



Dynamics of retrospective timing: A big data approach

Fuat Balci¹ · Hüseyin Ünübol² · Simon Grondin³ · Gökben Hızlı Sayar² · Virginie van Wassenhove⁴ · Marc Wittmann⁵

Accepted: 17 March 2023 / Published online: 3 April 2023
© The Psychonomic Society, Inc. 2023

Abstract

Most interval timing research has focused on prospective timing tasks, in which participants are explicitly asked to pay attention to time as they are tested over multiple trials. Our current understanding of interval timing primarily relies on prospective timing. However, most real-life temporal judgments are made without knowing beforehand that the durations of events will need to be estimated (i.e., retrospective timing). The current study investigated the retrospective timing performance of ~24,500 participants with a wide range of intervals (5–90 min). Participants were asked to judge how long it took them to complete a set of questionnaires that were filled out at the participants' own pace. Participants overestimated and underestimated durations shorter and longer than 15 min, respectively. They were most accurate at estimating 15-min long events. The between-subject variability in duration estimates decreased exponentially as a function of time, reaching the lower asymptote after 30 min. Finally, a considerable proportion of participants exhibited whole number bias by rounding their duration estimates to the multiples of 5 min. Our results provide evidence for systematic biases in retrospective temporal judgments, and show that variability in retrospective timing is relatively higher for shorter durations (e.g., < 30 min). The primary findings gathered from our dataset were replicated based on the secondary analyses of another dataset (*Blursday*). The current study constitutes the most comprehensive study of retrospective timing regarding the range of durations and sample size tested.

Keywords Big data · Psychophysics · Retrospective timing · Time perception

Introduction

The ability of organisms to keep track of seconds-to-minutes long durations is called *interval timing*. Many interval timing studies in the range of up to several seconds have focused on prospective timing by asking participants to repeatedly judge time intervals. Thus, in these tasks participants attend

to elapsed time and attempt to use chronometric methods (e.g., tapping, counting) to improve their performance (Block et al., 2018; Grondin et al., 2004; Grondin & Killen, 2009; Rattat & Droit-Volet, 2012). The results gathered from prospective studies have established that humans are on average accurate in their time estimates and that the variability in time estimates is overall proportional to the mean estimates (scalar property, e.g., Grondin et al., 1999; Wearden & Lejeune, 2008). But the above-mentioned features of prospective timing reduce its ecological validity regarding human temporal information processing in daily life (Brown & Stubbs, 1988). For instance, prospective timing cannot capture our ability to confidently estimate the duration of events without knowing in advance that we would be asked to make that judgment. Such instances of temporal information processing are referred to as retrospective timing that can be interrogated only once and after the to-be-timed event is experienced (Hicks, 1992). Arguably, this is the main reason retrospective timing has not been widely researched. However, we seldom experience events or stimuli with the instruction to judge duration. Instead, events with longer

✉ Fuat Balci
fuat.balci@umanitoba.ca; fuatbalci@gmail.com

¹ Department of Biological Sciences, University of Manitoba, 50 Sifton Road, BSB 222, Winnipeg, MB R3T 2M5, Canada

² Department of Psychiatry, Üsküdar University, Istanbul, Türkiye

³ École de Psychologie, Université Laval, Québec City, Québec, Canada

⁴ Cognitive Neuroimaging Unit, INSERM, CEA, CNRS, Université Paris-Saclay, NeuroSpin, 91191 Gif/Yvette, France

⁵ Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany

duration are often judged in retrospect when they have already passed (“this was a long wait”) (Jokic et al., 2018).

The distinction between retrospective and prospective timing was made as early as the 19th century by William James (1890). The former is argued to rely on memory processes (Block, 1985; Block & Zakay, 1997), while the latter requires the involvement of attention (Balci et al., 2009). We investigated the retrospective judgment of the duration of a wide range of intervals (5–90 min) in over 24,000 participants. We tested the accuracy of retrospective time estimates, whether the variability in retrospective time estimates exhibits scalar property and the duration that can be most accurately judged retrospectively.

Retrospective time perception has been investigated in a handful of studies with a relatively narrow range of elapsed durations (around a few minutes) concerning the experience of waiting (e.g., Jokic et al., 2018; Witowska et al., 2020). Studies that directly compared the prospective and retrospective timing performances showed that prospective time estimates are more accurate and precise (e.g., Block & Zakay, 1997; Tobin & Grondin, 2009), reinforcing the claim that these two types of timing are mediated by different processes. Other studies investigated how the accuracy and variability of retrospective estimates change with elapsed time using multiple activity tasks. This variant of retrospective timing allows one to test the retrospective time judgments multiple times in individual participants. To this end, Brown and Stubbs (1988) asked participants to retrospectively judge the duration of musical excerpts from a set (ranging from ~1.5 to 8 min), and found that the exponent of the power function ($k \cdot t^a$) that defined the subjective time ranged between .32 and .38. Using a very similar approach, Grondin and Plourde (2007) asked participants to retrospectively estimate the duration of a series of cognitive tasks using a within-subject task design (2–8 min). They found that the timing imprecision (i.e., the coefficient of variation - CV) decreased from .39 to .29 over the test durations, and the exponent of the power function was .47 for event segments (.79 when the total duration was estimated). These findings point to the overestimation of shorter durations and underestimation of longer intervals and lower temporal sensitivity to shorter intervals. Grondin and Laflamme (2015) showed similar patterns for time reproduction but not estimation (for much briefer intervals of 2–16 s).

Earlier studies of retrospective timing are limited in terms of the range of the target durations tested and the limited sample size (but see Chaumon et al., 2022). Thus, a full-scale characterization of retrospective timing is surely needed. The current study aims to draw a clearer picture of the nature of retrospective time judgments and how timing variability changes as a function of objective time by testing a very large number of participants with a very wide range of

intervals.¹ We also aimed to identify the duration that can be most accurately estimated by the participants. Earlier findings (e.g., Chaumon et al., 2022) suggested that participants would overestimate short intervals and underestimate long intervals, and highest accuracy and precision would be for a 15- to 20-min long interval.

Finally, earlier research has shown a systematic whole number bias both in real-life tasks (i.e., stock trades) and experimental settings (e.g., Converse & Dennis, 2018). For instance, the number of trades in the stock exchange showed systematic blips at 100 and 1,000s (particularly 1,000, 2,000, 5,000) compared to other trade sizes. Similar biases were observed experimentally under time pressure and high cognitive load. These biases are consistent with the predictions of the “prominent numbers” approach of Albers and Albers (1983), which suggests that multiples, powers, doubles of 10, and their halves are more prominent numerals based on their mathematical properties. We tested whether similar biases apply to retrospective time estimates.

Methods

Participants

In this study, 24,494 participants (49.8% females) across Türkiye were tested in person. The participants were selected from various groups, including schools, municipal buildings, private companies, and public places such as neighborhood units, courses, and charities. The study used a stratified cluster-sampling approach based on the NUTS classification (Nomenclature of Territorial Units for Statistics) to select participants. The study included people residing in 26 NUTS regions of Turkey, with at least 200 and at most 2,000 individuals in each region. Larger samples were selected from regions with higher population densities. Using a stratified cluster-sampling approach based on the NUTS classification ensures that the study is representative of the population of Turkey. A total of 24,990 people were interviewed for the study; however, 24,494 individuals met the criteria and completed the scales. All participants were at least 18 years old. The average age was 31.8 (SD 16.8)

¹ Note that in its strictest form, time-scale invariance (a theoretically critical form of modulation of variability) is a within-subject phenomenon. Our aim to test scalar property in data gathered from a between-subject design was motivated by the fact that when the variability of behavior is modulated in a similar way for each participant (but with variability in parameters between participants), the average of data points composed of single data points from each subject for one target would approximate the patterns in individual participants. In support of our assumption, recently we observed that timescale invariance emerges even when data are pooled between subjects, experiments, and studies (Aydoğan et al., 2023).

years. We used liberal inclusion criteria, which were to be able to answer the questions on the surveys and not having a mental illness that would prevent them from completing the questionnaires. There was no compensation for taking part in the experiment.

Procedure

This study was approved by the Institutional Review Board of Üsküdar University B.08.6.YÖK.2.ÜS.0.05.06 / 2018 / 800. Data were collected in 2018 (July–October) by 125 clinical psychologists as part of a large field study. All researchers received 8 h of training to ensure consistency in administering interviews. The data collection was coordinated by nine sub-region coordinators, four regional coordinators, and two faculty members. All participants provided written informed consent before data collection. After the consent form was signed, participants were given a booklet that contained the following surveys: Sociodemographic information form, Personal Well-Being Index-Adult, Brief Symptom Inventory, Positive and Negative Affect Survey, Toronto Alexithymia Scale, Experiences in Close Relationships Inventory, Behavioral Addiction Risk Questionnaire. The researchers provided verbal and written instructions to the participants, and any questions asked by the participants were answered. In exceptional cases, the researcher read the questions to the participants and filled out the questionnaires based on the responses verbally provided by the participants.

The booklet contained 208 questions (including the time judgment), which were completed on average in 28 (SD 24) min. There was no time limit for filling out the surveys. The data were quality checked by the coordinators during the data collection period. Participants were asked to record the clock time at the designated spot on the first page of the booklet (__:__). After completing all surveys, participants recorded their responses to the following question: “How many minutes did it take to complete the questionnaires? (e.g., 67 min have elapsed, 7 min have elapsed, etc.)”. On the last page of the booklet, participants recorded the clock time again (__:__). The objective duration was calculated as the difference between the two recorded clock times. Participants did not know that they would be asked to estimate the duration of filling out the questionnaires.

Data analyses

Participants' data were excluded from the analyses when they met one of the following conditions: estimates that were shorter than $.1 \times$ elapsed time (.1% of the cases) and longer than $10 \times$ elapsed time (1.1%), missing timing data (.3%), elapsed times of filling out the questionnaires (too slow) > 90 min (.5%) and < 8 min (too fast) (2.6%), and an estimate > 180 min (.02%). The exclusion criteria were determined before data analyses. As a result, $\sim 4.5\%$ of the data were excluded from the analysis. The lower limit of the elapsed time

(8 min) was determined as the first time point that is higher than the quarter of the average duration to finish the surveys. We used linear and power functions to characterize the psychophysical functions for non-transformed data. We particularly chose the power function since, in the perception domain, a classical way of analyzing the psychological reality to physical magnitude (or chronometric units in the case of duration) is to use a log-log function (Stevens' power law; Stevens, 1957), with the exponent being the signature of the sensory modality one is working with (e.g., a slope much larger than 1 (exponential function) when working with electric shock sensation). This function can capture linear and non-linear patterns (positively and negatively accelerating), which can capture biases in time estimates.

Secondary analysis of Blursday dataset – control session To test whether the same core patterns held in a previous study, we conducted secondary analyses of the dataset that we had recently collected as part of an international consortium (Chaumon et al., 2022). The analyses were repeated for the data that were collected from the control group in that study (data after the lockdown due to the COVID-19 pandemic). The same exclusion rules were applied. Data over 90 mins for the elapsed time were excluded from this dataset, as was done in our analysis of the data collected in the current study.

Results

Figure 1 (bottom left panel) shows that most of the estimates were rounded to multiples of 5 min, which refers to a typical whole number bias (vertical jumps on the red/lighter cumulative distribution). Although to a lesser extent, this was also the case for the recorded clock times, which could result from researchers' tendency to note down the clock time to the closest multiple of 5.

The best fit linear and power functions mapping retrospective estimates onto elapsed time were $5.92+.67 \cdot t$ and $1.84 \cdot t^{0.78}$ (R^2 s of .46; Fig. 1, left top panel), respectively. Fits to the average estimates led to similar conclusions with R^2 values $> .96$. The best fit linear function to the log-log data was $.73+.73 \cdot t$ (Fig. 1, left middle panel). Since we observed a tendency to report clock times at multiples of 5, we repeated the same analyses, excluding any elapsed durations that are multiples of 5 ($n = 10,196$; 43% of the data). We gathered nearly identical fits to the data $5.43+.68 \cdot t$ and $1.72 \cdot t^{0.80}$. The visual inspection of Fig. 1 (left top and middle panels) shows that subjective time estimates crossed the identity line at ~ 15 min and thus participants were most accurate in timing intervals between 14 and 16 min².

² When the same analyses were conducted without any exclusions, the best fit for linear and power functions were $14.1+.37t$ (R^2 of .31) and $2.77t^{0.66}$ (R^2 of .36). The results were $10.03+.51 \times t$ and $2.82t^{0.66}$ when multiples of 5 min in elapsed times were excluded.

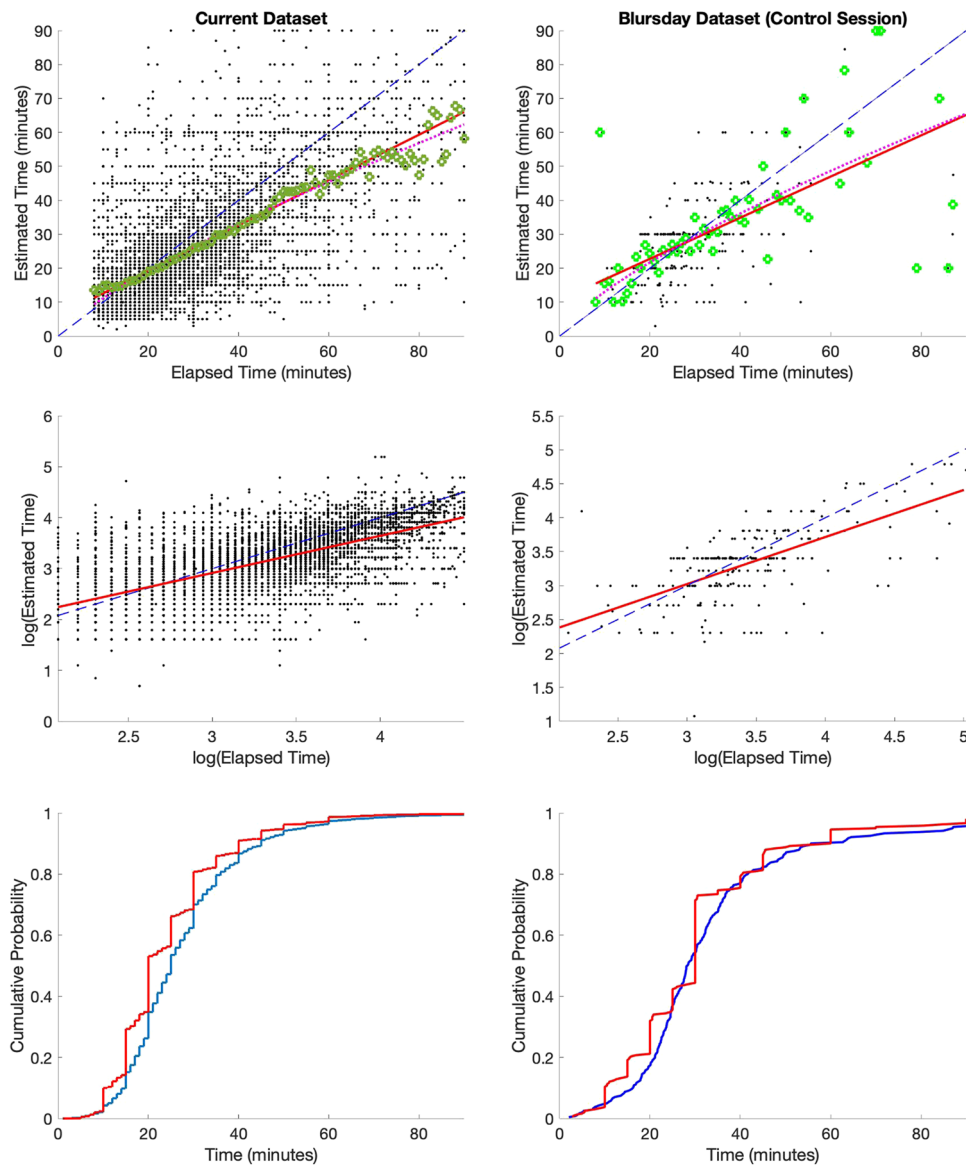


Fig. 1 The left column depicts the data collected in the current study. The right column depicts the data collected as part of the Blursday Database (Chaumon et al., 2022). Top Panels: Linear (solid red) and power function (dotted magenta) fits to the individual participants’ retrospective time estimates. The dashed blue line denotes the identity line between objective and subjective time. Green diamonds are aver-

age estimated durations for different elapsed times. Note that the analyses were conducted for all estimates up to 180 min for both datasets, but the figure is truncated at 90 min for ease of visualization. Middle Panels: Linear function fits to the log-transformed elapsed and estimated durations. Bottom Panels: Cumulative distribution of elapsed (blue/darker) and estimated times (red/lighter)

We also tested the effect of participant age on elapsed and estimated time (excluded were two participants > 80 years of age and one participant with an unreliable age record). Results pointed at a significant relation with elapsed time, $F(1,23646) = 1160.48, p < .001 (R^2 = .05, \beta = .25)$ and estimated time, $F(1,23646) = 864.43, p < .001 (R^2 = .04, \beta = .22)$, but not with proportion of estimated time to elapsed time, $F(1,23646) = .03, p = .87 (R^2 = .00, \beta = -.00)$.

Finally, we characterized how the variability in time estimates changed as a function of elapsed time (Fig. 2). To this end, we calculated the coefficient of variation (CV) of time estimates for different durations. A vertically shifted exponential function (with a constant added to the parent function) provided a good fit to the relationship between CV and elapsed time: $.35 + e^{-.16 * x} (R^2 = .49)$. Identical results were gathered when CVs corresponding to multiples of 5 min were excluded from the analysis: $.35 + e^{-.16 * x} (R^2 = .46)$.

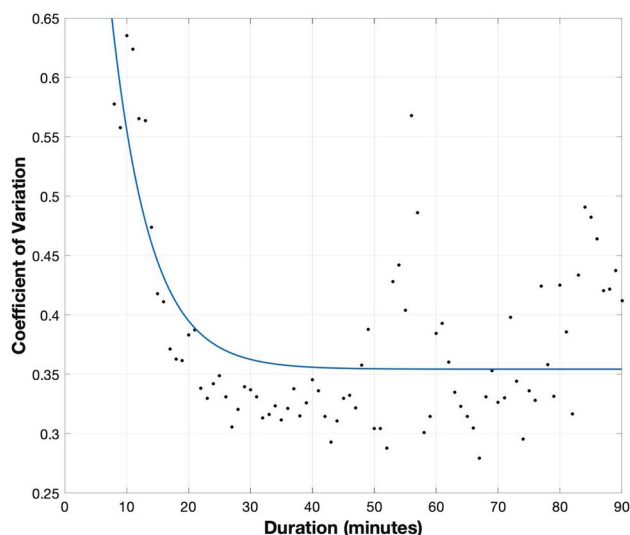


Fig. 2 Vertically shifted exponential function ($a + e^{-\lambda t}$) fit to the coefficient of variation (CV) as a function of elapsed time. The best-fit function to these data show that CVs were calculated separately for each estimated duration in terms of minute-long bins

This fit showed a minimum between-individuals variability in retrospective time estimates at around 45 min.

Blursday control session data The best fit for linear and power functions were $10.65 + 0.61 \cdot t$ and $2.41 \cdot t^{0.73}$ (R^2 s of .51), respectively. Fits to the averaged data supported the same conclusions. The best-fit linear function to the log-transformed elapsed and estimated time was $.94 + 0.69 \cdot t$. Due to the smaller number of data points, we could not analyze the CV values. The bias of estimating the elapsed time at multiples of 5 min was clear in the Blursday control session dataset. Finally, note that the crossover of the elapsed and estimated time was later than 15 min in this dataset.

Discussion

The current study aimed to bridge the historical gap in retrospective time perception by providing conclusive evidence regarding the psychophysical function that best describes subjective time and its variability patterns for an unprecedentedly wide range of durations (8–180 min). These big data are the first to sample a unique duration on a large sample of independent individuals. This differs from previous literature and truly responds to retrospective time.

We found that subjective and objective time were linearly related with a slope considerably lower than the unity line. Participants overestimated intervals that were shorter than 14 min and underestimated intervals that were longer than 16 min. These results were corroborated by the secondary analysis of the Blursday dataset (although the cross-over for

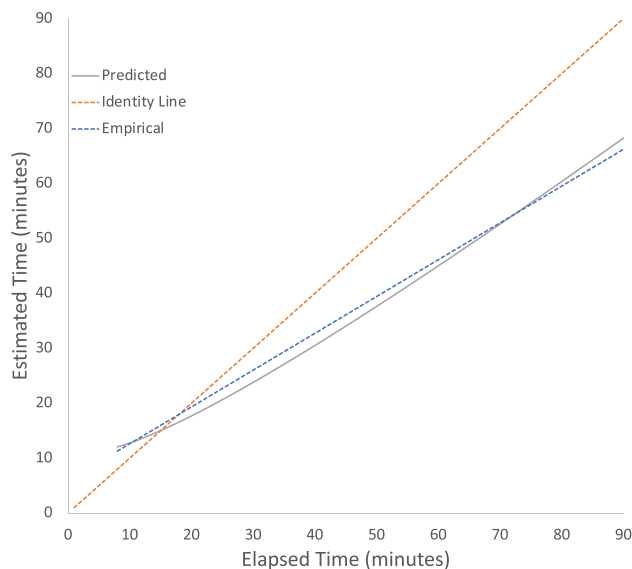


Fig. 3 Predictions of Bayesian averaging with a 15-min long prior

that dataset was longer than 15 min). This difference could be due to the different nature of the tasks in the two studies and/or due to the carry-over effects of the lockdown and continuing COVID-19 pandemic. These findings are also consistent with the conclusions of earlier work conducted in the lab with a narrower range of intervals (e.g., Brown & Stubbs, 1988; Grondin & Plourde, 2007). Briefly, our findings suggest that the retrospective timing is most accurate around 15–20 min, but longer intervals are severely underestimated. For instance, a 90-min long event is underestimated by half an hour in the current dataset. The overestimation of shorter durations is proportionally as prominent as the underestimation of long intervals, which was nearly 70% for 8 min and decreased to 45–48% for 9 and 10 min in the current dataset.

It is possible that 15 min served as a strong prior in the task used in the current study (e.g., due to the expected task duration), leading to an overestimation of shorter and an underestimation of longer intervals because of Bayesian integration. Within this framework, time estimates can be calculated as $15 \text{ min} * (\sigma_t^2 / (\sigma_{15m}^2 + \sigma_t^2)) + \text{Target} * (\sigma_{15m}^2 / (\sigma_{15m}^2 + \sigma_t^2))$. When this calculation is realized for an arbitrary level of CV and the resultant variance decreases according to the power function (i.e., $t^{-2.5}$), one can approximate the gathered relationship (Fig. 3). This prior would have to be considerably longer for the Blursday database. This assumption differs from the regression to the mean, where the prior is determined as the middle of the range of the experienced intervals (Jazayeri & Shadlen, 2010). In fact, the intervals would not be expected to regress to the mean of the intervals in a retrospective timing study since those intervals would be experienced by individual subjects and would not

be accessible to others. In any case, the fact that this pattern was observed in both datasets points to the robustness of these biases in retrospective time judgments. The cognitive process that might underlie such biases is offered by Stewart et al. (2006), who suggest that task-dependent values are compared to a sample of values from memory reflecting the distribution of the corresponding values in real-life experience. Unfortunately, we do not have access to real-world distributions of event durations (akin to word frequency analysis in language studies), although earlier attempts have been made to capture the duration of how long things last (Sandow et al., 1977). This would be an interesting future undertaking that could provide what informative priors participants might use for different types of tasks or contexts. An interesting prediction of this account is that participants characterized by hypo-priors (e.g., autistic individuals; Pellicano & Burr, 2012) might provide more veridical retrospective time estimates.

Glasauer and Shi (2022) showed that individual beliefs in how sensory stimuli are generated in a task could account for individual differences in perceptual biases (see also Glasauer and Shi (2021) for applying this idea to the central tendency effect). Future studies are needed to test how prior expectations (e.g., regarding task duration, cultural expectations) affect retrospective timing. This can be tested by manipulating instructions regarding how much, on average, the same task would last and then testing whether the time estimates show systematic biases based on these priors. For instance, this can be achieved using a similar approach to that used by Tanaka and Yotsu-moto (2017), who tested the violation of temporal expectations but using the passage of time measurements (see also Sackett et al., 2010).

The non-linear modulation of the CV as a function of time is also consistent with earlier research (see Grondin & Plourde, 2007, Fig. 2; up to 8 min). Such a mapping between CV and elapsed time (a particularly higher variability for shorter intervals) can be due to the impact of using 5-min reference units for estimation; such a verbalizing whole number strategy would have a larger effect for shorter intervals. But the fact that the timing variability reaches the lower asymptote much later than 15 min is surprising given that the highest accuracy was observed around 15-min long intervals.

For briefer intervals up to 25 min, the CV gradually decreases with an increase of the interval magnitude; from 25–45 min, the CV remains constant. Such a pattern is similar to observations with the discrimination of brief time intervals. According to Weber's law, the variability of estimates increases linearly with the magnitude of the interval (the scalar property), with the variability-to-time ratio remaining constant (the Weber fraction). However, it is the generalized form of Weber's law that is reported to

hold for a given duration range (Getty, 1975). The generalized form allows accounting for the increased value of the Weber fraction for very brief intervals. This increase is attributed to non-temporal factors such as the sensory properties of signals used to mark intervals (Grondin, 1993). In the present study, such a non-temporal factor would be linked to the need to use numerical units and to the tendency to approximate duration, for example, to the nearest 5 min; the magnitude of this non-temporal source of error weights is larger for briefer durations.

Interestingly, with intervals longer than 1.5 (Gibbon et al. 1997; Grondin, 2012, 2014), or 2 s (Getty, 1975), the Weber fraction is constant for a small duration range and then increases as intervals get longer. In the present study, something similar could be observed. The non-linear fit of the CV showed a variability minimum of retrospective time estimates at around 45 min, and then much more variability with longer intervals. This might be another case of priors affecting the task (e.g., class times), but future systematic studies are needed to test this possibility.

Finally, consistent with the prominent numbers approach (Albers & Albers, 1983) and the findings of Conserve and Dennis (2018), we found that participants were more likely to use multiples of fives to estimate durations. This observation also held for the Blursday control session dataset. Crucially, a similar systematic bias was observed in an earlier study of time interval estimates in real life. Specifically, 95% of the participants in Branas-Garza et al. (2004) rounded their estimates of wait times in the doctor's office to multiples of 5 min. Although this study used data from over 40,000 participants in a database, they did not have an objective benchmark for actual wait times and thus could not address the relationship between retrospective timing and elapsed time. Such rounding of intervals is also observed even with shorter intervals (Block et al., 2018). This tendency can simply be due to psychological factors in choosing prominent numbers (Conserve & Dennis, 2018), including the manifestation of priors in time estimation and/or reporting to the verbalization or numerization of analog values. Moreover, our timing system does not have the accuracy in estimating an interval to the minute, say 28 min, and instead we indicate 30 min.

One important point is that the to-be-estimated durations depended on the participant's behavior. Thus, participants estimating short and long intervals in this study may differ on one fundamental feature, namely the speed of reading and answering (filling out the questionnaires), which may relate to cognitive capabilities or an internal propensity to take time to do things. The former group consisted of those who completed the main task quickly and then overestimated that interval. For instance, individuals who are faster have stronger attentional focus, and faster information processing,

which in retrospect would lead to a relatively stronger build-up of contextual memory changes, which in turn would lead to a relative overestimation of duration (e.g., Zakay & Block, 1997). The crossover from overestimation to underestimation may, therefore, not be due to some systematic property of timing (or inferred prior) but may be due to this covariate. Although to a lesser extent, the same issue also applies to the Blursday data. Ideally, the interval to be retrospectively judged by participants should be under experimental control. Crucially, the regression of elapsed and estimated time on the age of the participants revealed a very weak positive relation but age did not predict the normalized estimated times, suggesting that age was not a predictor of these patterns. Future data collection should control for such factors.

Taking a different perspective, this feature of our study is also what constitutes the novelty and ecological value of the task: true retrospective timing can only be a single-trial experiment to prevent attentional orientation and a cognitive strategy from the participants (e.g., Azizi et al., 2021). In the absence of knowledge as to the duration estimation requirements that will be asked of them retrospectively, participants have no reason to perform a goal-oriented timing task or to develop particular biases intrinsic to timing. In daily life too many of our time judgments occur under such circumstances and lack of expectancies, and thus without the adoption of auxiliary mechanisms that might mask the true nature of temporal judgments. This would in turn lead to a more prominent manifestation of inherent links between subjective time and other cognitive processes (e.g., memory encoding/retrieval efficacy) when they are not directed towards temporal information processing. These might indeed be the conditions that are most sensitive to individual differences since cognitive compensation would not be expected without a timing-based task goals. To our knowledge, this study investigated the widest range and the longest target interval (see Tobin et al., 2010), and future lab studies would help validate these results. These studies can also utilize tasks to better elucidate the association between retrospective time estimates and cognitive traits.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13423-023-02277-3>.

References

- Albers, W., & Albers, G. (1983). On the prominence structure of the decimal system. In *Advances in Psychology* (vol. 16, pp. 271–287). North-Holland.
- Aydoğan T, Karşilar H, Duyan YA, Akdoğan B, Baccarani A, Brochard R, et al. (2023) The timing database: An open-access, live repository for interval timing studies. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-022-02050-9>
- Azizi, L., Polti, I., & van Wassenhove, V. (2021). Episodic timing: How spontaneous alpha clocks, retrospectively. *bioRxiv*, 2021-10. <https://doi.org/10.1101/2021.10.01.462732>
- Balci, F., Meck, W. H., Moore, H., & Brunner, D. (2009). Timing deficits in aging and neuropathology. In J. Woods & A. Bizon (Eds.), *Animal Models of Human Cognitive Aging* (pp. 161–201). Humana Press.
- Block, R. A. (1985). Contextual coding in memory: Studies of remembered duration. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 169–178). Springer-Verlag.
- Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, 4, 184–197.
- Block, R. A., Grondin, S., & Zakay, D. (2018). Prospective and Retrospective Timing Processes: Theories, Methods, and Findings. In A. Vatakis, F. Balci, M. Di Luca, & Á. Correa (Eds.), *Timing and Time Perception: Procedures, Measures, & Applications* (pp. 32–51). Brill. <https://doi.org/10.1163/j.ctvbqs54b.6>
- Brañas-Garza, P., Morales, A. J., & Serrano-del-Rosal, R. (2004). *Estimating time by counting hands*. IESA Working Papers Series 0401, Institute for Social Studies of Andalusia - Higher Council for Scientific Research.
- Brown, S. W., & Stubbs, D. A. (1988). The psychophysics of retrospective and prospective timing. *Perception*, 17(3), 297–310.
- Chaumon, M., Rioux, P.-A., Herbst, S., Spiouas, I., Kübel, S., Gallego Hiroyasu, E., Runyun, S., Micillo, L., Thanopoulos, V., Mendoza-Durán, E., Wagelmans, A., Mudumba, R., Tachmatzidou, R., Cellini, N., D'Argembeau, A., Giersch, A., Grondin, S., Gronfier, C., Alvarez Igarzábal, F., ... van Wassenhove, V. (2022). The Blursday Database: Individuals' Temporalities in Covid Times. *Nature Human Behaviour*, 6, 1587–1599.
- Converse, B. A., & Dennis, P. J. (2018). The role of “Prominent Numbers” in open numerical judgment: Strained decision makers choose from a limited set of accessible numbers. *Organizational Behavior and Human Decision Processes*, 147, 94–107.
- Getty, D. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, 18, 1–8.
- Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology*, 7(2), 170–184.
- Glasauer, S., & Shi, Z. (2021). The origin of Vierordt's law: The experimental protocol matters. *PsyCh Journal*, 10(5), 732–741.
- Glasauer, S., & Shi, Z. (2022). Individual beliefs about temporal continuity explain variation of perceptual biases. *Scientific Reports*, 12, 10746.
- Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics*, 54, 383–394.
- Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 880–890.
- Grondin, S. (2014). About the (non) scalar property for time perception. In *Neurobiology of interval timing* (pp. 17–32). Springer.
- Grondin, S., & Killeen, P. R. (2009). Tracking time with song and count: Different Weber functions for musicians and nonmusicians. *Attention, Perception & Psychophysics*, 71(7), 1649–1654.
- Grondin, S., & Laflamme, V. (2015). Stevens's law for time: A direct comparison of prospective and retrospective judgments. *Attention Perception & Psychophysics*, 77, 1044–1051.
- Grondin, S., & Plourde, M. (2007). Judging multi-minute intervals retrospectively. *Quarterly Journal of Experimental Psychology*, 60(9), 1303–1312.
- Grondin, S., Ouellet, B., & Roussel, M. E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals.

- Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 58(1), 1–12.
- Hicks, R. E. (1992). Prospective and retrospective judgments of time: A neurobehavioral analysis. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, action and cognition: Towards bridging the gap* (pp. 97–108). Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-94-017-3536-0_12
- James, W. (1890). *The Principles of Psychology*. Henry Holt and Company the Principles of Psychology. <https://doi.org/10.1037/11059-000>
- Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, 13(8), 1020–1026.
- Jokic, T., Zakay, D., & Wittmann, M. (2018). Individual differences in self-rated impulsivity modulate the estimation of time in a real waiting situation. *Timing & Time Perception*, 6(1), 71–89.
- Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510.
- Rattat, A. C., & Droit-Volet, S. (2012). What is the best and easiest method of preventing counting in different temporal tasks? *Behavior Research Methods*, 44(1), 67–80.
- Sackett, A. M., Meyvis, T., Nelson, L. D., Converse, B. A., & Sackett, A. L. (2010). You're having fun when time flies: The hedonic consequences of subjective time progression. *Psychological Science*, 21(1), 111–117.
- Sandow, S. A., Bamber, C., & Rioux, J. W. (1977). *Durations: The encyclopedia of how long things take*. Times Book.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64(3), 153–181.
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1), 1–26.
- Tanaka, R., & Yotsumoto, Y. (2017). Passage of time judgments is relative to temporal expectation. *Frontiers in Psychology*, 8, 187.
- Tobin, S., & Grondin, S. (2009). Video games and the perception of very long durations by adolescents. *Computers in Human Behavior*, 25(2), 554–559.
- Tobin, S., Bisson, N., & Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: A study involving gamers. *PLoS ONE*, 5(2), e9271.
- Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, 61(4), 569–587.
- Witowska, J., Schmidt, S., & Wittmann, M. (2020). What happens while waiting? How self-regulation affects boredom and subjective time during a real waiting situation. *Acta Psychologica*, 205, 103061.
- Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science*, 6(1), 12–16.
- The authors are ordered alphabetically after the second author based on the last name. This study was not pre-registered. This work is funded by NSERC Discovery Grant, RGPIN/3334-2021 to FB.
- Open practices statement** The data for the current experiment are available in the Timing Database at <https://osf.io/vrwjz/>. The Blursday data are currently available at <https://osf.io/359qm/>. None of the experiments was preregistered.
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.