No. of choices (without first pair):	**25**	**44**	**43**	**38**	**23**	**77**	**15**	**64**
Percentage	*36%*	*64%*	*53%*	*47%*	*23%*	*77%*	*19%*	*81%*
No. of judgments „less difficult" (total):	**40**	**83**	**54**	**69**	**31**	**99**	**16**	**105**
Percentage	*33%*	*67%*	*44%*	*56%*	*23%*	*76%*	*13%*	*87%*

exhibit 7: Choices and difficulty judgments in experiment 2.

8 In fact, it turned out that in 65% of all cases there was an identical peak loss size of 60 in both sequences. In only 20% of the pairs, all tasks in both sequences were solved correctly and in time.

One might guess that this is partly due to the specific experimental setting: Our computer algorithm does ensure the total loss difference to be in an appropriate range. It cannot make sure, however, that the task difficulties always induce the intended payment profiles. This problem looks more severe for the end-pairs, in which the intended pattern can be destroyed by single tasks (e.g., an unfortunate slip in the final easy task). It thus could be argued that the effect is less pronounced for the pairs Q and R since there may not be an end effect in the payment sequences for a considerable number of pairs. An analysis of the data, however, rejects this explanation. In only 4 of the 246 presented end-pairs (Q and R), the sum of two final losses in sequence $*_2$ was higher than the sum of the final losses in $*_1$. We find a similar number of inappropriate profiles for the duration-pairs S and T. In only 3 out of 251 pairs, the actually displayed sequence $*_2$ was shorter than the respective sequence $*_1$.

In the choice data, we observe an order effect (s. exhibit 7). For the pairs Q, R and T, the percentage of $*_2$-choices is significantly higher if $*_2$ is presented as the second sequence (and vice versa). There is no significant order effect for the pair S. A more thorough analysis shows that the order effect is almost entirely driven by the pairs that appeared as the first pair in an individual treatment..[9] To eliminate the effect of the first pair, exhibit 7 presents choice data for the second and third pairs as well. When one restricts the attention only to the second and third pairs, the results and their significance remain unchanged for the pairs Q (64%), S (77%) and T (81%). In the pair R (47%), we now find a majority of R_1-choices, but the results are statistically insignificant. Our hypotheses regarding the impact of sequence length on the strength of the effect are only confirmed for the end-pairs Q and R. While the lower end losses of $*_2$ strongly influence choices in the long pair Q, the effect is much weaker for the short pair R. In the duration-pairs S and T, we have the same strong effect for both sequence lengths. It could be argued that the short duration pair T is already too long (the median number of tasks actually displayed in pair T is 21, compared to 13 in the short end-pair R) to allow a recall of the experienced losses for the retrospective evaluation. Finally, exhibit 7 presents data on the „judged difficulty“ of the sequences. By this formulation, we switch to a different type of stimuli (perceived difficulty instead of explicit payments) without giving up the existence of a clear normative aggregation rule. The results for the difficulty question are very similar to the results for the payment choices. Thus, the type of question asked for retrospective evaluation, choice or degree of difficulty, does not influence the direction or the strength of our results.

---

9 This pattern could be explained by learning: In the course of the experiment, subjects get acquainted with the tasks and improve their performance (The data shows that the average loss in the first pair is 40 Pfennig higher than in the other pairs). Anticpating their personal improvement and not knowing that the second sequence is adjusted to performance by the computer, subjects with no clear preference make a smart decision by choosing the second sequence in the pair.

In the motivation for the design of this experiment we have already mentioned that we cannot expect any profile features of the payment sequences to influence choices if the difference in total losses between the sequences is too high and thus too obvious. More general, it can be predicted that end effects and duration neglect will be the more pronounced the smaller there is a difference in total losses. In exhibit 8, we present aggregated data on the relation between total loss differences and choices, which confirms this prediction. In the cases with a loss difference not greater than 40 Pfennig (i.e., in the clusters 0-20 and 21-40), more than 76% of the subjects chose the sequence with the higher losses and thus behaved in line with our hypotheses. It is interesting to note that even for the worst cluster, in which the loss difference was higher than 60 Pfennig, we still have a majority of subjects choosing in line with our hypotheses.

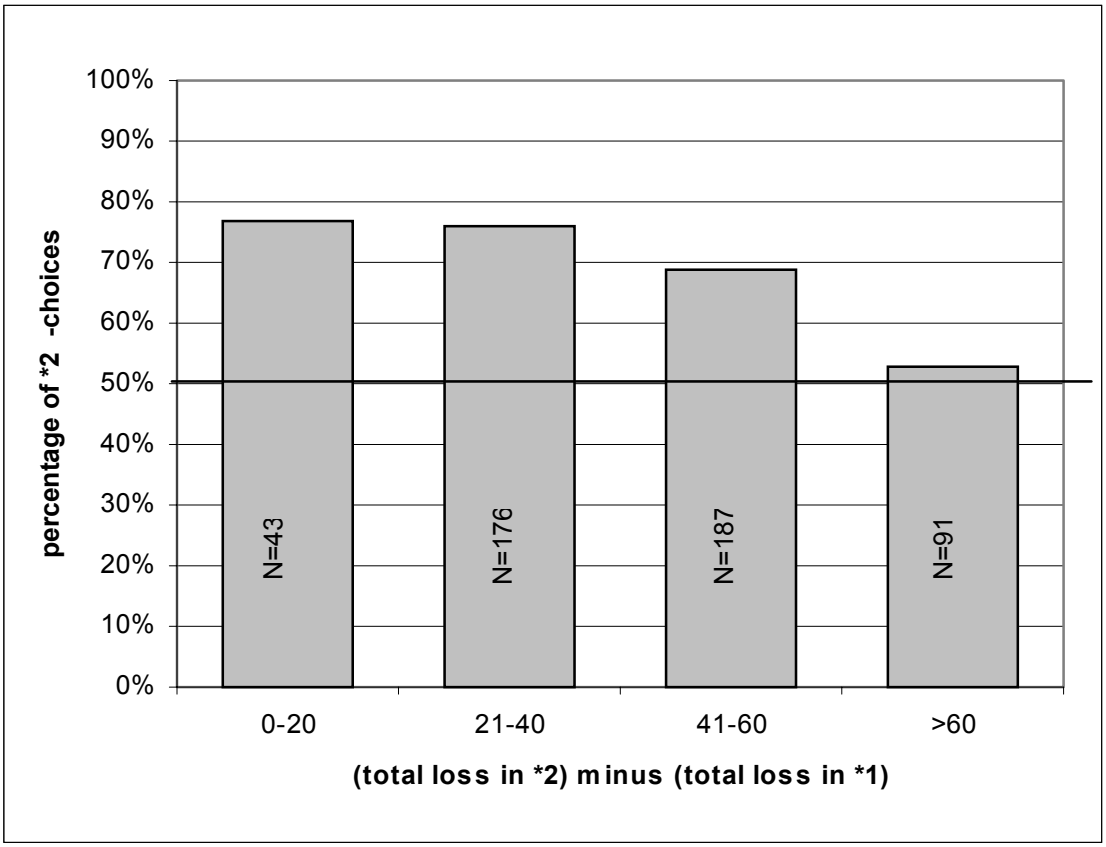

exhibit 8: Impact of total loss difference on choices in study 2.

Finally, we present some results on the impact of the peak losses on the choices. There is not much variability in peak size in the data. In more than 65% of all cases there was at least one error and thus an identical peak loss size of 60 Pfennig in both sequences. Less than 6% of the pairs were solved without any errors. Given this data an examination of peak size differences does not promise to be insightful. Instead, we analyze how the difference in the number of peaks influences the choices. We computed for each pair and each individual the difference in the number of errors between sequence $*_2$ and $*_1$ (s. exhibit 9). We find that the percentage of $*_2$

choices monotonically decreases in this difference. If the intended payment pattern (constructed to result in end or duration neglect effects) is additionally supported by a lower number of peak losses in $*_2$ the choice anomaly is even more pronounced.

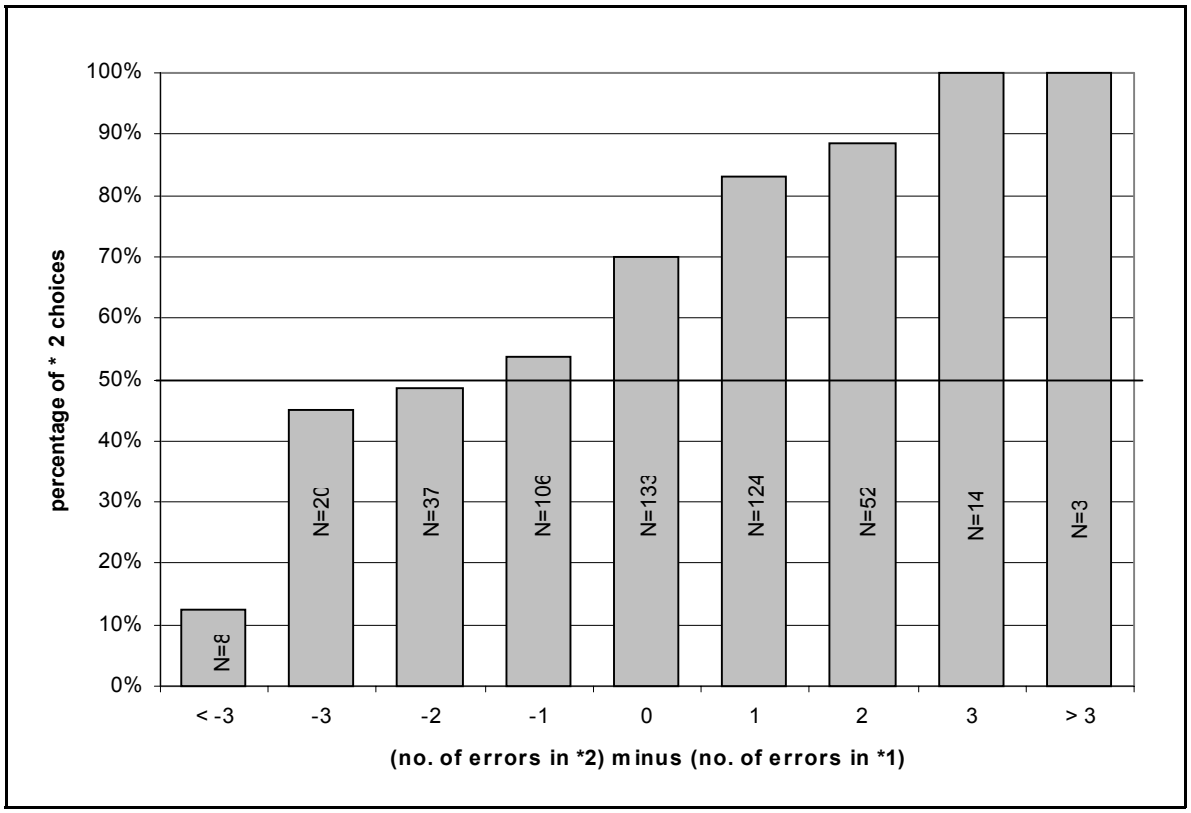

exhibit 9: Impact of error number on choices in study 2.

IV. Summary

In this paper, we address the question of whether individuals are able to make good retrospective evaluations of payment sequences. We particularly examine if systematic biases (peak and end effects and duration neglect) identified in empirical research on the retrospective evaluation of affective episodes, as extended pain experiences translate to the monetary context. Retrospective evaluations of monetary sequences differ in many important ways from retrospective evaluations of affective episodes. The most important difference is the existence of a unique, obvious and easy to apply aggregation rule for the payment sequences. This is in contrast to the scenario of affective episodes. It is not obvious at all, how a profile of instant utility levels (pain or pleasure) should be aggregated into an overall evaluation.

In the monetary scenario individuals are aware that they should choose the sequence with the highest total payment. These facts were confirmed in our first experiment. When confronted with sequences of numbers (with or without a monetary interpretation), most subjects tried to avoid

the need for a possibly biased retrospective evaluation by performing an online aggregation of the payments. Here we did not observe a neglect of sequence length (duration) or any peak or end effects.

In our second experiment, we linked the payments to personal effort and performance in strenuous tasks. There are two important consequences: first, this linkage makes the payments more affective stimuli; second, the distraction through the tasks forces subjects to construct their evaluation indeed retrospectively, based on their memory of the payment sequences. In this experiment, we observe a strong effect of duration neglect. A significant majority of subjects chose for their real payment the longer sequence that had, on average, less difficult tasks, but the higher total loss at the same time. We also observed an end effect. Most subjects chose for their real payment the sequence with the lower end losses, though it had the higher total loss. Finally we observed a specific type of peak effect. The number of peak losses in a sequence determined its (un)attractiveness. A sequence with the higher total loss was more likely to be chosen for real payment if it had the lower number of peak losses at the same time.

Our experimental research demonstrates that systematic biases like peak and end effects or duration neglect do not occur only in the retrospective evaluation of continuous pain and pleasure streams, but also persist in monetary settings with its discrete and objective stimuli and its straightforward normative aggregation procedure. However, the occurrence is restricted to situations in which online aggregation cannot take place and individuals have to make indeed retrospective evaluations of the payment sequences.

In future research it should be examined whether other results on the robustness of the biases in pain and pleasure settings also translate to the payment scenario. Ariely (1998) found that the duration of an episode is less neglected if the volatility in the stimulus intensity is higher. He also derived a slightly different conclusion regarding the peak-and-end effect. From his data, he concluded that „...*the retrospective evaluations of painful experiences are influenced primarily by a combination of the final pain intensity and the intensity trend during the latter half of the experience*." Both these results are easily transferable and testable in a monetary setting. Ariely and Loewenstein (forthcoming) pinpoint the varying strength of the duration neglect effect (for aversive sounds) dependent on the implemented elicitation procedure (rating or choice) and/or the comparative status (joint or isolated evaluation). It should be examined whether their results translate to monetary settings.

Our results have provided insights into the role of duration neglect, end effects and the effect of number of peak losses in the choice of payment sequences. In many real economic situations, retrospective judgments of past payment sequences are made. It would be interesting to study the economic situations where the biases produced by duration neglect, peak and end effects and possibly the number of peak losses are most pronounced.

Literature

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

Ariely, D. and Carmon, Z. (2000): *Gestalt characteristics of experiences: The defining features of summarized events*. Journal of Behavioral Decision Making, Vol. 13, 191-201.

Ariely, D. and Loewenstein, G. (in press): *When does duration matter in judgment and decision making?* Journal of Experimental Psychology: General.

Ariely, D. and 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, Vol. 13, 219-232.

Baumgartner, H. Sujan, M. and Padget, D. (1997): *Pattern of affective reactions to advertisements: The integration of moment-to-moment responses into overall judgments*. Journal of Marketing Research, Vol. 34, 219-232.

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

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

Kahneman, D.; Wakker, P. and Sarin, R. (1997): *Back to Bentham? Explorations of Experienced Utility*. Quarterly Journal of Economics, Vol. 112, 375-405.

Loewenstein, G. and Prelec, D. (1991): *Negative time preference*. American Economic Review: Papers and Proceedings, Vol. 82, 347-352.

Loewenstein, G. and Prelec, D. (1993): *Preferences for sequences of outcomes*. Psychological Review, Vol. 100, 91-108.

Loewenstein, G. and Sicherman, N. (1991): *Do workers prefer increasing wage profiles*? Journal of Labor Economics, Vol. 9, 67-84.

Redelmeier, D.A. and Kahneman, D. (1996): *Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures.* Pain, Vol. 66, 3-8.

Varey, C.A. and Kahneman, D. (1992): *Experiences extended across time: Evaluation of moments and episodes*. Journal of Behavioral Decision Making, Vol. 5, 169-185.

Nr.	Author	Title
01-12	Peter Albrecht Raimond Maurer Ulla Ruckpaul	On the Risks of Stocks in the Long Run: A Probabilistic Approach Based on Measures of Shortfall Risk
01-11	Peter Albrecht Raimond Maurer	Zum systematischen Vergleich von Rentenversicherung und Fondsentnahmeplänen unter dem Aspekt des Kapitalverzehrrisikos - der Fall nach Steuern
01-10	Gyöngyi Bugàr Raimond Maurer	International Equity Portfolios and Currency Hedging: The Viewpoint of German and Hungarian Investors
01-09	Erich Kirchler Boris Maciejovsky Martin Weber	Framing Effects on Asset Markets - An Experimental Analysis -
01-08	Axel Börsch-Supan Alexander Ludwig Joachim Winter	Aging, pension reform, and capital flows: A multi-country simulation model
01-07	Axel Börsch-Supan Annette Reil-Held Ralf Rodepeter Reinhold Schnabel Joachim Winter	The German Savings Puzzle
01-06	Markus Glaser	Behavioral Financial Engineering: eine Fallstudie zum Rationalen Entscheiden
01-05	Peter Albrecht Raimond Maurer	Zum systematischen Vergleich von Rentenversicherung und Fondsentnahmeplänen unter dem Aspekt des Kapitalverzehrrisikos
01-04	Thomas Hintz Dagmar Stahlberg Stefan Schwarz	Cognitive processes that work in hindsight: Meta-cognitions or probability-matching?
01-03	Dagmar Stahlberg Sabine Sczesny Friederike Braun	Name your favourite musician: Effects of masculine generics and of their alternatives in german

Nr.	Author	Title
01-02	Sabine Sczesny Sandra Spreemann Dagmar Stahlberg	The influence of gender-stereotyped perfumes on the attribution of leadership competence
01-01		Jahresbericht 2000
00-51	Angelika Eymann	Portfolio Choice and Knowledge
00-50	Oliver Kirchkamp Rosemarie Nagel	Repeated Game Strategies in Local and Group PrisonersÍDilemmas Experiments: First Results
00-49	Thomas Langer Martin Weber	The Impact of Feedback Frequency on Risk Taking: How general is the Phenomenon?
00-48	Niklas Siebenmorgen Martin Weber	The Influence of Different Investment Horizons on Risk Behavior
00-47	Roman Inderst Christian Laux	Incentives in Internal Capital Markets
00-46	Niklas Siebenmorgen Martin Weber	A Behavioral Approach to the Asset Allocation Puzzle
00-45	Thomas Langer Rakesh Sarin Martin Weber	The Retrospective Evaluation of Payment Sequences: Duration Neglect and Peak-and-End-Effects
00-44	Axel Börsch-Supan	Soziale Sicherung: Herausforderungen an der Jahrhundertwende
00-43	Rolf Elgeti Raimond Maurer	Zur Quantifizierung der Risikoprämien deutscher Versicherungsaktien im Kontext eines Multifaktorenmodells
00-42	Martin Hellwig	Nonlinear Incentive Contracting in Walrasian Markets: A Cournot Approach
00-41	Tone Dieckmann	A Dynamic Model of a Local Public Goods Economy with Crowding
00-40	Claudia Keser Bodo Vogt	Why do experimental subjects choose an equilibrium which is neither risk nor payoff dominant