UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Task-oriented Bayesian inference in interval timing: People use their prior reproduction experience to calibrate time reproduction

Permalink

Journal

Authors
Cai, Zhenguang
Wang, Ruiming
Liu, Huihui
et al.

Publication Date

Peer reviewed

Task-oriented Bayesian inference in interval timing: People use their prior reproduction experience to calibrate time reproduction

Zhenguang Cai (Zhenguang.cai@uea.ac.uk)
School of Psychology, University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK
Ruiming Wang (wangrm@scnu.edu.cn)
School of Psychology, South China Normal University, Tianhe District, Guangzhou, 510631, China
Huihui Liu (1425357597@qq.com)
School of Psychology, South China Normal University, Tianhe District, Guangzhou, 510631, China
Maarten Speekenbrink (m.speekenbrink@ucl.ac.uk)
Department of Experimental Psychology, University College London, London, WC1H 0AP, UK

Abstract

The estimation of duration has been shown to follow Bayesian inference, where people use their prior belief to calibrate the estimation. This explains timing biases such as the range bias where a duration is reproduced as longer when previously encountered durations were longer than shorter. However, it is unclear whether prior belief is based on previously perceived or reproduced durations. In 4 experiments, we show that the range bias occurs between short and long *reproduction* ranges but not between short and long *perception* ranges. Further analyses also show that the prior is updated by the most recent reproduced (but not perceived) duration. Together these results support a task-oriented Bayesian inference account of time reproduction, where people use the perceived duration and their past reproduction experience to make an inference about how much time to reproduce.

Keywords: time perception; Bayesian inference; memory; psychophysics

Introduction

The mind is good at estimating quantitative dimensions of the physical world: we are able to estimate how much time has elapsed, how much distance has been traveled, how large an area is, etc. Indeed, our superb capacity to quantify things enables us to better adapt to the environment. However, these quantitative intuitions are not without errors. Systematic biases in human magnitude estimation have been identified (for reviews see Poulton, 1979, and Petzschner, Glasauer, & Stephan, 2015). Of these, the most robust is probably the bias of central tendency (Hollingworth, 1910), a phenomenon which has also been known by a variety of other names (e.g., contraction, regression effect, regression toward the mean). Central tendency refers to the observation that people tend to make estimates closer to the mean of the magnitudes to be estimated, leading to the (relative) overestimation of lower magnitudes in the stimulus set and underestimation of higher magnitudes. A central tendency bias has been observed in the estimation of distance (Jou et al., 2004; Radvansky et al., 1995), brightness (Fotios & Cheal, 2007), weight (Jones & Hunter, 1982), and loudness (Algom & Marks, 1990), to mention a few. In particular, the central tendency bias has been most often observed in the estimation of time intervals, with the tendency for people to relatively overestimate shorter durations and underestimate longer durations (Bausenhart, Dyjas, & Ulrich 2014; Gu & Meck, 2011;

Jazayeri & Shadlen, 2010; Lejeune & Wearden, 2009; Moon & Anderson 2013). More interestingly, as a result of the central tendency bias, a stimulus duration tends to be reproduced as longer if it occurs as a member of a longer range (e.g., 1000 ms in the range of 847 - 1200 ms) than as a member of a shorter range (1000 ms in the range of 671 - 1023 ms) (Jazayeri & Shadlen, 2010), a phenomenon which we refer to as the range bias.

Jazayeri and Shadlen (2010) argue that the central tendency and the range bias occur because time estimation follows Bayesian inference: as memories of durations (indeed magnitudes in general) are inherently noisy (Gallistel & Gelman, 2000), people resort to their prior belief about how likely a duration is in order to calibrate their estimate of the magnitude of a perceived duration. In Bayesian inference, a posterior belief is the product of the "likelihood" (reflecting the variability of perceptions of a given duration) and the prior, and this posterior (the estimated duration) is necessarily pulled toward the mean of the prior distribution (e.g., the midpoint of experienced durations), hence the central tendency and the range bias. Such a Bayesian inference account of timing has been endorsed in many subsequent related research on timing behaviour, e.g., animal timing (Li & Dudman, 2014), time prediction (Griffiths & Tenenbaum, 2011; Di Luca & Rhodes, 2016) and delay discounting (McGuire & Kable, 2012) (see Shi, Church & Meck, 2014, for a review). Indeed, Petzschner et al. (2015) argue that Bayesian inference is used in estimating all kinds of physical magnitudes and can accommodate a wide range of behavioural effects in magnitude estimation, central tendency included.

A common assumption in Bayesian accounts of timing (and indeed magnitude estimation in general) is that people make an optimal estimate about how long the perceived duration is by incorporating their belief about how likely a duration is as a function of previously presented magnitudes; in a time reproduction task, they then reproduce an amount of time to match this estimate (Cicchini et al., 2012; Jazayeri & Shadlen, 2010; Di Luca & Rhodes, 2016). These accounts thus assume that the estimation of a perceived duration makes reference to the previously *perceived* durations.

However, these accounts ignore another source of information that participants can rely upon when trying to optimally reproduce durations, i.e. their past experience with

reproduced durations. According to this alternative account, instead of making an estimate of a perceived duration and then use this estimate to guide time reproduction, people directly make an estimate about how long the reproduced duration should be. In this *task-oriented* inference, people make use of their past reproduction experience rather than their past perception experience; after all, when one is to reproduce a duration, the history of the reproduced durations may provide a better constraint for optimally determining how much time to reproduce.

The current study investigates whether people resort to previous perception or reproduction experience as the prior in their inferences. To do so, we take advantage of the range bias (a duration is reproduced as longer if it is placed in a context of long than short stimulus durations). If people use their perception experience to calibrate their time estimation, then we should expect the range bias to occur between contexts of long vs. short perceived durations, even when the context of reproduced durations is kept constant. If instead people use their reproduction experience in their inference, we should expect the range bias to occur between contexts of long vs. short reproduced durations, even when the context of perceived durations is kept constant.

Experiment 1

Methods

Participants. 32 volunteers (24 females, 20.3 ±1.6 in age) from the South China Normal University community volunteered for a small monetary reward. Participants in this experiment (and indeed in each experiment reported here) did not take part in any other experiment, though participants for all the experiments came from the same participant pool. These participants (and also those in other experiments) had normal or corrected-to-normal vision and received a small monetary reward for their participation

Design. We manipulated the reproduction context while keeping the perception context constant. To do this, we used an alternative-task paradigm where participants perceived a duration and then, upon a cue, *either* reproduced the stimulus duration *or* compared it to a new duration. Participants were presented with a set of *short durations* (600 – 2200 in steps of 200 ms) a set of *long durations* (1800 – 3400 in steps of 200 ms), all interleaved. Half of the participants reproduced the short durations and compared the long durations; the other half did the opposite (i.e. reproducing long durations and comparing short durations). Note that, in such a design, while the two groups of participants differed in their reproduction contexts, they had identical perception contexts. Critically, the two reproduction contexts overlapped in three stimulus durations (1800, 2000, 2200 ms), which allowed us to determine whether different reproduction contexts lead to a range bias, even when the perception context was kept constant.

Materials. For each of the 18 durations, a shorter (0.1 log shorter) and a longer (0.1 log longer) comparison duration was created. Each participant completed 5 blocks of trials. In each block, every stimulus duration was presented twice, either both as reproduction or comparison trial. Half of the comparison trials used a shorter comparison duration and the other half had a longer one. Trials in each block were presented in an individually randomized order. In total, there were 180 experimental trials.

Procedure. The experiment was run on a desktop using E-Prime 2.0. Participants sat about 50 cm away from the monitor. The experiment began with a practice session of 4 trials (2 reproduction and 2 comparison trials) followed by the main experiment. In a trial, a black cross (Courier New 48) was presented for a stimulus duration, followed by a blank screen of 300 ms. Then, a cue (an asterisk "*" or the phrase "第二段时间", meaning "second duration") was presented. An asterisk informed participants to reproduce the stimulus duration by holding down the spacebar for the same duration. At the press of the spacebar, the asterisk turned into three asterisks which remained on screen until the release of the spacebar. The phrase "第二段时间" informed participants to compare the first (stimulus) duration with an upcoming (comparison) duration. The text cue stayed on screen for 1 s and was replaced by a blank screen of 300 ms. The comparison duration was then presented with a blue cross (Courier New 48), followed by a blank screen of 300 ms. Then a judgment screen was displayed asking participants to decide whether the first (stimulus) or second (comparison) duration was longer by pressing the "F" or "J" key.
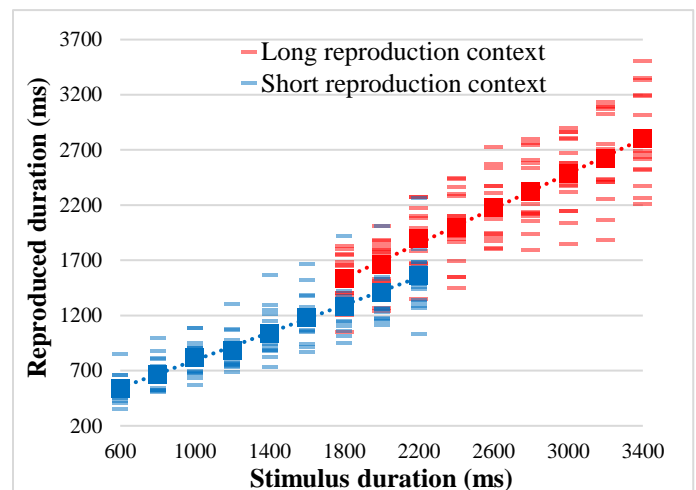

Fig. 1. Results of Experiment 1. Each dash represents a participant's averaged reproduction for each stimulus duration in the long (red) or short (blue) reproduction context; the squares represent the averaged reproductions at the group level.

Results and discussion

As the comparison data does not address our theoretical interest, to save space, we only report analyses on the reproductions. We first excluded as outliers reproductions that were 1/3 or 3 times the stimulus duration, leading to the loss of 2% (65 reproductions) of the data (this trimming

criterion was adopted for all experiments reported in this study). For the remaining data, a participant's mean reproduction for each stimulus duration was computed. We compared reproductions for the three overlapping stimulus durations (1800, 2000, and 2200 ms) between the two reproduction contexts. An ANOVA with reproduction context as a between-participant factor and stimulus duration as a within-participant continuous variable revealed a significant main effect of reproduction context ($F_{(1,30)} = 12.41$, $p = .001$, $\eta^2 = .29$): durations were reproduced for longer in the long than the short reproduction context (see Fig. 1). Reproduced durations increased as a function of the stimulus duration ($F_{(1,30)} = 72.03$, $p < .001$, $\eta^2 = .71$) and did not significantly interact with reproduction context ($F_{(1,30)} = 1.44$, $p = .240$, $\eta^2 = .05$).

The main effect of reproduction context suggested a range bias in the reproduction of the overlapping durations: even when the perception context was kept constant, durations were reproduced as longer when prior reproductions were longer. Such a *reproduction* range bias is inconsistent with previous Bayesian inference accounts which posit that, to estimate a duration, people use their memory of the perceived duration and experience of previously *perceived* durations. Instead, the results suggest that people make use of their experience of previously *reproduced* durations in order to calibrate their reproductions.

In Experiment 2, we aimed to replicate the reproduction range bias using a within-participant design. In particular, we distinguished the long and short reproduction contexts using different modalities of reproduction: for half of the participants, people reproduce long durations with motor reproduction and short durations with an auditory reproduction (see below); for the other half, the paring was reversed. If reproduction experience calibrates duration reproduction, we should again expect a range bias for overlapping durations between the two reproduction contexts.

Experiment 2

Methods

Participants. 20 volunteers (10 females, 20.8 ±2.5 in age) took part in the experiment.

Design. This experiment was similar to Experiment 1 except that we replaced the comparison task in Experiment 1 with an auditory reproduction task (and we manipulated reproduction task within-participants). As in Experiment 1, there were two duration ranges (short range: 600 – 2200 in steps of 200 ms; long range: 1800 – 3400 in steps of 200 ms). Two experimental versions were created such that one version had the short range paired with motor reproduction and the long range with auditory reproduction and the other version had the reverse. As in Experiment 1, we were interested whether people would be susceptible to the reproduction range bias when reproducing the overlapping durations (i.e. 1800, 2000 and 2200 ms) under different reproduction contexts.

Materials. As in Experiment 1, there were 5 blocks of trials and each block contained two occurrences of each of the 18 stimulus durations (i.e. 36 trials in each block). Trials in each block were presented in an individually randomized order. For auditory reproduction, a 10s sine-wave pure tone sampled at a rate of 44100 Hz was created using Audacity.

Procedure. The experimental setting and overall experimental procedure were the same as those in Experiment 1, except that participants always reproduced a stimulus duration. After a cross was presented for a stimulus duration, followed by a blank, an image of a keyboard (as a cue for motor reproduction) or a mouse (as a cue for auditory reproduction) was displayed. For motor reproduction, as in Experiment 1, participants held down the spacebar to reproduce the stimulus duration. For auditory reproduction, participants clicked the mouse (at which point the mouse image disappeared) to initiate a tone and clicked again to terminate it when they felt that tone had been played for the same length as the stimulus duration. The experiment lasted for about 25 min.
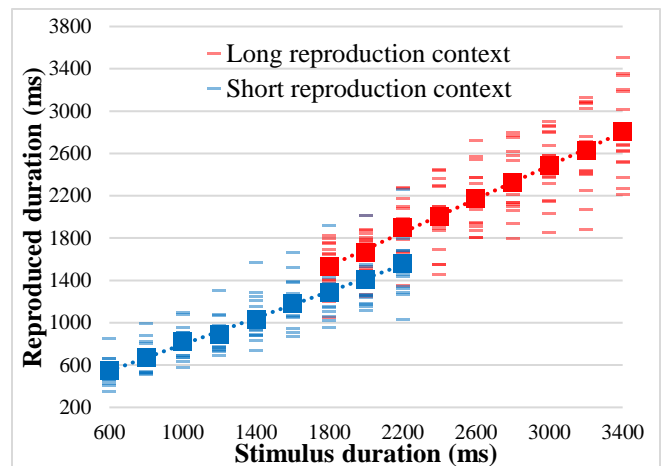

Figure 2. Results of Experiment 2. Each dash represents a participant's averaged reproduction for each stimulus duration in the long (red) or short (blue) reproduction range; the squares represent the averaged reproductions at the group level.

Results and discussion

About 6% (216 reproductions) of the data were excluded as outliers (the high exclusion rate was due to the fact that participants sometimes accidentally pressed rather than held down the spacebar due to the influence of the auditory reproduction method). For the overlapping durations, an ANOVA using reproduction context and stimulus duration as within-participant variables reveals a significant main effect of reproduction context ($F_{(1,19)} = 27.73$, $p < .001$, $\eta^2 = .59$), with longer reproductions of the overlapping stimulus durations in the long compared to the short reproduction context. Reproductions also increased as a function of stimulus duration ($F_{(1,19)} = 131.5$, $p < .001$, $\eta^2 = .87$). There was no significant interaction between reproduction context and stimulus duration ($F_{(1,19)} = 0.13$, $p = .724$, $\eta^2 = .01$).

194

The results thus replicated, with a within-participant design, the finding in Experiment 1 that reproduction experience calibrates duration reproductions. That is, a duration was reproduced as longer if it was done in the context of long reproductions compared to short reproductions. They also suggest that people construct task-specific priors (i.e. past motor vs. auditory reproduction experience in the current experiment) in their time estimation, an issue that awaits further empirical verification.

In Experiment 3, we further explore whether manipulating the perception context alone leads to a range bias. If reproduction experience, but not perception experience, calibrates reproduction, we should not see a range bias in this experiment.

Experiment 3

Methods

Participants. 20 volunteers (14 females, 20.0 ±1.2 in age) took part in Experiment 3.

Design. We manipulated perception context (short vs. long) within-participants in a blocked design. As in Experiment 1, Experiment 3 used an alternative-task paradigm (reproduction or comparison). The two perception contexts were created using three ranges of durations: the short perception context consisted of 6 short durations (600 – 1600 in steps of 200 ms) serving as comparison durations and 6 mid durations (1200 – 2200 in steps of 200 ms) serving as reproduction durations; the long perception context consisted of the 6 mid durations serving as reproduction durations and 6 long durations (1800 – 2800 in steps of 200 ms) serving as comparison durations. Thus, the two perception contexts had the same range of durations to be reproduced (i.e. both had the mid durations for reproduction) but differed in the range of durations to be perceived (long and mid durations for the long perception context but short and mid durations for the short perception context). If the perception context manipulation leads to a range bias, we should expect the mid durations to be reproduced as longer in the long than in the short perception context. Alternatively, if the range bias is driven by reproduction experience only, we should expect the mid durations to be reproduced as equally long between the two perception contexts.

Materials. As in Experiment 1, a shorter (0.1 log shorter) and longer (0.1 log longer) comparison duration were created for each of the comparison durations. Three blocks of materials were created for both the short and the long perception context. In each block, each stimulus duration was presented twice for reproduction and twice for comparison (once with a longer comparison duration and once with a shorter comparison duration), amounting to 24 trials in each block. Two experimental versions were created: the three blocks of the short perception context preceded those of the long perception context in one version and the order was reversed in the other. A short practice block of 4 trials preceded the first block. In order to prevent possible spillover of the perception context in the first three blocks to the last

three blocks, a compulsory 2-min break was inserted after the first three blocks; additionally, a practice block of 12 trials preceded the 4th block.

Procedure. The experimental setting and the trial structure were identical to those in Experiment 1; that is, after the presentation of a stimulus duration, depending on the ensuing cue, participants either reproduced the stimulus duration or compared it with an upcoming duration. During the 2-min break, participants were allowed to do whatever they liked as long as they remained seated in the test cubicle. The experiment took about 25 min.
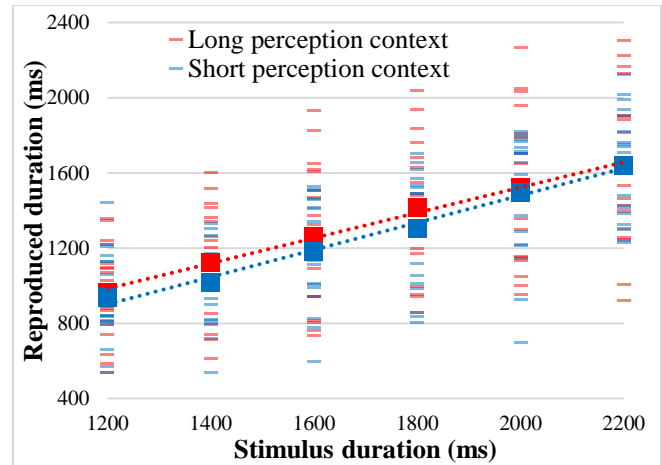

Figure 3. Results of Experiment 3. Each dash represents a participant's averaged reproduction for each stimulus duration in the long (red) or short (blue) perception context; the squares represent the averaged reproductions at the group level.

Results and discussion

The trimming criterion led to the exclusion of 5% (80 data points) of all the reproduced durations. An ANOVA with perception context as a within-participant factor and reproduction durations (1200 – 2200) as a within-participant continuous variable revealed no significant main effect of perception context ($F_{(1,19)} = 0.82$, $p = .377$, $\eta^2 = .04$), suggesting that the reproductions of stimulus durations were similar between the two perception contexts. Reproduction increased as a function of the stimulus duration ($F_{(1,19)} = 255.3$, $p < .001$, $\eta^2 = .93$). There was no significant interaction ($F_{(1,19)} = 1.18$, $p = 290$, $\eta^2 = .06$).

The failure for the perception context manipulation to induce a range bias suggests that participants did not use their perception experience to infer stimulus duration for their reproductions. In Experiment 4, we changed all the comparison trials in Experiment 3 into reproduction trials so that the long and short perception context had respectively a long and short reproduction range (i.e. we additionally introduced the reproduction context manipulation). If people use reproduction experience to calibrate their time estimation, we should restore the range bias that was missing in Experiment 3.

Experiment 4

Methods

Participants. Another 20 volunteers (14 females, 19.9±1.1 in age) took part in the experiment.

Design, materials and procedure. These were the same as those in Experiment 3 except that the comparison trials in Experiment 3 were changed into reproduction trials. Thus, the short perception context had 6 short durations (600 – 1600 in steps of 200 ms) and 6 mid durations (1200 – 2200 in steps of 200 ms), all to be reproduced; the long perception context had 6 mid durations (1200 – 2200 in steps of 200 ms) and 6 long durations (1800 – 2800 in steps of 200 ms), all to be reproduced.
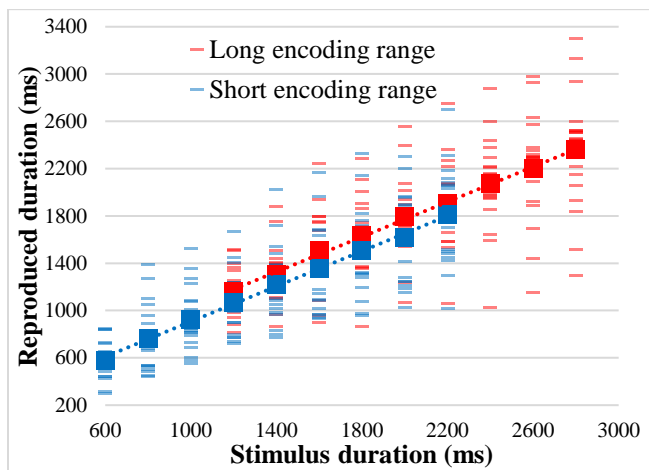

Figure 4. Results of Experiment 4. Each dash represents a participant's averaged reproduction for each stimulus duration in the long (red) or short (blue) perception/reproduction context; the squares represent the averaged reproductions at the group level.

Results and discussion

We excluded 3% (106 data points) of all the reproduced durations as outliers. We compared the reproductions of the 6 overlapping stimulus durations (1200 – 2200 in steps of 200 ms) between the two perception (and indeed reproduction) contexts. In contrast to the finding in Experiment 3, the ANOVA showed a significant main effect of perception/reproduction context ($F_{(1,19)} = 10.00$, $p = .005$, $\eta^2 = .34$), with longer reproductions of the stimulus durations when they were part of the long than the short perception/reproduction context. Reproductions increased as a function of the stimulus duration ($F_{(1,19)} = 255.6$, $p < .001$, $\eta^2 = .93$). The two variables did not interact significantly ($F_{(1,19)} = 0.44$, $p = .514$, $\eta^2 = .02$), suggesting a central tendency in the reproduced durations in both contexts.

The most striking observation is the return of the range bias for the overlapping stimulus durations when different reproduction ranges were introduced, in contrast to Experiment 3, where the reproduction range was the same between the two perception contexts. Such a finding clearly suggests that the reproduction experience, rather than the perception experience, drives the range bias.

Prior updating

A crucial prediction of Bayesian inference in time perception is that the prior is constantly updated. If the prior in time reproduction is based on previously reproduced rather than perceived durations, as our experiments have shown, we should predict the most recently reproduced (but not perceived) duration to have an influence on the prior, and hence on the posterior, such that a longer reproduced (but not perceived) duration in the preceding trial leads to a longer reproduced duration in the current trial.

The comparison trials in Experiments 1 and 3 allowed us to examine the possible influence of the preceding perceived (i.e. stimulus) duration on the prior belief (and the reproduced duration in the current trial). We used linear mixed effects modelling for these analyses, where we included as predictors the stimulus duration of the current trial and the stimulus duration in the preceding trial. For Experiment 1, though reproductions increased as a function of the current trial's stimulus duration ($\beta = 556.28$, $SE = 35.56$, $t_{(31.0)} = 15.65$, $p < .001$), they were insensitive to the magnitude of the stimulus (i.e. perceived) duration in the preceding comparison trial ($\beta = -1.75$, $SE = 17.64$, $t_{(28.1)} = -0.10$, $p = .922$). The same pattern was also observed in Experiment 3 ($\beta = 241.36$, $SE = 16.73$, $t_{(19.4)} = 14.43$, $p < .001$, for current stimulus duration; $\beta = -36.90$, $SE = 26.35$, $t_{(19.1)} = -1.40$, $p = .177$, for preceding stimulus duration).

We next analyzed reproductions taking into account the preceding *reproduced* duration (i.e. when the preceding trial was a reproduction trial) in all the 4 experiments. Reproductions always increased as a function of the stimulus duration of the current trial ($\beta = 570.52$, $SE = 23.86$, $t_{(31.1)} = 23.91$, $p < .001$; $\beta = 643,41$, $SE = 45.47$, $t_{(19.0)} = 14.15$, $p < .001$; $\beta = 225.05$, $SE = 20.50$, $t_{(17.6)} = 10.98$, $p < .001$; : $\beta = 413.54$, $SE = 22.89$, $t_{(19.0)} = 18.07$, $p < .001$; for Experiments 1-4 respectively) and also of the preceding trial ($\beta = 191.46$, $SE = 16.99$, $t_{(38.2)} = 11.27$, $p < .001$; $\beta = 89.34$, $SE = 16.64$, $t_{(18.4)} = 5.37$, $p < .001$; $\beta = 151.48$, $SE = 19.95$, $t_{(20.6)} = 7.59$, $p < .001$; $\beta = 125.94$, $SE = 15.96$, $t_{(17.7)} = 7.89$, $p < .001$; for Experiments 1-4 respectively). These findings consistently suggest that the prior was updated by a recent reproduction output such that a recent longer reproduction increased the prior mean, which in turn increased the posterior mean, resulting in a longer reproduction in the current trial.

General discussion

In four experiments, we showed that people use their past (in particular the most recent) reproduction experience to calibrate their duration reproduction. These results are inconsistent with previous Bayesian inference accounts of timing (and magnitude estimation in general), whereby people use previously perceived durations to calibrate their noisy memory of stimulus duration, which in turn is used to guide reproduction (Jazayeri & Shadlen, 2010; Di Luca & Rhodes, 2016; Cicchini et al., 2015; Petzschner et al., 2015).

Such a memory-optimizing account is explicitly spelled out in Petzschner et al. (2015) as a Bayesian inference account of magnitude estimation in general. In this account, people perceive a magnitude (e.g., a duration) and keep a noisy memory of it. Later they use their prior belief to infer an optimal estimate based on this noisy memory, and this estimate represents the *inferred* stimulus duration that is used to guide subsequent response (e.g., time reproduction).

Instead, the current findings support a *task-oriented* Bayesian inference account, where people directly use the noisy memory of the stimulus duration and their past reproduction experience to infer a reproduction estimate. Note that, unlike previous accounts, such an estimate is not an updated version of the stimulus duration but should instead be viewed as a *planned* reproduced duration.

If it is the case that Bayesian inference is task-oriented (i.e. the inference serves the task at hand), then we should expect the source of the prior information to vary across different magnitude tasks. For instance, it is possible that, whereas time reproduction recruits prior reproduction experience, time comparison may instead recruits prior perception experience as the task would involve making inferences about *perceived* durations. It is also possible, as Experiment 2 suggested, that different reproduction tasks may resort to task-specific priors. These remain to be tested in future studies.

Acknowledgments

References

Algom, D., & Marks, L. E. (1990). Range and regression, loudness scales, and loudness processing: Toward a range-bound psychophysics. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 706.

Bausenhart, K. M., Dyjas, O., & Ulrich, R. (2014). Temporal reproductions are influenced by an internal reference: Explaining the Vierordt effect. *Acta Psychologica, 147*, 60-67.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *Journal of Neuroscience, 32*, 1056-1060.

Di Luca, M., & Rhodes, D. (2016). Optimal perceived timing: integrating sensory information with dynamically updated expectations. *Scientific Reports, 6*.

Fotios, S. A., & Cheal, C. (2007). Evidence for response contraction bias in side-by-side matching tasks. *Lighting Research & Technology, 39*, 159-169.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4*, 59-65.

Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: people combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General, 140*, 725.

Gu, B. M., & Meck, W. H. (2011). New perspectives on Vierordt's law: memory-mixing in ordinal temporal comparison tasks. In *Multidisciplinary aspects of time and time perception* (pp. 67-78). Springer Berlin Heidelberg.

Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*, 461-469.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal range calibrates interval timing. *Nature Neuroscience, 13*, 1020-1026.

Jones, L. A., & Hunter, I. W. (1982). Force sensation in isometric contractions: a relative force effect. *Brain Research, 244*, 186-189.

Jou, J., Leka, G. E., Rogers, D. M., & Matus, Y. E. (2004). Contraction bias in memorial quantifying judgment: Does it come from a stable compressed memory representation or a dynamic adaptation process? *American Journal of Psychology, 117*, 543–564.

Lejeune, H., & Wearden, J. H. (2009). Vierordt's The Experimental Study of the Time Sense (1868) and its legacy. *European Journal of Cognitive Psychology, 21*, 941-960.

Li, Y., & Dudman, J. T. (2013). Mice infer probabilistic models for timing. *Proceedings of the National Academy of Sciences, 110*, 17154-17159.

McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition, 124*, 216-226.

Moon, J., & Anderson, J. R. (2013). Timing in multitasking: Memory contamination and time pressure bias. *Cognitive Psychology, 67*, 26-54.

Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences, 19*, 285-293.

Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin, 86*, 777-803.

Radvansky, G. A., Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Uncertainty in estimating distances from memory. *Memory & Cognition, 23*, 596-606.

Shi, Z., Church, R. M., & Meck, W. H. (2013). Bayesian optimization of time perception. *Trends in Cognitive Sciences, 17*, 556-564.