Judgment and Decision Making, Vol. 5, No. 5, August 2010, pp. 347–364

Encoding, storage and judgment of experienced frequency and duration

Tilmann Betsch,* Madlen Glauer, Frank Renkewitz
University of Erfurt, Germany

Isabell Winkler, Peter Sedlmeier
Chemnitz University of Technology, Germany

Abstract

This paper examines conditions that do or do not lead to accurate judgments of frequency (JOF) and judgments of duration (JOD). In three experiments, duration and frequency of visually presented stimuli are varied orthogonally in a within-subjects design. Experiment 1 reveals an *asymmetric* judgment pattern. JOFs reflected actual presentation frequency quite accurately and were unbiased by exposure duration. Conversely, JODs were almost insensitive to actual exposure duration and were systematically biased by presentation frequency. We show, however, that a tendency towards a *symmetric* judgment pattern can be obtained by manipulating encoding conditions. Sustaining attention during encoding (Experiment 2) or enhancing richness of the encoded stimuli (Experiment 3) increases judgment sensitivity in JOD and yields biases in both directions (JOF biased by exposure duration, JOD biased by presentation frequency). The implications of these findings for underlying memory mechanisms are discussed.

Keywords: frequency processing, time perception, quantitative judgment, long-term memory.

1 Introduction

Event frequency and duration are of fundamental importance in behavioral adaptation (Gallistel, 1989; Schwartz & Reisberg, 1991). Predators, for instance, increase their chances of survival if they approach areas where the prey resides more frequently and for longer periods of time. Conversely, prey should strive to avoid areas where predators roam frequently and consistently. Consider an individual who must decide where to go hunting: the probability of maximizing the individual's yield is a function of the frequency and length of time (duration) prey lingers in a certain area. Assume the individual is an experienced hunter who has often visited different areas in the past and observed the behavior of the prey animals. Further assume that the observations made would represent a valid sample of the prey animals' migrations. In order to properly adapt his own behavior to the environment, the hunter should use frequentistic *and* temporal information in his subsequent choice (Attneave, 1953; Herrnstein, 1961). This requires two abilities: first, the individual must be capable of *discriminating* differences in frequency and duration; second, representations of frequency and duration must be *stored* in memory so that they can be used later for *judgment and decision making*.

Psychologists from various fields have studied processing of duration and frequency extensively. Their approaches, however, differ. Research in animal cognition and neuropsychology focuses primarily on discrimination and short term storage (e.g., Meck, 2003). Cognitive psychologists are often interested in memory processes (e.g., Lewis & Miall, 2006; Zacks & Hasher, 2002), whereas JDM researchers mainly consider judgmental heuristics (e.g. Kahneman, Slovic & Tversky, 1982). These fields host different theoretical perspectives, apply different research paradigms and arrive at diverging interpretations of the phenomena.

Consider, for example, the case of frequency judgment. Cognitive psychologists normally employ learning procedures to induce frequency knowledge in the laboratory and usually find that participants subsequently make quite accurate judgments. They explain their findings with reference to general models of memory (e.g., Hintzman, 1988; Hasher & Zacks, 1984). Conversely, JDM researchers often investigate factors that cause changes in judgment accuracy and explain their results with reference to judgmental heuristics (see Sedlmeier, Betsch & Renkewitz, 2002, for a discussion).

---

*This research was supported by grants from the German Science Foundation (Deutsche Forschungsgemeinschaft) to Tilmann Betsch (BE 2012/5–1) and Peter Sedlmeier (SE 686/8–1). We thank two anonymous reviewers for valuable comments. We are grateful to Judith Bernauer, Stefanie Häusig, Marei Krüger, Justus Julius Wolfgang Kunz, Frank Lindow and Nicola Schindler for running the experiments. Address: Tilmann Betsch, Department of Psychology, University of Erfurt, PO Box 900 221, D-99105 Erfurt, Germany. Email: tilmann.betsch@uni-erfurt.de.

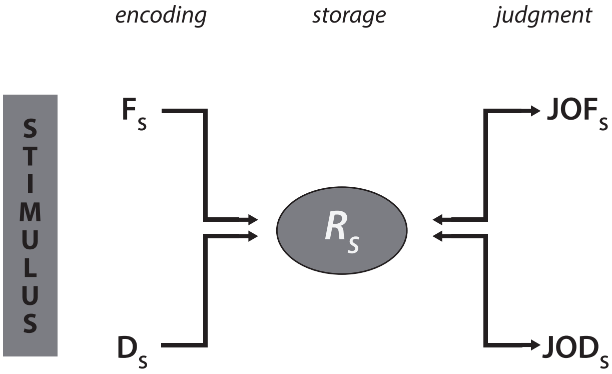
The separation between research fields has been often lamented, but there is an increasing number of researchers striving towards integration both on the methodological and the theoretical level (e.g., Dougherty, Gettys & Ogden, 1999; Fiedler, 2002, Thomas, Dougherty, Sprenger & Harbison, 2008; Weber, Goldstein & Barlas, 1995; Weber, Johnson, Milch, Chang, Brodscholl & Goldstein, 2007)

In this paper, we report studies using a research paradigm from cognitive psychology to induce knowledge about frequency and duration. We compare judgments in two related domains, judgments of duration (JOD) and judgments of frequency (JOF). We manipulate conditions of encoding and assess their effects on judgmental patterns. Specifically, we compare JOD and JOF regarding differences in their retrospective sensitivity and differences in their susceptibility for biases. With this approach, we seek to increase our knowledge of fundamental memory processes. Such knowledge may help to achieve a better understanding of variations in judgment. In the remainder of the introduction, we give a brief overview of the state of research on processing of frequency and duration and arrive at a working hypothesis regarding storage of these entities in memory.

Ample evidence indicates that organisms are remarkably good at *discriminating* differences in frequency and duration (e.g., Gallistel, 1989). Like animals (Meck & Church, 1983), adult humans (Hasher & Zacks, 1984; Whalen & Gallistel, 1999; Zacks & Hasher, 2002) and even young children (Droit-Volet & Wearden, 2001) are capable of effectively discriminating low range frequencies ($n < 10$). Similarly, humans, like many other species, are able to discriminate temporal information across a wide range of intervals — from circadian timing to the timing of seconds and even milliseconds (Buhusi & Meck, 2005). The discrimination of very short durations in the millisecond range, in particular, seems to be very accurate (Lewis & Miall, 2006; Rammsayer, 2003). All together, research findings suggest that a common mental mechanism drives discrimination of both stimulus dimensions — frequency and duration (Walsh, 2003; but see Dormal et al., 2006).

Granting the findings on discrimination, one may be tempted to assume that frequency and duration would also affect *storage* in memory in a similar fashion. According to such a common-path hypothesis (Figure 1), frequency and duration would have commensurate effects on a stimulus' representation (R_S) and subsequent judgment (e.g., Cooper & Pantle, 1967; but see Hintzman, 2004a, for a critical discussion of this hypothesis.[1])

Figure 1: A common-path hypothesis for storage and judgment of frequency (F) and duration (D). S= stimulus; R_S = representation of the stimulus in memory.



For instance, they may both change the accessibility or strength of an event in memory.

Changes in strength could be modeled by different theoretical approaches to memory. In connectionist-network models, frequency and duration may be conceived to change the strength of associations among R_S, its features and other constituents in the network (e.g., PASS: Sedlmeier, 1999, 2002, 2008). In multiple-trace models (e.g., MINERVA DM: Dougherty, Gettys and Ogden, 1999; MINERVA 2: Hintzman, 1988; Dougherty & Franco-Watkins, 2002), strength is primarily conceived as a function of the number of traces representing a stimulus or event, at least if encoding and response criteria are constant. Other things being equal, trace number is assumed to increase as a function of encounter frequency. It may also increase as a function of duration because prolonged exposure may entail repetitive encoding of the event.

Regardless of the underlying mechanism, a common-path hypothesis assumes that the same property of memory is used in subsequent judgments. Accordingly, JOF and JOD should be made through assessments of the strength of R_S in memory. One main prediction to be derived from such a common-path hypothesis is that JOF and JOD should be *symmetrically* biased. Accordingly, JOF should be biased by exposure duration, while JOD should be biased by exposure frequency. Empirical evidence for the common-path hypothesis, however, is mixed. Supporting evidence stems mostly from JDM research on heuristics that are assumed to use memory strength as a predictor for quantitative judgments. For example, when applying the availability heuristic (Tversky & Kahneman, 1973), an individual is thought to estimate a criterion by assessing the ease with which the stimulus (stimuli) can be rehearsed from memory. The fluency heuristic (e.g., Jacoby & Brooks, 1984) employs a similar process: the more fluently a stimulus is processed (due to

---

[1] Note, however, that Hintzman (2004a) discussed a common-path model with reference to recognition confidence ratings (RCR) and judgments of frequency (JOF) but not with respect to judgments of duration (JOD).

previous exposure), the higher the estimation of the criterion (Schooler & Hertwig, 2005, p. 612).

In line with this notion, some studies have provided evidence that exposure duration biases subsequent JOF. For example, Williams and Durso (1986, Exp. 2) varied the duration of words (0.5, 3, 6 sec) between subjects. They found that subsequent frequency judgments were systematically biased by duration. Specifically, the longer the category exemplars were shown in the presentation, the higher the estimated size of the category, and the higher the probability that exemplars could be reproduced in a free recall test. Lewandowsky and Smith (1983) found similar effects earlier. Employing a within-subjects design, they varied exposure duration (1.55, 3.75, 5 sec) and assessed the effects on frequency judgments and recall. Both measures were positively correlated with exposure duration. Similarily, JOD were found to be systematically biased by frequency. Ornstein (1969) showed that retrospective judgments of time intervals reflected the number of available events that filled the episode. Hintzman (1970) reported evidence showing that JOD systematically reflected exposure frequency of the stimulus (but see the discussion of this paper in the next section).

Taken together, these findings support the assumptions of a common-path hypothesis, because JOF and JOD seem to behave *symmetrically* with regard to biases. A conclusive test, however, requires us to consider two other aspects of the judgment pattern. One is the relative strength of the biases, the other is judgment sensitivity, Judgment sensitivity refers to the degree to which experienced *differences* in duration and frequency can be reproduced in judgments. To this end, one must compare JOF and JOD within the same experiment. Such an approach was applied earlier in memory research.

In a seminal paper, Hintzman (1970) varied frequency on four levels (1, 2, 3, 5 repetitions) and duration on five levels (2, 3, 4, 5, 6 seconds) in a within-subjects design. After the presentation he asked participants to either judge the frequency with which each word occurred or the duration it was shown on an average trial. Hintzman found that JOF showed a remarkable sensitivity for actual differences in exposure repetition. Moreover, JOF were only weakly affected by differences in duration. Conversely, judgment sensitivity was comparatively lower in JOD. As another difference, and already noted in the previous section, JOD were systematically influenced (biased) by actual frequency. Hintzman (1970, p. 442) concluded that the asymmetric pattern of results (differences in retrospective sensitivity, uni-directional bias) violates a common-path hypothesis and suggests that frequency and duration affect memory in a *different* fashion. Hintzman reported similar results in subsequent studies (e.g., Hintzman, Summers & Block, 1975; for an overview: Hintzman, 2000). All together, empirical evidence from this line of research suggests that frequency and duration are stored differently in memory.

The findings reported so far form an incoherent picture. Duration biases on JOF sometimes occur and sometimes do not. JOD seem to be more susceptible to biases than JOF, but there is also evidence for the opposite (Hintzman, 2004b). There are several possible explanations for the mixed evidence. An important one is that the type of judgment task differs between studies. As for JOD, some studies use retrospective tasks (e.g., Hintzman, 1970: participants do not expect to judge duration) and some use prospective tasks (e.g., Hintzman, 2004b: participants expect to judge duration; see Hicks, Miller & Kinsbourne, 1976, for a discussion of task distinction). As for JOF, most studies documenting duration biases on JOF used set-size tasks. That is, participants estimated the size of categories containing different exemplars (e.g., Lewandowsky & Smith, 1983; Williams & Durso, 1986). In contrast, duration biases on JOF are less pronounced in tasks that assess judgment of event frequencies (e.g., Hintzman, 1970, Hintzman et al., 1975; see Betsch et al., 1999 and Manis et al., 1993, for a discussion of task distinction).

Another possible explanation is that *encoding conditions* may differ across tasks and studies. Encoding of frequency requires comparatively little attention to the stimulus. One must simply realize that the stimulus occurs. Encoding of duration, however, requires that the individual additionally sustains attention to the stimulus for the entire period of its presence. Differences in procedures and materials may differentially affect the individual's inclination to attend to the stimuli proportionate to actual duration. In the literature on time perception, it is a widespread notion that attention is a crucial variable in time processing (e.g., Block, 1990; Macar, 2005; Zakay, 2005; Zakay & Block, 1996, 1997). Consequently, differences in result patterns might be caused by *differences in encoding* rather than by *differences in memory*. Without systematically considering encoding conditions, it might be a premature conclusion to dismiss a common-path hypothesis.

The present research focuses on the role of encoding conditions in JOD and JOF. In three experiments, we varied total duration and frequency of stimuli orthogonally within the same experimental design. Specifically, the presentation of the stimuli differed with regard to repetition frequency (2, 4, 8 times) and the total (summed) duration they appeared on the screen (8, 16, 24 sec). The resulting 3 (frequency) x 3 (duration) factorial design was always implemented within subjects. Following stimuli presentation, participants made JOF and JOD. Specifically, they judged the frequency of the event's occurrence, its total duration (Exps. 1, 2, 3), and its single (average) duration (Exp. 3).

We generally assess the pattern of judgment as the central dependent variable. It takes two dimensions into account — *judgment sensitivity* and *biases*. Judgment sensitivity refers to the extent to which individuals are able to reproduce differences on stimulus dimensions in their judgments. Biases are indicators for the type of memory representation underlying the judgment. Experiment 1 serves as a baseline study in which we did not control for encoding conditions. Experiment 2 compares different conditions of the focus of attention during encoding. Experiment 3 compares different types of stimulus formats that differ in likelihood of holding the observer's attention on the stimuli during encoding.

2 Experiment 1

2.1 Method

2.1.1 Participants and design

Sixty undergraduates (39 female; mean age: 24.7) from different academic majors at Chemnitz University of Technology participated in the experiment, lasting about 30 minutes, and were paid € 6 each. The design was a 3 (presentation frequency) x 3 (duration) full-factorial within-subjects design. Accordingly, the stimuli differed with regard to repetition frequency (2, 4, 8 times) and the total duration (8, 16, 24 sec) they are presented.

2.1.2 Materials

Participants saw 36 forenames, 18 female and 18 male. In each condition of the 3x3 design, two female and two male names were shown. In total, the presentation list contained 168 items. In a pretest, names were selected according to prevalence, liking and readability. First, a sample of names was randomly drawn from a list of one hundred names representing those currently most prevalent in Germany (based on a 10-year statistic of civil registry offices). We then selected those names with scores within the second and third quartile on a scale of likeability (Rudolph, 2001). Finally, names were tested for readability. Specifically, we selected only those names that could easily be read aloud within the time constraints resulting from our manipulation of duration. According to the above design, the shortest duration of a single stimulus was one second (if the stimulus appeared 8 times, and the total duration was 8 seconds). Thus, participants had to be able to read a name aloud within one second. The 36 names that passed pre-testing and were used in the experiments are shown in the Appendix.

2.1.3 Procedure

The presentation of stimuli and assessment of dependent measures were completely computer-directed. The software was programmed in E-Prime (version 1.1.4.15) and was run under a Windows XP environment on IBM compatibles equipped with 17-inch flat-screen monitors. For each individual session, the program randomly drew forenames from the pre-tested base list to build a stimulus sample. Additionally, the stimuli were randomly assigned to conditions resulting from the 3x3 design. We also made sure that male and female names appeared equally often within each condition. Stimulus words always appeared in the center of the screen.

Participants were individually tested in separate cubicles. Upon arrival, they were informed that they would view a list of words appearing consecutively on the screen and that they would be asked some questions about the presented stimuli later. They were not informed that later they would judge frequency and duration in order to prevent them from explicitly counting or timing (post-experimental interviews indicated that none of the participants was aware about the actual purpose of the study). We ensured encoding of the stimuli by having participants read each stimulus word aloud (Johnson et al., 1989). To foster commitment to this instruction, we installed microphones in the cubicles and informed participants that their utterances would be recorded. During a training phase, participants were familiarized with the presentation format and the read-aloud procedure. The training presentation lasted 2 minutes and contained a list of forenames that were not contained in the subsequent presentation.

After the presentation, participants were presented with all stimuli again in random order. They were asked to judge the frequency and the total duration (in seconds) with which each name appeared on the screen. Judgments were blocked. Accordingly, participants first estimated the frequency of all stimuli and then judged their duration (or vice versa). In two pilot studies (Fischbach, 2004) using a similar design (but with different kind of stimuli) we found no order effects for judgment blocks. Therefore, we refrained from manipulating order as an independent variable but we counterbalanced the order of presentation.

2.2 Results

2.2.1 Memory-based sensitivity for frequency

Mean judgments of frequency (JOF) are presented in the upper part of Table 1. Apparently, JOF systematically co-vary with actual presentation frequencies. The observation is corroborated by individual trend analyses. All three linear trends for the factor "frequency" in the three

Table 1: Results of Experiment 1. Standard deviations are in parentheses; duration judgments are in seconds.

Mean JOF

	Presentation frequency		
Total duration	2	4	8
8 sec	4.32 (2.97)	5.71 (3.30)	7.77 (4.02)
16 sec	4.70 (3.41)	6.19 (4.17)	7.87 (4.13)
24 sec	4.39 (2.75)	5.88 (3.18)	8.04 (4.18)

Mean JOD

	Presentation frequency		
Total duration	2	4	8
8 sec	6.95 (6.41)	9.67 (8.15)	12.64 (10.24)
16 sec	8.50 (8.72)	10.37 (9.53)	12.41 (9.92)
24 sec	8.37 (7.75)	10.15 (9.07)	13.36 (11.61)

duration conditions are significant and strong in effect size: $F_{8\,\text{SEC}}$ (1, 58) = 84.4, $p < .01$, $\eta^2 = .59$; $F_{16\,\text{SEC}}$ (1, 58) = 91.7, $p < .01$, $\eta^2 = .61$; $F_{24\,\text{SEC}}$ (1, 58) = 88.5, $p < .01$, $\eta^2 = .60$. Accordingly, the effect for the repeated measure factor "frequency" is also significant: F (1.4, 82.4) = 102.2, $p < .01$, $\eta^2 = .63$. The results demonstrate that participants are capable of reliably differentiating between the presentation frequencies of the 36 names. In other words, memory-based judgments reveal a high sensitivity to differences in actual frequencies.

2.2.2 Impact of duration on JOF (duration bias)

A column-by-column comparison of means in the upper part of Table 1 reveals that frequency judgments are *not* systematically influenced by duration. According to this observation, none of the single trend analyses produces reliable effects for the factor "duration" (all $F < 1$). Neither the main effect for the repeated measure factor "duration" nor the duration-frequency-interaction effect is significant (both $F < 1.6$; $p > .20$). These insignificant effects indicate that our participants' JOF are virtually immune to biases of duration.

2.2.3 Memory-based sensitivity for duration

Mean judgments of duration (JOD) are presented in the lower part of Table 1. Inspection of the means and standard deviations suggest that participants are almost insensitive to differences in duration. Only when presentation frequency is small do participants detect the difference in duration, namely, between 8 sec and 16 sec. In spite of the fact that they are insensitive to a 16-sec to 24-sec difference, the single trend analysis for the first frequency condition produces a significant effect: $F_{2\,\text{TIMES}}$ (1, 58) = 6.63; $p < .05$, $\eta^2 = .10$. In the other two conditions, the linear trend analyses produce insignificant effects ($F_{4\,\text{TIMES}}$, $F_{8\,\text{TIMES}} < 1.4$). The main effect for the repeated measure factor "duration" is weak and only marginally significant: F (1.96, 116.2) = 2.5, $p < .10$, $\eta^2 = .04$. The results indicate that participants can only occasionally reproduce differences in actual durations in their judgments.

2.2.4 Impact of frequency on JOD (frequency bias)

A row-by-row comparison of means in the lower part of Table 1 reveals that JOD systematically increase with increasing presentation frequency, even though total duration is held constant. All single linear trends for the factor "frequency" in the three duration conditions are significant and of considerable effect size: $F_{8\,\text{SEC}}$ (1, 58) = 42.9, $p < .01$, $\eta^2 = .43$; $F_{16\,\text{SEC}}$ (1, 58) = 26.7, $p < .01$, $\eta^2 = .32$; $F_{24\,\text{SEC}}$ (1, 58) = 31.5, $p < .01$, $\eta^2 = .35$. Accordingly, the main effect for the repeated measure factor "frequency" is also significant, F (1.7, 98.9) = 36.3, $p < .01$, $\eta^2 = .39$. There is no interaction effect ($F < 1.5$). The results indicate that in our experiment JOD are strongly and systematically biased by presentation frequency.

2.3 Discussion

The results of the first experiment revealed an asymmetrical judgment pattern. JOF were generally and strongly sensitive to actual differences in frequency, whereas JOD were almost insensitive to actual differences in duration. Moreover, JOD were consistently biased by frequency, whereas JOF were immune against biases by duration. The finding that JOD and JOF behave asymmetrically supports findings reported by Hintzman (1970). In addition, the results challenge the common-path hypothesis, which assumes that storage and judgment of frequency and duration are driven by common mechanisms of storage.

The asymmetrical judgment pattern seems to corroborate the notion that individuals are not able to store unintentionally differences in duration as effectively as they can store differences in frequency. One reason for this could be that variations in frequency are more likely to be encoded than variations in duration. Why should this be the case?

Hintzman (1970) proposed (but never tested) an interesting explanation. He argued that participants attend to certain stimuli for only a constant time period: "… it ordinarily takes [a participant] only a short time (less than 2 seconds) to process a word into memory, and, after this has been done, he stops processing the stimulus until a different word is presented. According to this explanation, exposure duration has no effect because during any interval exceeding the minimal processing time, the consolidation mechanism is essentially inactive" (Hintzman, 1970, p. 443). If this explanation is valid, we should obtain a symmetrical judgment pattern with bi-directional biases under conditions that promote individuals' processing of a stimulus *for the entire exposure period.* To test this assumption, we conducted a second experiment. Its design contains different manipulations to increase the likelihood that participants remain focused on the crucial stimulus dimensions at the time of encoding. If the common-path hypothesis is valid, we should find a symmetric judgment pattern under those conditions that encourage participants to focus on the stimulus for its entire presentation duration.

3 Experiment 2

The purpose of the second experiment was to establish conditions that enhance encoding of temporal information. Research on time perception in humans indicates that the focus of attention mediates accuracy in discrimination of durations (e.g., Zakay, 2005; Zakay & Block, 1996). We attempted to manipulate encoding focus without making participants explicitly aware of the nature of the subsequent judgment task. Because we are interested in unintentional memory processes, we avoided establishing conditions of prospective timing (Hicks, Miller & Kinsbourne, 1976) that would encourage participants to actively store and remember the duration of stimuli. Rather, we had participants perform a motor task contingent upon either the temporal or frequency-related aspects of the stimulus presentation. In a *run-stop* condition, participants were instructed to press the space bar on the keyboard as long as a stimulus remained on the screen and release the key when it disappeared. In order to perform this task properly, the individual had to sustain attention to the stimulus for the entire time during which it was presented. As a control condition, the other half of the participants hit the space bar only briefly each time a stimulus appeared on the screen (*event* condition). With this manipulation we intended to move the focus of attention away from the temporal dimension. The two motor manipulations adopt a behavioral component to direct attention towards duration and frequency during encoding. The technique mimics properties of experimental procedures in animal research in which the subject is required to reproduce stimulus properties in behavior, for example, via pressing a lever for a certain number of times or duration (e.g., Meck, 2003).

As a second between-subjects factor, we varied the focus instruction. In an instruction-present condition, participants were explicitly told to focus on duration or frequency, corresponding to the motor-task condition to which they were assigned. A condition without further instruction served as a control.

3.1 Method

3.1.1 Participants and design

Seventy-seven undergraduates (60 female; mean age: 20.2) from different academic majors at the University of Erfurt participated in the experiment lasting about 20 minutes. They were paid € 6 each and were randomly assigned to four experimental groups resulting from a 2 (motor task: run-stop vs. event) x 2 (focus instruction: yes vs. no) x 3 (presentation frequency) x 3 (duration) factorial design. The last two factors were manipulated within subjects. We employed a nested within-subjects design in which duration was only varied in the mid-frequency condition (4 times) and frequency was only varied in the mid-duration condition (16 sec.).

3.1.2 Materials

Participants saw 20 forenames, 10 female and 10 male, randomly drawn from the list of names from Experiment 1. In each of the five conditions of the nested within-subjects design, two female and two male names were shown. In total, the presentation list contained 88 items.

3.1.3 Procedure

Presentation of the stimuli followed the procedure of Experiment 1. Participants were individually tested in separate cubicles. As in Experiment 1, they were informed that they would view a list of words appearing consecutively on the screen and that they would be asked some questions about the presented stimuli later. Participants assigned to the run-stop task were told: "If a word appears on the screen you have to immediately hit the red colored key (the space-bar). Continue pressing the key as long as the word appears on the screen." Participants assigned to the event task were told: "If a word appears on the screen you have to immediately hit the red colored key (the space-bar). Press the key only briefly and release it immediately." Moreover, participants were instructed to repeat the motor task each time a word appeared on the screen. In the focus-instruction condition, event-task participants were asked to "concentrate on the occurrence of the words", whereas run-stop-task participants were asked to "concentrate on the length of time the words appear". The controls did not receive any additional focus instruction.

After the presentation, participants were presented with all stimuli again in random order. They were asked

Table 2: Results of Experiment 2. Standard deviations are in parentheses; duration judgments are in seconds. Between-subjects conditions: EVENT mode with focus INSTRuction (n = 19), RUN mode with focus INSTRuction (n = 20), EVENT mode without focus instruction (n = 18), RUN mode without focus instruction (n = 20).

Mean JOF			
	Presentation Frequency		
Total duration	2	4	8
8 sec			
EVENT, INSTR		5.68 (2.39)	
RUN, INSTR		4.78 (1.34)	
EVENT		5.88 (3.07)	
RUN		5.04 (1.87)	
Grand mean		**5.32 (2.24)**	
16 sec			
EVENT, INSTR	4.38 (2.35)	5.30 (2.38)	8.42 (3.77)
RUN, INSTR	4.18 (1.86)	5.13 (2.23)	6.24 (2.30)
EVENT	3.90 (1.38)	5.67 (3.21)	8.60 (5.41)
RUN	3.94 (1.60)	5.03 (1.74)	6.76 (2.90)
Grand mean	**4.10 (1.81)**	**5.27 (2.39)**	**7.46 (3.80)**
24 sec			
EVENT, INSTR		5.88 (3.48)	
RUN, INSTR		5.35 (2.11)	
EVENT		6.36 (4.54)	
RUN		5.51 (2.01)	
Grand mean		**5.76 (3.12)**	

Mean JOD			
	Presentation frequency		
Total duration	2	4	8
8 sec			
EVENT, INSTR		10.92 (6.28)	
RUN, INSTR		10.59 (5.19)	
EVENT		10.81 (5.88)	
RUN		11.78 (9.91)	
Grand mean		**11.03 (6.97)**	
16 sec			
EVENT, INSTR	8.25 (4.14)	10.68 (5.90)	14.78 (7.90)
RUN, INSTR	11.14 (5.70)	12.64 (7.14)	13.09 (5.64)
EVENT	8.36 (4.44)	10.53 (4.78)	13.65 (6.19)
RUN	11.20 (11.37)	12.41 (9.60)	15.25 (13.35)
Grand mean	**9.79 (7.14)**	**11.60 (7.08)**	**14.20 (8.76)**
24 sec			
EVENT, INSTR		10.57 (5.63)	
RUN, INSTR		13.43 (6.26)	
EVENT		12.78 (9.22)	
RUN		14.30 (12.17)	
Grand mean		**12.80 (8.68)**	

to judge the frequency and the total duration (in seconds) that each name appeared on the screen. Judgment tasks were blocked. Block order was counterbalanced.

3.2 Results

3.2.1 Memory-based sensitivity for frequency

Mean estimates of frequency are presented in the left part of Table 2. Clearly, JOF systematically co-vary with actual presentation frequencies; the observation is corroborated by trend analyses. All linear trends for the factor "frequency" in the four between-subjects conditions are significant and strong in effect size: $F_{\text{EVENT, INSTR}}$ (1, 18) = 47.1, $p < .01$, $\eta^2 = .74$; $F_{\text{RUN, INSTR}}$ (1, 19) = 29.2, $p < .01$, $\eta^2 = .62$; F_{EVENT} (1, 17) = 21.5, $p < .01$, $\eta^2 = .57$; F_{RUN} (1, 19) = 31.6, $p < .01$, $\eta^2 = .64$. Accordingly, the effect for the repeated measurement factor "frequency" is also significant F (1.5, 106.7) = 86.1, $p < .01$, $\eta^2 = .54$. The 3x2x2 analysis of variance also reveals a significant interaction effect between the repeated measure factor and the type of motor task: F (1.5, 106.7) = 7.6, $p < .01$, partial $\eta^2 = .10$. This interaction effect indicates that differences between estimates are more pronounced if individuals judged high-frequency events after having placed their focus of attention on the frequency dimension during encoding (event task). No other effect reached an acceptable level of significance (all $F < 2.3$, all $p > .13$). The present findings replicate results from Experiment 1 and show that individuals are highly sensitive to the frequency of events in their memory-based judgments.

3.2.2 Impact of duration on JOF (duration bias)

Inspection of the means in the three time conditions in the upper part of Table 2 reveals that JOF are quite robust against influences of duration ($M_{8\text{ SEC}} = 5.32$, $M_{16\text{ SEC}} = 5.27$, $M_{24\text{ SEC}} = 5.76$). Accordingly, three of four linear trend analyses produce insignificant results: $F_{\text{EVENT, INSTR}}$, F_{EVENT}, $F_{\text{RUN}} < 1.5$; all $p > .25$. The linear trend for the explicit time-focus condition is marginally significant: $F_{\text{RUN, INSTR}}$ (1, 19) = 3.5, $p < .10$, $\eta^2 = .06$. The effect for the repeated measure factor "duration" is significant: F (1.7, 123.4) = 3.7, $p < .05$, $\eta^2 = .05$. The 3x2x2 analysis of variance yields no other significant effect (all $F < 1$).

3.2.3 Memory-based sensitivity for duration

Mean estimates of duration are depicted in the right part of Table 2. When participants are led to focus on the time dimension, JOD monotonically increase with increasing actual duration. This observation is corroborated by linear trend analyses: $F_{\text{RUN, INSTR}}$ (1,19) = 9.8, $p < .01$, $\eta^2 = .35$; F_{RUN} (1, 19) = 5.2, $p < .05$, $\eta^2 = .22$. Effects for other trends do not reach an acceptable level of significance $F_{\text{EVENT, INSTR}}$ (1, 18) = 0,43, $p = .52$, $\eta^2 = .03$; F_{EVENT} (1, 17) = 2,06, $p = .17$, $\eta^2 = .13$. The main effect for the repeated measure factor "duration" is reliable: F (1.5, 108.7) = 6.7, $p < .01$, partial $\eta^2 = .08$. The 3x2x2 analysis of variance yields no other significant effect and, in particular, no significant interaction effect between motor task and focus instruction (all $F < 2.2$, all $p > .15$). The present study shows that individuals are able to discriminate between differences in duration in their judgments if they are required to focus attention on the stimulus for the entire span of occurrence, either by performing a corresponding motor task only or additionally receiving an instruction to focus on time.

3.2.4 Impact of frequency on JOD (frequency bias)

Inspection of mean duration estimates in the right part of Table 2 reveals that JOD increase with increasing presentation frequency ($M_{2\text{ TIMES}} = 9.79$, $M_{4\text{ TIMES}} = 11.60$, $M_{8\text{ TIMES}} = 14.20$). Likewise, the four linear trends for the factor "frequency" are significant: $F_{\text{EVENT, INSTR}}$ (1, 18) = 40.4, $p < .001$, $\eta^2 = .70$; $F_{\text{RUN, INSTR}}$ (1, 19) = 3.1, $p < .10$, $\eta^2 = .15$; F_{EVENT} (1, 17) = 21.0, $p < .01$, $\eta^2 = .57$; F_{RUN} (1, 19) = 16.2, $p < .01$, $\eta^2 = .47$. Note, however, that the measure of effect size is quite small in the condition in which participants are explicitly instructed to focus on the time dimension (run-stop task with focus instruction). Over all conditions, the effect for the repeated measure factor "frequency" is significant and strong in effect size F (1.8, 129.6) = 42.9, $p < .01$, $\eta^2 = .40$. The effect is qualified by a two-way interaction effect between the motor task and the repeated measure factor: F (1.8, 129.6) = 4.7, $p < .01$, $\eta^2 = .06$. No other effects are reliable (all $F < 1$). The interaction reflects the result that the frequency bias on JOD decreases when attention is explicitly focused on the time dimension. This attenuation effect notwithstanding, JOD are still systematically biased by presentation frequency. In three out of four conditions, these biases are strong in effect size.

3.3 Discussion

The purpose of Experiment 2 was to explore the effects of different encoding conditions on the patterns of JOF and JOD. Recall that findings of the first experiment clearly violated the predictions of the common-path hypothesis. However, outright rejection of the hypothesis seemed to be a premature conclusion. Following Hintzman's (1970) reasoning, we speculated that participants might have failed to encode differences in duration simply because they did not keep their attention on the stimulus in proportion to its actual duration. If, as Hintzman suggested, they attended to each stimulus only for a constant period of time (e.g., as long as necessary to read), actual differences in duration should not change memory and, subsequently, should have null effects on judgment pattern.

Our second study demonstrated that the manner of encoding does indeed matter. Participants became sensitive in their JOD if they had been encouraged to keep their attention on a stimulus at the time of encoding. Note that merely performing a run-stop motor task was sufficient to cause a reliable effect of the factor "duration" on JOD. This effect increased in size if the motor task was assisted by an explicit focus instruction. In the latter condition, we also found a tendency towards duration biases on JOF, although the effect was comparatively weak in size. These findings clearly deviate from those obtained in the first experiment. The prior pattern of results could be replicated in the frequency-focus conditions that served as controls. In the event task conditions (with or without instruction to focus on frequency), we found only retrospective sensitivity in JOF and a unidirectional bias from presentation frequency to JOD. Judgment-sensitivity in JOF was generally high, irrespective of the motor-task and instruction conditions to which the participants were assigned.

Although encoding condition factors had an undeniable effect on the judgment pattern, the present findings do not convincingly corroborate the common-path hypothesis. First, evidence for a duration bias on JOF was only weakly reliable and found in only one condition (run-stop task with explicit focus instruction). Second, the judgment pattern in the time-focus conditions is still strongly asymmetric with regard to effect sizes. Third, the common-path hypothesis cannot account for the finding that frequency biases on duration decrease in the run-stop motor task condition (see the last result section: two-way interaction effect between motor task and repeated measure factor) given the other finding that participants can accurately differentiate between frequencies in all conditions.

Two misgivings, however, might be raised against the second experiment, which may cast doubt on the conclusiveness of the findings. First, the motor task imposes a dual-task situation. Participants must simultaneously attend to stimuli and control their motor reactions, which may have resulted in a division of attention. Consequently, cognitive resources for attention may have been only partially devoted to the stimuli themselves. Second, sustaining attention to a stimulus beyond reading time might have added only a marginal change to memory; recall that the stimuli provide only minimal information. As stimuli we presented typical forenames without context. All were presented in the same size, color and font. Reading the word might have been sufficient to encode the entire set of features representing the stimulus. Thus, one might wonder if any additional change in memory could be achieved by sustaining attention to such an impoverished stimulus presentation. A richer presentation, however, would have given the focus manipulations a greater chance to develop their impact on memory. If stimuli would provide more information than the names, the number of features to be encoded should increase with the length of time attention is maintained.

In Experiment 3 we subjected the common-path hypothesis to another test. We varied the *codality* of the stimulus material to enhance the probability that maintaining encoding efforts could effectively result in memory changes. Codality refers to the coding system used to present information within a certain mode. For example, in the visual mode, information can be presented literally, by pictograms or pictures. In this study we compared two codality conditions, words with pictures of the same entity. Because pictures are richer in regard to information conveyed (more details, contextual information, etc.), variations in exposure duration should cause corresponding variations in the time devoted to process the stimuli. If frequency and duration had commensurate effects on memory (as predicted by the common-path hypothesis), we should at least obtain a *symmetrical judgment pattern* in those codality conditions in which stimuli are presented as pictures together with context information.

4 Experiment 3

The study employed the same within-design as Experiment 1. In addition, we manipulated the codality of the stimuli as a between-subjects factor (word naming object, picture of object, picture of object in context). Instead of forenames, we presented participants with a list of objects (furniture, tools, food, toys).[2]

In this study we also assessed free recall (either before or after judgment). As another change, we took two kinds of JOD. Besides asking for total duration, we also had participants judge the single (mean) exposure duration of the stimulus (in accordance with Hintzman, 1970).

4.1 Method

4.1.1 Participants and design

Ninety undergraduates (27 female; mean age: 22.9) from different majors at Saarland University[3] participated in the experiment lasting about 35 minutes. They were paid € 6 each. The design was a 3 (codality of stimulus: word naming object, picture of object, picture of object in context) x 2 (position of recall test: before or after judgment) x 3 (presentation frequency) x 3 (total duration) factorial design. The last two factors were manipulated within subjects. Participants were randomly assigned to the six between-subjects conditions.

4.1.2 Stimuli

Participants were presented with 27 objects (see Appendix) shown on a computer screen. In accordance with data from several pretests using a large sample from our pool of participants (Glauer et al., 2008), objects were selected so that they were comparable with regard to prevalence, readability and affective load. In order to vary codality, we presented words or color pictures. The pictures of the objects were taken with a digital camera in real-life contexts. The objects filled out between 75 and 85 % of the area and were always shown in the foreground. In the "context free" condition, context information was digitally removed. The size of the pictures was held constant. Each picture covered one quarter of the screen and was presented in the center. In the "context" condition, the background contained features of environments that typically host the objects in real life settings. For example, a wardrobe was shown in a living room with parts of a carpet and wallpaper. For each experimental session, the stimuli were randomly assigned to the 9 conditions resulting from the 3 (frequency) x 3 (duration) within-subjects design, with the restriction that each condition contained three objects.

4.1.3 Procedure

The procedure was essentially the same as in the previous experiments. We counterbalanced the order of judgments. Half of the participants made JOF in first position and JOD (total) in the second. For the other half, the order was reversed. Mean JOD were always assessed in the third position. Mean JOD pertains to the average time interval the single word or picture was shown on the screen. This was carefully explained to participants in a similar fashion as described by Hintzman (1970, Exp. 3). As in the previous experiments, single (mean) duration was

---

[2] Changing codality in the previous material would have required us to present participants with pictures of women and men. It is difficult if not impossible to prevent those pictures from evoking affective reactions — feelings of liking/disliking, attraction, sympathy, etc. — in the observer. It cannot be ruled out that these affective reactions, themselves, may affect memory processes. Thus, we used stimuli with a generally lower affective potential. Moreover, we carefully pre-tested these stimuli on affective dimensions.

[3] We are indebted to Ralf Rummer and his research group at Saarland University for collecting the data.

held constant in each condition of the (presentation frequency) x 3 (duration) within-subjects design. Consider, for example, the first cell of the design in which stimuli were presented twice for a total duration of eight seconds. Accordingly, a stimulus word assigned to this condition was always shown for 4 seconds when it appeared on the screen.

4.2 Results

4.2.1 Memory-based sensitivity for frequency

Mean JOF are presented in Table 3. Evidently, JOF systematically co-vary with actual presentation frequencies. The observation is corroborated by trend analyses. All linear trends for the factor "frequency" in the six between-subjects conditions are significant and strong in effect size (all F (1,13) > 38. 4, all $p < .01$, all $\eta^2 > .72$). Moreover, we performed a 3 (codality) x 2 (position of recall) x 3 (frequency) x 3 (duration) analysis of variance. All together, the analysis produces four significant effects (for all other effects: $F < 1.8$, $p > .10$). Firstly, in accordance with linear trend analyses, the effect for the repeated measurement factor "frequency" is also high in effect size: F (1.73, 145.1) = 422.45, $p < .01$, $\eta^2 = .83$. Secondly, we found a weak but significant codality-by-frequency interaction effect F (3,45, 145.1) = 3.2, $p < .05$, $\eta^2 = .07$, reflecting the observation that regression to the mean[4] in JOF is slightly less pronounced in the "picture with context" condition, compared to the condition in which participants saw a written presentation of the objects. Thirdly, we obtained a recall-position-by-frequency interaction effect that is also small in effect size: F (1.73, 145.1) = 4.1, $p < .02$, $\eta^2 = .05$. Again, this effect reflects a systematic variation in the range of frequency judgments. Specifically, the regression effect in frequency judgment is slightly less pronounced if the recall measure is taken after participants have made their judgments than in the other recall-position condition. We will not consider the two interaction effects further. The fourth effect, however, is of theoretical importance and will be presented in the next paragraph.

4.2.2 Impact of duration on JOF (duration bias)

Inspection of mean JOF in the three time conditions in Table 3 indicates that frequency estimates tend to be biased by duration. This observation is corroborated by a significant main effect for the repeated measure factor "duration" on frequency judgments F (1.98, 166.2) = 15.2, $p < .01$, $\eta^2 = .15$. Six linear trend analyses reveal that this effect is caused mainly by the codality condition in which participants were presented with pictures showing the object in context (PIC-CONTEXT). In the two PIC-CONTEXT conditions, both linear trends (recall before/after judgment) are significant and of moderate effect size: $F_{\text{PIC-CONTEXT, RECALL-BEFORE}}$ (1, 13) = 8.9, $p <$

Table 3: Results of Experiment 3: Judgments of Frequency (JOF). Standard deviations are in parentheses. Between-subjects factors: Recall (before / after judgment), Codality (PICtures of objects with CONTEXT / PICtures of objects without context / WORDS); n = 15 in each of the six between-subjects conditions.

Mean JOF			
	Presentation Frequency		
Total duration	2	4	8
8 sec			
Recall before judgment			
PIC-CONTEXT	3.29 (0.90)	4.04 (0.86)	5.82 (0.93)
PIC	3.38 (0.94)	4.04 (0.83)	5.87 (1.13)
WORDS	3.49 (0.72)	4.07 (0.61)	5.42 (1.23)
Mean	**3.39 (0.84)**	**4.05 (0.76)**	**5.70 (1.10)**
Recall after judgment			
PIC-CONTEXT	2.98 (0.77)	4.36 (0.96)	6.38 (0.87)
PIC	3.00 (0.63)	4.29 (0.77)	5.91 (0.91)
WORDS	3.02 (0.86)	4.22 (0.89)	5.33 (0.90)
Mean	**3.00 (0.74)**	**4.29 (0.86)**	**5.87 (0.98)**
Grand mean	**3.19 (0.81)**	**4.17 (0.82)**	**5.79 (1.04)**
16 sec			
Recall before judgment			
PIC-CONTEXT	3.62 (1.23)	4.62 (1.05)	5.76 (0.90)
PIC	3.11 (0.90)	4.38 (0.96)	5.80 (0.99)
WORDS	3.22 (0.98)	4.58 (1.57)	5.89 (1.28)
Mean	**3.32 (1.05)**	**4.53 (1.20)**	**5.81 (1.05)**
Recall after judgment			
PIC-CONTEXT	3.33 (1.21)	4.67 (0.91)	6.71 (1.04)
PIC	3.24 (0.76)	4.24 (1.01)	6.20 (1.23)
WORDS	3.51 (0.85)	4.31 (1.12)	5.71 (1.33)
Mean	**3.36 (0.94)**	**4.41 (1.01)**	**6.21 (1.25)**
Grand mean	**3.34 (0.99)**	**4.47 (1.10)**	**6.01 (1.16)**
24 sec			
Recall before judgment			
PIC-CONTEXT	3.73 (1.07)	4.80 (0.94)	6.53 (1.04)
PIC	3.29 (0.93)	4.44 (0.92)	5.95 (1.02)
WORDS	4.13 (1.36)	4.67 (1.23)	5.33 (1.05)
Mean	**3.72 (1.16)**	**4.64 (1.03)**	**5.94 (1.13)**
Recall after judgment			
PIC-CONTEXT	3.56 (1.12)	4.89 (1.35)	6.67 (1.11)
PIC	3.16 (0.55)	4.62 (1.07)	5.93 (1.07)
WORDS	3.27 (1.24)	4.47 (1.15)	5.96 (1.01)
Mean	**3.33 (1.01)**	**4.66 (1.18)**	**6.19 (1.09)**
Grand mean	**3.52 (1.10)**	**4.65 (1.10)**	**6.06 (1.11)**

---

[4] The regression effect pertains to the well-established finding that low frequencies are likely to be overestimated, whereas high frequencies tend to be underestimated with regard to actual presentation frequency (Sedlmeier, Betsch & Renkewitz, 2002).

.01, η^2 = .39; $F_{\text{PIC-CONTEXT, RECALL-AFTER}}$ (1, 13) = 6.2, p < .01, η^2 = .31. There is also a significant effect in one other codality condition: $F_{\text{WORDS, RECALL-AFTER}}$ (1, 13) = 3.6, p = .05, η^2 = .20. The three remaining linear trends produced no significant effects, all F < 1.3. In a nutshell, these findings show that JOF are most likely to be biased by duration if the stimuli are presented visually together with context information.

4.2.3 Memory-based sensitivity for total duration

In this and the following subsection we first report results on judgments of *total* duration (this measure accords with those in the previous experiments). The results on judgments of single (mean) duration show the same pattern. Thus, we will only briefly consider them at the end of the results section.

Mean JOD (total) are shown in Table 4. Participants were sensitive to differences in duration in their memory-based judgments. This observation is corroborated by a 3x2x3x3 ANOVA that produces a significant and moderately strong effect for the within-subjects factor "duration", F (1.9, 160.2) = 57.7, p < .01, η^2 = .41. Linear trend analyses in the six between-subjects conditions, however, show that sensitivity to duration is most pronounced in one codality condition. If the objects are depicted in context, participants are extremely successful at reproducing differences in duration in their judgments. Accordingly, in the "context" conditions, both linear trends (recall before vs. after judgment) are significant and of strong effect size: $F_{\text{PIC-CONTEXT, RECALL-BEFORE}}$ (1, 13) = 28.8, p < .01, η^2 = .67; $F_{\text{PIC-CONTEXT, RECALL-AFTER}}$ (1, 13) = 27.3, p < .01, η^2 = .66. The remaining four trend analyses produce effects that are smaller in effect size and partially not significant: $F_{\text{PIC, RECALL-BEFORE}}$ (1, 13) = 9.5, p < .01, η^2 = .40; $F_{\text{PIC, RECALL-AFTER}}$ (1, 13) = 2.7, ns, η^2 = .16; $F_{\text{WORD, RECALL-BEFORE}}$ (1, 13) = 2.5, p > .10, η^2 = .15; $F_{\text{WORD, RECALL-AFTER}}$ (1, 13) = 8.3, p < .01, η^2 = .37.

In line with the pattern of effects obtained from trend analyses, we also find a significant codality-by-duration interaction in the multivariate ANOVA: F (3.8, 160.2) = 4.8, p < .01, η^2 = .10. The interaction effect statistically substantiates the observation that sensitivity in duration judgments is most pronounced in the codality condition in which participants watched pictures in context. The ANOVA produced only one further effect, which is considered in the following paragraph (for all other effects: F < 2, p > .13).

Table 4: Results of Experiment 3: Judgments of total Duration (JOD). Standard deviations are in parentheses; duration judgments are in seconds. Between-subjects factors: Recall (before / after judgment), Codality (PICtures of objects with CONTEXT / PICtures of objects without context / WORDS); n = 15 in each between-subjects condition.

Mean JOD (total)			
	Presentation frequency		
Total duration	2	4	8
8 sec			
Recall before judgment			
PIC-CONTEXT	12.11 (1.94)	13.09 (2.36)	16.05 (2.71)
PIC	11.89 (2.91)	13.09 (2.67)	15.69 (3.44)
WORDS	12.69 (2.83)	14.58 (3.16)	17.11 (3.68)
Mean	**12.23 (2.56)**	**13.58 (2.78)**	**16.28 (3.28)**
Recall after judgment			
PIC-CONTEXT	10.44 (0.96)	13.09 (2.38)	16.46 (4.24)
PIC	11.80 (2.48)	13.22 (3.31)	17.51 (3.21)
WORDS	11.05 (2.24)	12.42 (2.03)	15.73 (3.10)
Mean	**11.10 (2.04)**	**12.91 (2.60)**	**16.57 (3.55)**
Grand mean	**11.66 (2.37)**	**13.25 (2.70)**	**16.43 (3.40)**
16 sec			
Recall before judgment			
PIC-CONTEXT	14.13 (2.54)	14.82 (2.31)	18.42 (2.86)
PIC	11.89 (2.68)	14.73 (2.95)	16.49 (2.67)
WORDS	13.13 (2.76)	15.18 (3.60)	16.80 (3.54)
Mean	**13.05 (2.76)**	**14.91 (2.94)**	**17.24 (3.10)**
Recall after judgment			
PIC-CONTEXT	11.98 (1.98)	14.42 (2.51)	18.75 (3.58)
PIC	12.96 (2.22)	14.47 (3.04)	16.80 (3.55)
WORDS	12.40 (3.13)	13.40 (2.88)	15.31 (2.92)
Mean	**12.44 (2.47)**	**14.10 (2.80)**	**16.96 (3.58)**
Grand mean	**12.75 (2.62)**	**14.50 (2.88)**	**17.10 (3.33)**
24 sec			
Recall before judgment			
PIC-CONTEXT	15.20 (2.29)	16.38 (3.03)	19.51 (3.59)
PIC	14.20 (3.80)	15.05 (2.58)	17.84 (2.26)
WORDS	14.51 (3.56)	16.56 (3.02)	16.93 (3.42)
Mean	**14.64 (3.23)**	**15.99 (2.90)**	**18.10 (3.26)**
Recall after judgment			
PIC-CONTEXT	14.33 (2.96)	16.42 (3.13)	19.49 (2.73)
PIC	13.98 (3.29)	15.09 (2.94)	17.42 (3.56)
WORDS	12.84 (3.31)	13.93 (2.89)	17.29 (2.99)
Mean	**13.72 (3.18)**	**15.15 (3.10)**	**18.07 (3.20)**
Grand mean	**14.18 (3.22)**	**15.57 (3.01)**	**18.08 (3.21)**

4.2.4 Impact of frequency on JOD (frequency bias)

As in the previous experiments, we observe a frequency bias in total duration judgments. The multivariate ANOVA produces a strong effect on the repeated measure factor "frequency": F (1.6, 133.9) = 140.5, p < .01, η^2 = .63. The six linear trends for this factor are significant in all six between-subjects conditions (all F (1, 13) > 18.6; all p < .01, all η^2 > .56).

4.2.5 Memory-based sensitivity for single duration

Mean JOD$_{\text{SINGLE}}$ are shown in Table 5. The pattern of results provides a similar picture to that for judgments of total duration (see above), the only difference being that effect sizes are slightly smaller for single duration judgments. Again, participants were able to reproduce differences in duration in their memory-based judgments. The observation is corroborated by a 3x2x3x3 ANOVA that produces a significant and moderately strong effect for the within-subjects factor "duration"; F (1.97, 165.9) = 49.9, $p < .01$, $\eta^2 = .37$. Linear trend analyses in the six between-subjects conditions again show that sensitivity to duration is most pronounced in one picture-codality-condition. If the stimulus objects are placed in context, participants are extremely good at reproducing differences in single durations in their judgments — both linear trends (recall before/after judgment) are significant and of strong effect size: $F_{\text{PIC-CONTEXT, RECALL-BEFORE}}$ (1, 13) = 28.7, $p < .01$, $\eta^2 = .67$; $F_{\text{PIC-CONTEXT, RECALL-AFTER}}$ (1, 13) = 17.9, $p < .01$, $\eta^2 = .56$. The remaining four trend analyses produce effects that are smaller in effect size: $F_{\text{PIC, RECALL-BEFORE}}$ (1, 13) = 6.4, $p < .01$, $\eta^2 = .32$; $F_{\text{PIC, RECALL-AFTER}}$ (1, 13) = 4.98, $p < .05$, $\eta^2 = .26$; $F_{\text{WORD, RECALL-BEFORE}}$ (1, 13) = 5.4, $p < .05$, $\eta^2 = .28$; $F_{\text{WORD, RECALL-AFTER}}$ (1, 13) = 2.4, $p > .10$, $\eta^2 = .14$. Note that the last trend fails to reach an acceptable level of significance. The last condition (WORD, RECALL AFTER) is the replication condition of the former experiments. As in the previous studies, in this condition participants' ability to remember differences in duration is weak. The ANOVA produced no further reliable effect (all $F < 2.5$, all $p > .09$, all $\eta^2 < .05$).

Table 5: Results of Experiment 3 - Judgments of Single Duration. Standard deviations are in parentheses; duration judgments are in seconds. Between-subjects factors: Recall (before / after judgment), Codality (PICtures of objects with CONTEXT / PICtures of objects without context / WORDS); n = 15 in each between-subjects condition.

JOF$_{\text{SINGLE}}$			
	Presentation frequency		
Total duration	2	4	8
8 sec			
Recall before judgment			
PIC-CONTEXT	4.07 (1.77)	3.69 (1.39)	3.78 (1.81)
PIC	4.67 (1.99)	4.42 (2.23)	4.84 (1.67)
WORDS	5.07 (1.62)	4.93 (1.38)	5.22 (2.19)
Mean	**4.60 (1.81)**	**4.35 (1.75)**	**4.62 (1.96)**
Recall after judgment			
PIC-CONTEXT	4.87 (1.43)	4.49 (1.01)	4.15 (2.38)
PIC	4.25 (1.77)	4.38 (1.82)	4.31 (1.79)
WORDS	5.18 (1.02)	5.60 (1.60)	5.71 (1.74)
Mean	**4.76 (1.46)**	**4.82 (1.58)**	**4.73 (2.07)**
Grand mean	**4.68 (1.64)**	**4.59 (1.68)**	**4.67 (2.01)**
16 sec			
Recall before judgment			
PIC-CONTEXT	5.20 (2.20)	4.76 (1.74)	5.36 (2.17)
PIC	5.53 (2.80)	5.20 (1.85)	5.42 (2.23)
WORDS	5.98 (2.74)	5.53 (1.90)	5.62 (1.80)
Mean	**5.15 (2.21)**	**5.03 (1.68)**	**5.39 (2.06)**
Recall after judgment			
PIC-CONTEXT	5.64 (1.57)	4.84 (1.18)	5.47 (2.18)
PIC	5.74 (1.71)	4.78 (1.98)	4.29 (2.02)
WORDS	5.33 (1.78)	5.80 (1.18)	6.07 (1.80)
Mean	**5.57 (1.66)**	**5.14 (1.56)**	**5.27 (2.10)**
Grand mean	**5.36 (1.95)**	**5.08 (1.61)**	**5.33 (2.07)**
24 sec			
Recall before judgment			
PIC-CONTEXT	5.80 (1.79)	5.89 (2.12)	5.56 (2.27)
PIC	5.98 (2.74)	5.53 (1.90)	5.62 (1.80)
WORDS	5.96 (2.17)	5.78 (2.05)	5.64 (1.53)
Mean	**5.91 (2.22)**	**5.73 (1.99)**	**5.61 (1.85)**
Recall after judgment			
PIC-CONTEXT	6.38 (1.60)	5.91 (1.19)	5.82 (2.32)
PIC	6.00 (1.93)	4.87 (1.16)	4.15 (1.55)
WORDS	6.47 (1.82)	5.80 (1.69)	6.11 (1.82)
Mean	**6.28 (1.76)**	**5.53 (1.42)**	**5.36 (2.07)**
Grand mean	**6.10 (2.00)**	**5.63 (1.72)**	**5.48 (1.96)**

4.2.6 Impact of actual frequencies on JOD$_{\text{SINGLE}}$ (frequency bias)

In our design, *total* duration and frequency were uncorrelated. To achieve this, *single* durations of stimuli were decreased with increasing presentation frequency. For example, consider the conditions for which *total* duration was held constant at the level of 8 seconds. If the stimulus was presented 2 times one single appearance of the stimulus lasted 4 seconds. If it was presented 4 times, however, its single duration was decreased to 2 seconds. Due to the negative correlation between *single* duration and presentation frequency, the *absence* of a frequency effect is an indicator of a frequency bias that compensates influences of duration. Accordingly, only one of six linear trend analyses yields a significant effect: If the recall measure is taken following judgments and the picture is presented without context, participants reliably reproduce decreasing durations across frequency conditions in their judgments: $F_{\text{PIC, RECALL-AFTER}}$ (1, 13) = 7.6, $p < .01$, $\eta^2 = .35$. In all other conditions, participants fail to do this (all $F < 1$ all $p > .30$, all $\eta^2 < .07$). The pattern of results suggests that judgments of single durations are likely to be biased by frequency (although we are aware that this interpretation exploits null effects).

4.2.7 Free recall

As expected, percentages of recalled stimuli tend to be higher in the picture-codality conditions. None of the differences between conditions, however, reached an acceptable level of significance. This may be due to a ceiling effect because participants were able to recall 21 of the 27 stimuli on average. We therefore refrain from reporting descriptive and inferential statistics and from discussing free recall data.

4.3 Discussion

The results replicate and extend previous findings. First, consider the replication conditions. If participants were presented with words, we observed an asymmetrical judgment pattern. Specifically, memory-based sensitivity was generally strong for frequency but weak or absent for duration. Moreover, judged word duration was strongly biased by frequency. The bias in the reverse direction was weak or absent.

In contrast, consider the conditions in which stimulus codality was changed: we observed a tendency towards a symmetric pattern in terms of statistical reliability and, occasionally, even in terms of effect size. Particularly in the picture-with-context conditions, individuals showed a sound capability to reproduce differences in duration in their judgments, regardless of whether participants were to judge total or single (mean) duration. The finding that the presence of contextual information plays a crucial role in encoding converges with results from research on recognition memory (e.g., Murnane, Phelps & Malmberg, 1999). The bias of duration on frequency judgment is also remarkably strong in these conditions. Moreover, frequency and duration showed similar effects on recall. The number of correctly recalled words increased with increasing presentation frequency and duration.

All together, the findings converge with those from Experiment 2, which showed that bivariate sensitivity and bidirectional biases occurred if attention during the encoding phase was sustained to the stimuli proportionate to their actual duration. Moreover, the present study showed that explicit focus instructions were not necessarily a condition for these effects to appear. Rather, the evidence suggests that differences in duration were unintentionally registered in the picture-with-context conditions. Therefore, it seems to be the case that stimuli containing more detailed information can encourage individuals to sustain attention and maintain encoding of the stimuli in proportion to the length of time they are encountered.

These results tend to be in line with the prediction of the common-path hypothesis. However, some pieces of evidence still challenge the model. The strength of effects still deviates between the two judgment dimensions. Again, consider the picture-with-context conditions in which we found the strongest tendency towards a symmetric judgment pattern. Neither effect sizes for judgment sensitivity (all $\eta^2_{\text{JOF}} > .72$, all $\eta^2_{\text{JOD}} > .64$) nor for biases (all $\eta^2_{\text{JOF}} < .40$, all $\eta^2_{\text{JOD}} > .55$) converge. These differences cannot be accounted for by a strict common-path hypothesis.

5 General discussion

5.1 Judgment pattern and the primacy of frequency storage

In three experiments we tested the common-path hypothesis for JOD and JOF. Accordingly, both experienced frequency and duration should change the strength by which the corresponding stimulus is represented in memory. At the time of judgment, the individual can rely on memory strength to make JOF and JOD, retrospectively. This being the case, experienced frequency and duration should have commensurate effects on judgments. Specifically, the judgment pattern is expected to be *symmetric*. Individuals should be able to reproduce differences in experienced duration and frequency in their judgments (bivariate sensitivity). Moreover, JOF should be susceptible to bias of experienced duration, while JOD should be susceptible to bias of experienced frequency (bidirectional bias).

The results, however, were mixed. Participants were generally retrospectively sensitive to differences in frequency, but were not always sensitive to differences in duration. JOF were often but not always robust against biases of duration. In contrast, JOD were consistently and strongly biased by frequency, indicating that participants were prone to rely on frequency memory when forming judgments of duration. In other words, there was a strong, yet inconsistent tendency towards an *asymmetric* judgment pattern. All in all, the results suggest that, with regard to storage in long-term memory, there is a *primacy* of frequency over duration.

5.2 Violation of the common-path hypothesis

The overall observed pattern of results is in opposition to the predictions of a strict common-path hypothesis. Note, however, that the common-path hypothesis only makes predictions concerning memory and judgment; it does not spell out the processes of encoding. Hence, it does not exclude the possibility that individuals may be differentially likely to detect the frequency-related and time-related aspects of stimuli. Effective storage of quantitative information requires processes of encoding to be synchronized with the appropriate stimulus dimension to produce proportionate changes in memory.

As Hintzman (1970) speculated, the time of encoding may be a constant under such conditions. Hence, differences in duration that elapse beyond that period may fail to be systematically transferred into long-term memory. We proposed that such differences in encoding may account for the primacy of frequency-related aspects of information in encoding and storage. Moreover, this interpretation is in line with the literature that deals with time perception. A number of researchers have shown memory for duration to be generally poor if individuals are not encouraged to attend to the temporal dimension during encoding (e.g., Hicks, 1992).

In fact, encoding conditions substantially affected the pattern of judgment. We obtained evidence for bivariate sensitivity and bidirectional biases when participants were encouraged to sustain attention to the stimuli in proportion to their presentation duration, either by a corresponding motor task (Exp. 2), an explicit instruction (Exp. 2) or because the stimuli were presented in a picture-with-context format (Exp. 3). Notwithstanding these results, the judgment pattern could not be considered truly symmetric. Effect sizes consistently indicate that sensitivity was more pronounced in JOF than JOD and that biases were stronger in JOD than JOF.

5.3 Violation of a dual-path hypothesis

Granting the violation of the strict common-path hypothesis, one may speculate whether results corroborate a *dual-path hypothesis,* which states that frequency and duration are stored independently in memory. Models of independent storage have been suggested, for example, in literature on learning. According to the mode control model (Meck & Church, 1983), frequency and time are processed in different modes yielding independent representations for frequency and duration (although a similar mechanism is used for counting and timing).[5] Note, however, that the mode control model is *not* assumed to account for long-term memory storage of time and frequency. Rather, it has been suggested as a model for prospective timing and counting in working memory (see Wearden, 2005, for a discussion). Nevertheless, a dual-path approach may also apply to long-term storage of frequency and duration.

Unfortunately, our findings clearly violate the dual-path hypothesis. Although a dual path approach could account for differences in sensitivity, it could *not* account for bi-directional biases. Accordingly, if distinct representations existed for frequency and temporal information, subsequent JOF and JOD could rely on their corresponding representation in memory. Due to such independent representations, judgments of one stimulus dimension should be *not* affected by the other.

All together, our results contradict both a strict common-path and a strict dual-path hypothesis.

5.4 A conditional-common-path hypothesis

Despite the failure of the strict common-path hypothesis, the tendency towards a symmetric judgment pattern gives rise to the speculation that, under certain conditions, both frequency and duration may change a common representation in memory. Obviously, the *tuning of attention* at encoding seems to be one such crucial condition. Minimal attention (e.g., the orientation reaction towards the stimulus) was a sufficient condition for effective storage of frequency in memory. Storage of duration, however, seems to also require that attention towards the stimulus is sustained proportionate to its actual duration. Accordingly, judgmental sensitivity for duration should be conditional on attention during encoding. Attentional processes are resource consuming and error prone. Due to a variety of factors (distraction, fatigue, etc), the individual will often fail to direct attention perfectly towards the stimulus for the entire duration. Therefore, there should be a primacy in encoding and storing differences in frequency compared to differences in duration. Given that attention is sustained proportionate to actual duration, frequency and duration may both affect memory in a common fashion (e.g., by changing the strength of R_S in memory). When making judgments, individuals may rely on such changes in memory to estimate frequency or duration. Our findings are largely in line with such a conditional-common-path (CCP) hypothesis.

In Experiment 1, neither the features of the task nor the stimuli themselves encouraged participants to sustain attention proportionate to stimulus duration. JOD failed to reveal any retrospective sensitivity for duration but reflected differences in experienced frequency. On the contrary, JOF were rather accurate and unbiased by experienced duration. According to CCP hypothesis, any factor promoting the individual to attend to stimuli proportionate to their actual duration should cause a tendency to-

---

[5] The mode control model is a variant of scalar timing theories (Gibbon, 1977; see Church, 2003, for a brief introduction). It postulates that the brain can discriminate frequency and duration independently but uses a common accumulation mechanism. The input to the process stems from inside the organism and not directly from the stimulus. A pacemaker (e.g., an internal clock) emits pulses that are summed by an accumulator mechanism. A switch controls the accumulation process. If the switch operates in the *run-and-stop mode,* it serves as a timer. Pulses are gated to the accumulator as long as an event occurs. If the switch operates in the *event mode*, it serves as a counter. Pulses are gated to the accumulator for a fixed interval as often as an event occurs. The same accumulator mechanism can produce measures for both duration and frequency that are independently represented and experienced as mental magnitudes (for the concept of mental magnitudes, see Whalen, Gallistel & Gelman, 1999).

wards a symmetric judgment pattern. In line with this assumption, we found evidence for bivariate sensitivity and bidirectional biases regardless of whether the task (focus conditions in Exp. 2) or the stimuli (e.g., picture in context condition; Exp. 3) encouraged participants to sustain attention. Even under these conditions, however, we found a stronger sensitivity for JOF compared to JOD as well as a stronger frequency bias. These differences in effect sizes should be expected if one assumes that sustaining attention is a process prone to error. Merely detecting the occurrence of an event requires less attentional resources and is less susceptible to the effects of distracting variables than maintaining attention proportional to the entire duration of stimulus presentation. In other words, it is easier to encode differences in frequency than differences in duration.

The important role of attention has been widely recognized in research on timing. Particularly, in prospective timing task, attention is proposed to be a necessary condition for forming valid measures of time in working memory (see the attentional gate model, Block & Zakay, 1996). The CCP hypothesis assumes that attention is also an important condition in retrospective tasks when judgments capitalize on long-term memory. The effects of attention on long-term memory can be modeled within existing theories, for example within Hintzman's Minerva model (Hintzman, 1988). Minerva proposes that each encounter of the stimulus results in trace building. Other things being equal, frequency of occurrence will correspond to the number of traces that, all together, represent the stimulus in memory. Assume now that sustaining attention resulted in forming multiple traces for the same stimulus within the same encoding episode (e.g., via repetitively reading the stimulus word). As a consequence, both duration and frequency would change the number of traces that represent the stimulus in memory. If JOF and JOD relied on echo strength (which in turn is assumed to be a function of trace number), we should expect bivariate sensitivity and bidirectional biases in judgments under the above described conditions. However, we should not expect perfect symmetry, since event occurrence and event duration are differentially effective in adding traces to memory.

5.5 Implications for research on judgmental heuristics

Our results are in line with a heuristic approach assuming that individuals use memory strength as a predictor for quantitative judgments. Prominent examples are the availability heuristic (Tversky & Kahneman, 1973) and the fluency heuristic (e.g., Jacoby & Brooks, 1984; Schooler & Hertwig, 2005). Doughertey and colleagues (Dougherty & Franco-Watkins, 2002; Dougherty, Gettys & Ogden, 1999) showed how these and other heuristics can be tied together in an overarching theory of memory. Their MINERVA-DM model advances Hintzman's multiple trace theory of memory towards an integrative account for several phenomena in judgment and decision making. With the auxiliary assumption that sustaining attention results in multiple trace building for the same stimulus, our results are also in line with MINERVA-DM.

In JDM research, biases are often used as indicators for heuristics (Kahneman, Slovic & Tversky, 1982). From a memory perspective, however, one may doubt the indicative power of biases on the judgment level. Our results demonstrate that *variations in encoding* conditions change judgmental sensitivity and biases. Consequently, variations in judgmental patterns may indicate *changes in memory* rather than changes in judgmental processes.

Fiedler reported evidence from empirical studies and computer simulation convincingly showing that many biases in judgment can be explained with reference to basic pre-judgmental processes such as sampling, storing and retrieval (1996, 2000, 2002).

In all three of the studies reported above, duration judgments reflected experienced frequency. Presumably, participants relied on a common memory structure for different kinds of judgments (JOD and JOF). Together with the (partial) evidence for symmetry in judgment pattern, one may suspect that individuals employed a similar or even the same judgment strategy when making JOF and JOD (i.e., inferring frequency and duration from memory strength). It is up to future research to pursue this issue further or to identify conditions under which individuals use different strategies for JOF and JOD. Recent contributions to the literature on decision strategies cast some doubt on the view that individuals employ different rules for judgment and decision making, as the multiple strategy approach suggests (e.g., Glöckner & Betsch, 2008a,b; Bergert & Nosofsky, 2007; Lee & Cummins, 2004). These researchers do not deny that individuals employ different information search methods in memory and the environment. They suggest, however, that an all-purpose rule is applied to make a decision, regardless of how the input information has been gathered or formed. It has been shown that variations in the input to the decision processes can account for variations in decision outcome without making it necessary to evoke different heuristics. Regardless of whether one resides with a single or a multiple strategy view, it should have become clear that encoding and storage processes are crucial. Thus, "let us not forget memory" (Weber et al., 1995) when studying judgment and decision processes.

References

Attneave, F. (1953). Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology,46*, 81–86.

Betsch, T., Siebler, F., Marz, P., Hormuth, S., & Dickenberger, D. (1999). The moderating role of category salience and category focus in judgments of set size and frequency of occurrence. *Personality and Social Psychology Bulletin, 25*, 463–481.

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 107–129.

Block, R. A. (1990). Models of psychological time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 1–35). Hillsdale, NJ: Erlbaum.

Block, R. A., & Zakay, D. (1996). Models of Psychological Time Revisited. In: H. Helfrich (Ed.), *Time and Mind*. (pp. 171–195). Toronto: Hogrefe Huber.

Buhusi, C. V., & Meck, W.H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews, 6*, 755–765.

Brown, N. R. (2002). Encoding, representing, and estimating event frequencies: A multiple strategy perspective. In P. Sedlmeier & T. Betsch (2002). *Etc.: Frequency processing and cognition* (pp.37–54). Oxford: Oxford University Press.

Church, R. M. (2003). A concise introduction to scalar timing theory. In In W. H. Meck (ed.), *Functional and neural mechanisms of interval timing* (pp. 3–22). Boca Raton: CRC Press.

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 4*, 221–234.

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta Psychologica, 121*, 109–124.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. (1999). Minerva-DM: A memory model for judgments of likelihood. *Psychological Review, 106*, 180–209.

Dougherty, M. R. P., & Franco-Watkins, A. M. (2002). A memory-models approach to frequency and probability judgment: Applications of Minerva 2 and Minerva DM. In P. Sedlmeier & T. Betsch (2002). *Etc.: Frequency processing and cognition* (pp.121–136). Oxford: Oxford University Press.

Droit-Volet, S., & Wearden, J. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology, 80*, 142–159.

Fiedler, K. (2002). Frequency judgments and retrieval structures: Splitting, zooming and merging the units of the empirical world. In P.Sedlmeier & T.Betsch (eds.). *Etc.: Frequency processing and cognition* (pp. 67–87). Oxford: Oxford University Press.

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107*, 659–676.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review, 103*, 193–214.

Fischbach, J. (2004). *The effects of attention on judgments of frequency and duration*. University of Erfurt: Unpublished Master Thesis.

Gallistel, C. R. (1989). Animal cognition; Representation of space, time and number. Annual *Review of Psychology, 40*, 155–189.

Gibbon, J. (1977). Scalar expectancy theory and Weber´s law in animal timing. *Psychological Review, 84*, 279–325.

Glauer, M., Häusig, S., Krüger, M., Betsch, T., Renkewitz, F., Sedlmeier, P., & Winkler, I. (2007). Typizitätsnormen für Vertreter von 30 Kategorien. *Neurolinguistik,2*, 5–46.

Glöckner, A., & Betsch, T. (2008a). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making, 3*, 215–228.

Glöckner, A., & Betsch, T. (2008b). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1055–1075.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist, 12*, 1372–1388.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267–272.

Hicks, R. E. (1992). Prospective and retrospective judgments of time: A neurobehavioral analysis. In F.Macar,V. Pouthas, and W. J. Friedman (Eds.), *Time, action, and cognition: Towards bridging the gap*. (pp. 97–108). Dordrecht: Kluwer Academic Press.

Hicks, R. E.,Miller, G.W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of the amount of information processed. *American Journal of Psychology, 89*, 719–730.

Hintzman, D. L. (1970). Effects of repetitions and exposure duration on memory. *Journal of Experimental Psychology*, *83*, 435–444.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple trace model. *Psychological Review, 95*, 528–551.

Hintzman, D. L. (2000). Memory judgments. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 165–177). New York: Oxford University Press.

Hintzman, D. L. (2004a). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory and Cognition, 32*, 336–350.

Hintzman, D. L. (2004b). Time versus items in judgment of recency. *Memory and Cognition, 32*, 1298–1304.

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975) What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures. *Memory and Cognition, 3*, 287–294.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception and concept learning. In G. H. Bower (Ed.), *Psychology of learning and motivation* (*Vol. 18*, pp. 1–47). New York: Academic Press.

Johnson, M. K., Peterson, M. A., Yap, E. C., & Rose, P. M. (1989). Frequency judgments: The problem of defining a perceptual event. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 230–240.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. *Psychonomic Bulletin & Review, 11*, 343–352.

Lewandowsky, S., & Smith, P. W. (1983). The effect of increasing the memorability of category instances on estimates of category size. *Memory & Cognition, 11*, 347–350.

Lewis, P. A., & Miall, R. C. (2006). Remembering the time: a continuous clock. *Trends in Cognitive Sciences, 10*, 401–406.

Macar, F. (2005). Time passing, attention, and internal timers. In A.-N. Perret-Clermont, (ed.), *Thinking time — a multidisciplinary perspective* (pp. 15–21). Cambridge, MA: Hogrefe & Huber Publishers.

Manis, M. Shedler, J. Jonides, J., & Nelson, T. E. (1993). Availability heuristic in judgments of set-size and frequency of occurrence. *Journal of Personality and Social Psychology*, *65*, 448–457.

Meck, W. H. (2003) (Ed). *Functional and neural mechanisms of interval timing*. Boca Raton: CRC Press.

Meck, W. H., & Church, R.M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320-334.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-Dependent Recognition Memory: The ICE Theory. *Journal of Experimental Psychology: General, 128*, 403–415.

Ornstein, R. E. (1969). *On the experience of time*. Harmondsworth: Penguin.

Rammsayer, T. (2003). Sensory and cognitive mechanisms in temporal processing elucidated by a model system approach. In H. Helfrich (Ed.), *Time and mind II: Information processing perspectives.* (pp. 97–113). Ashland, OH: Hogrefe & Huber Publishers.

Rudolph, U. (2001). Die Wahrnehmung von Altersstereotypen, Attraktivität und Intelligenz für Vornamen im Deutschen. In: J. Eichhoff (ed.)*, Name und Gesellschaft — Soziale und historische Aspekte der Namensgebung und Namensentwicklung.* Heidelberg: Brockhaus.

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112*, 610–628.

Schwartz, B., & Reisberg, D. (1991). *Learning and Memory*. New York: Norton.

Sedlmeier, P. (1999). Improving statistical reasoning. Theoretical models and practical implications. Mahwah: Lawrence Erlbaum.

Sedlmeier, P. (2002). Associative learning and frequency judgement. In P. Sedlmeier, & T. Betsch (Eds.), *Etc.: Frequency processing and cognition*. Oxford: University Press.

Sedlmeier, P. (2008). *PASS-T: An associative learning model that simulates judgments of frequency and duration* . Talk at the XXIX International Congress of Psychology, Berlin.

Sedlmeier, P, Betsch, T., & Renkewitz, F. (2002). Frequency processing and cognition: Introduction and overview. In P.Sedlmeier & T.Betsch (eds.). *Etc.: Frequency processing and cognition* (pp. 1–17). Oxford: Oxford University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483–488.

Wearden, J.H. (2005). The wrong tree: Time perception and time experience in the elderly. In J. Duncan, L. Phillips, & P. McLeod (eds.), *Measuring the mind: Speed, age, and control* (pp. 137–158). Oxford: Oxford University Press.

Weber, E.U., Goldstein, W. M., & Barlas, S. (1995). And let us not forget memory: The role of memory processes and techniques in the study of judgement and choice. In J. R. Busemeyer, R. Hastie, and D. L. Medin (Eds.), *Decision making from a cognitive perspective: The psychology of learning and motivation* (pp. 33–82). New York: Academic Press.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130–137.

Williams, K.W., & Durso, F. T. (1986). Judging category frequency: automaticity or availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 387–396.

Zacks, R., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.). *Etc.: Frequency processing and cognition (pp. 21–36).* Oxford: Oxford University Press.

Zakay, D. (2005). Attention and duration judgment. *Psychologie Bulletin Francaise, 50*, 65–79.

Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. In M.A. Pastor & J.Artieda (eds.), *Time, internal clocks and movement* (pp 143–164). Amsterdam: North-Holland, Elsevier.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science, 6,* 12–16.

Zacks, R., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.). *Etc.: Frequency processing and cognition (pp. 21–36).* Oxford: Oxford University Press.

Appendix

Stimuli used in Experiment 1

	Male Forenames	Female Forenames
1.	Andreas	Anna
2.	Christoph	Claudia
3.	Daniel	Ella
4.	Fabian	Karin
5.	Felix	Katja
6.	Florian	Klara
7.	Julian	Laura
8.	Lukas	Lea
9.	Marcel	Lena
10.	Martin	Lisa
11.	Matthias	Lydia
12.	Moritz	Maria
13.	Peter	Nadin
14.	Philipp	Nina
15.	Robert	Sandra
16.	Stefan	Sarah
17.	Thomas	Sonja
18.	Tobias	Sophie

Stimuli used in Experiment 3

1. Aubergine [eggplant]
2. Sessel [armchair]
3. Trompete [trumpet]
4. Puppe [doll]
5. Zange [pliers]
6. Locher [hole punch]
7. Gurke [cucumber]
8. Glühbirne [light bulb]
9. Pullover [sweater]
10. Apfel [apple]
11. Couchtisch [table]
12. Cello [cello]
13. Lampe [lamp]
14. Hammer [hammer]
15. Tacker [stapler]
16. Möhre [carrot]
17. Kerze [candle]
18. Mantel [coat]
19. Regal [shelf]
20. Kleiderschrank [wardrobe]
21. Klavier [piano]
22. Schaukelpferd [rocking-horse]
23. Hobel [plane]
24. Lineal [ruler]
25. Tomate [tomato]
26. Spiegel [mirror]
27. Hose [trousers]