Title
Shared temporal expectation across higher- and lower-level cognition

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Wang, Muzhi
Zhang, Hang

Publication Date
2021

Peer reviewed

Shared temporal expectation across high- and low-level cognition

Muzhi Wang (wangmuzhi@pku.edu.cn), Hang Zhang (hang.zhang@pku.edu.cn)
School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road
Beijing, 100871 China

Abstract

Temporal expectation for future events allows people to prepare more efficiently for the future. In sensorimotor tasks, it has been considered as an important factor that influences the accuracy and speed of responding to specific sensory events. However, there was no consensus whether the temporal expectation functioning in sensorimotor tasks is simply an emergent property of task-specific, low-level circuits, or an abstract representation shared by higher-level cognition. In four experiments, we asked whether two simultaneously processed tasks—one of lower-level and the other of higher-level cognition—would be influenced by the same temporal expectation. One task was speeded response to a target stimulus, where the target was cancelled on 30% of the trials. The other task was a real-time gambling task, where participants needed to predict from time to time whether the current trial would end up with target or cancellation. Both the target and cancellation latencies followed specific distributions, with the distribution of cancellation latencies varied across blocks. Participants' choices in gambling provided real-time measures of the updating of temporal expectation over time, which suggest imperfect representation of temporal distributions. Importantly, we found that on a trial when participants predicted an ending of cancellation instead of target, their subsequent response to the target was strikingly slower (up to 1/3 increase in response time). It implies temporal expectation is shared across higher-level and lower-level cognitive tasks.

Keywords: Temporal expectation; Probability learning; Temporal distribution; Dual task; Ideal observer

Introduction

We often wait for events whose occurrence is uncertain but increasingly predictable. We update our temporal expectations from time to time and change our behaviors accordingly. For example, as a procrastinating academic, you are desperately waiting for the extension of a conference deadline. However, when it comes to the last week and you realize that an extension is improbable, you start to wake up early and stay up late, trying to squeeze out every minute to write.

Temporal expectation works on many different time scales (McGuire & Kable, 2013). On the time scale of hundreds of milliseconds to seconds, it has long been found that people can perceive more accurately or respond faster when they learn the latency of the target event in advance (Niemi & Näätänen, 1981; Rohenkohl et al., 2012). In perceptual and motor tasks, temporal expectation also influences the allocation of attention (Ghose & Maunsell, 2002) or motor preparation (Cui et al., 2009) over time, and modulates other time-dependent behavioral phenomena such as priming effects (Wang et al., 2020). In light of psychophysical and neurophysiological evidence that sensory and motor timing on the scale of hundreds of milliseconds is largely task-specific (Merchant et al., 2013), several theories propose that the effects of temporal expectation in lower-level cognitive tasks may be emergent properties from task-specific, low-level circuits that involve no abstraction of temporal information (Burr et al., 2007; Dragoi et al., 2003; Jepma et al., 2012; Machado, 1997).

However, it is largely unknown whether the temporal expectation implicit in lower-level cognition (i.e., sensorimotor tasks) is really task-specific, or is the same as that used in higher-level cognition. In other words, there are two alternative hypotheses: global temporal expectation hypothesis and local temporal expectation hypothesis, differing in whether temporal expectation is shared across lower- and higher-level cognition.

To test between these two hypotheses, we designed a dual task where two tasks—a simple response task and a novel real-time gambling task—are based on the same temporal structure. In the simple response task, also known as foreperiod task in the literature (Niemi & Näätänen, 1981), participants make speeded response to a visual target presented after varying delays, whose response time (RT) provides a conventional but limited measure of temporal expectation. On some trials the target may be cancelled.

When participants are waiting for the target, they also need to predict from time to time whether the target will be present or absent and will receive reward for correct prediction or penalty for incorrect prediction. This real-time gambling task provides a more direct and continuous measure of temporal expectations at different moments.

The simple response task is considered as a lower-level cognitive task, for which temporal expectation is not explicitly required and can in theory only function in task-specific local circuits. In contrast, the gambling task is a higher-level cognitive task, which would benefit from explicit expectation of how the probabilities of occurrence of different events change with time. If a global temporal expectation is shared across tasks, the performances of the two tasks should covary; otherwise, they should be independent of each other.

We collected data from four Web-based experiments (371 participants in total) that had the same temporal structure and simple response task but varied in the settings of the real-time gambling task. Our goal here was two-fold. First,

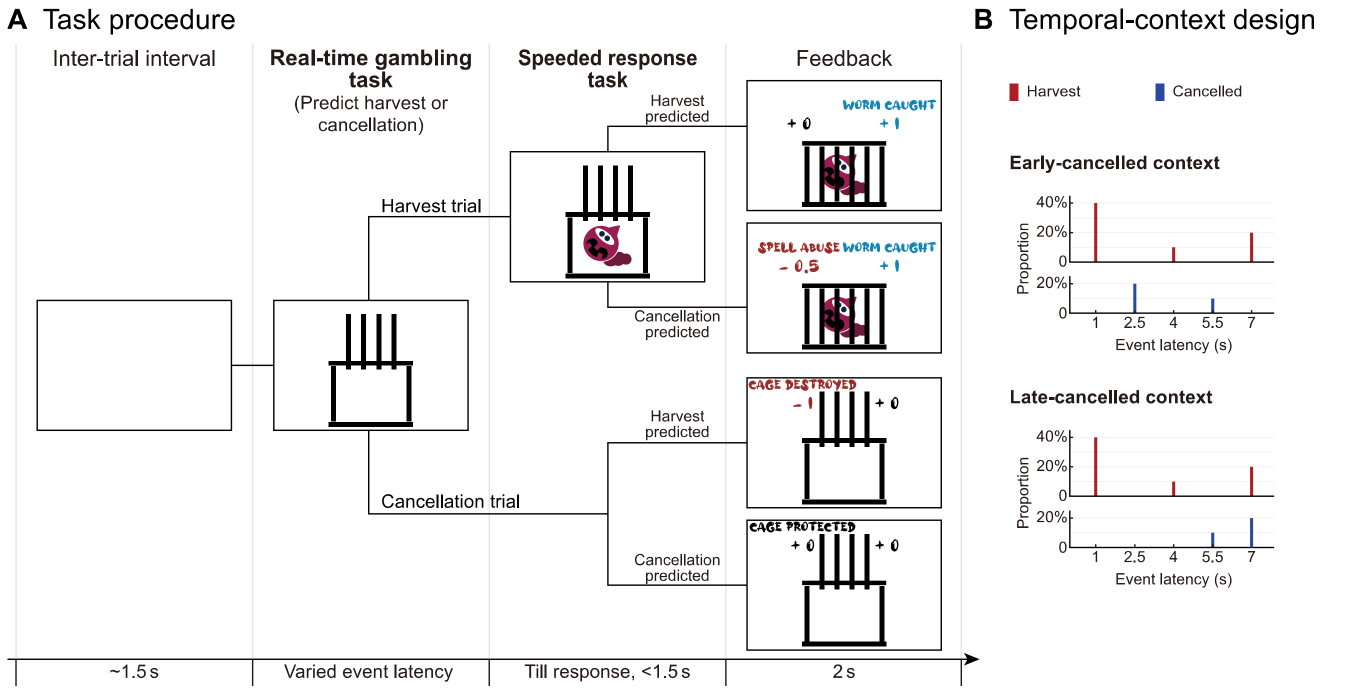

Figure 1: Task procedure and temporal-context design. (**A**) For speeded response task, the participants were asked to capture worms by pressing J on keyboard as quick as possible when a worm appears in harvest trials. However, the worm would not appear in cancellation trials. And the participants did not which type of trial they were in. The participants need to gamble on whether a worm will appear before event onset. (**B**) The same two temporal contexts were used in all experiments. Red bars denote the trial proportion of harvest trial; blue bars denote the proportion of cancellation trials. The event latency of cancellation trials in early-cancelled context is more likely to be short than that in late-cancelled context.

participants' predictions in the real-time gambling task would allow us to measure their temporal expectations on a real-time basis. Second, the relationship between participants' predictions in the real-time gambling task and their RTs in the simple response task would allow us to distinguish between the global and the local hypotheses of temporal expectation. The results of our four experiments provided converging evidence for shared temporal expectation across higher- and lower-level cognition: when participants predicted that the trial would end up with cancellation, their subsequent response to the target was strikingly slower (up to 1/3 increase in RT).

Methods

Participants

A total of 371 English-speaking participants (aged 18 to 40, 240 male, 125 females, 6 unknown) were recruited on the UK-based online platform Prolific. Participants had run our experiments on computers or laptops. Among them, 40 participants were excluded from further analysis due to high invalidity rate (>5% premature and time-out responses) in the speeded response task. Our study had been approved by the ethics committee of School of Psychological and Cognitive Sciences at Peking University. All participants provided informed consent prior to experiments. Participants received a basic payment of 4.5 GBP and a performance-dependent bonus (mean of valid participants: 1.22 GBP).

Stimulus presentation and data recording

Experiments were implemented in PsychoPy, hosted on Pavlovia.org and run on Web browsers. The standard deviation of timing for visual duration and reaction time under this setting is under 5 ms (Bridges et al., 2020).

We used a game-like interactive tutorial to provide task instructions to participants, whose answers to post-experiment questionnaires as well as whose performances during the experiment showed that they had well understood the instructions.

Task procedure

According to our cover story, participants were asked to catch worms using magic cages (Figure 1A). On each trial, an open cage was presented for 1–7 s until being filled with a worm (if in *harvest trial*) or vanishing (if in *cancellation trial*). Participants needed to perform dual tasks, both of which provided performance-dependent bonus points that were additive.

Speeded response task. In a harvest trial, a worm would appear in the cage and participants were instructed to close the cage by pressing the J key on keyboard as quick as possible. Their RT was recorded. Each valid response would result in a reward of +1 point. Time-out response (> 2 s) would incur a penalty of -1 point. No speeded response would be needed for a cancellation trial, where the cage vanished without a worm.

Real-time gambling task. During the open-cage stage, whether the trial would be a harvest or cancellation trial was unknown to participants and they needed to decide whether to protect the cage with a protective spell. The choices of to protect and not to protect the cage would be favorable respectively for cancellation and harvest trials and will be respectively referred to as *predicting cancellation* and *predicting harvest*, for simplicity. In particular, there were four consequences: harvest trial with harvest predicted, 0 point; harvest trial with cancellation predicted, –0.5 point; cancellation trial with harvest predicted, –1 point; cancellation with cancellation predicted, 0 point. Participants were explicitly informed of this payoff matrix before the experiment. The penalty ratio of incorrectly predicting harvest to incorrectly predicting cancellation was close to the proportion of harvest to cancellation trials (70% to 30%, see below), so that overall the two choices were almost equally profitable.

Decision types and action mapping

In different experiments, there were two decision types: participants made the prediction of harvest or cancellation either for one time or continuously during the open-cage stage. The four experiments also differed in whether pressing key was mapped to predicting harvest or cancellation. Decisions and their timing were recorded.

One-time harvest (OneHarv) experiment. Participants may press the F key on keyboard once a trial at any time during the open-stage stage to choose predicting harvest. No key press was meant to choose predicting cancellation.

Continuous harvest (ContHarv) experiment. Participants may switch their choices between predicting harvest (by holding down F) and predicting cancellation (by releasing F) as many times and at any time as they wish. Their payoff was determined by their final choice at the onset of the ending event (i.e., worm onset).

One-time cancellation (OneCanc) experiment. The same as one-time harvest experiment, except that pressing F was for predicting cancellation and no key press for harvest.

Continuous cancellation (ContCanc) experiment. The same as continuous harvest experiment, except that holding down F was for predicting cancellation and releasing F for harvest.

Temporal-context design

The same two temporal contexts (Figure 1B) were used in all four experiments. Each block consisted of 35 harvest trials and 15 cancellation trials (i.e., 70% and 30%). In harvest trials, the latency for worm appearing could be 1 s, 4 s, or 7 s, whose trial numbers followed the ratio 4:1:2. In cancellation trials, the latency for cage vanishing could be 2.5 s or 5.5 s for the *early-cancelled context* and 7 s or 5.5 s for the *late-cancelled context*, whose trial numbers followed the ratio 2:1. Harvest and cancellation trials with different event latencies were randomly mixed in each block. The temporal distributions of harvest and cancellation trials were chosen in such a way that the relative expected gain of predicting harvest versus cancellation would change with time to a considerable extent. Each participant completed four blocks, with two blocks for each temporal context. They were randomly assigned to the order of E(arly)L(ate)EL or LELE. There was a 30-second break between two blocks. Participants were not explicitly informed about the temporal structure of the experiment.

Ideal observer analysis

To evaluate participants' choices in real-time gambling, we modeled an ideal observer who knows the temporal structure of each temporal context (i.e., joint probability distribution of trial type and event latency, as we describe above) but whose time perception is corrupted by a Gaussian noise following Weber's law (Jazayeri & Shadlen, 2010).

At time t of a trial, define $\varphi_t(\tau',\tau)$ as a Gaussian kernel centered at τ with standard deviation kt (k as Weber fraction):

$$\varphi_t(\tau',\tau) = \frac{1}{\sqrt{2\pi}kt}\exp\left(-\frac{(\tau'-\tau)^2}{2(kt)^2}\right). \quad (1)$$

In continuous decision experiments where participants could change their choice at any time as they wish until the end of the trial, the ideal observer needs only to consider the outcome distribution at the next moment, i.e., given that the trial ends immediately after t. According to Bayes rule, the ideal observer's posterior estimate at time t for the conditional probability that the trial will be trial type T (harvest or cancellation) if it ends at time t is

$$p(T|t, t_{\text{end}} \geq t) = \frac{\Sigma_{i_T} q_{T,i_T} \int_t^\infty \varphi_t(\tau, \tau_{T,i_T}) d\tau}{\Sigma_T \Sigma_{i_T} q_{T,i_T} \int_t^\infty \varphi_t(\tau, \tau_{T,i_T}) d\tau}, \quad (2)$$

where subscript i_T indexes event latency of T-type, q_{T,i_T} and τ_{T,i_T} respectively denote the proportion and event latency of T-type trials with the i_T-th latency in the temporal context.

In one-time decision experiments where participants can only choose once, the ideal observer needs to consider the outcome distribution in all moments after t, thus computing

$$p(T|t, t_{end} \geq t) = \frac{\Sigma_{i_T} q_{T,i_T} \int_t^\infty \varphi_t(\tau, \tau_{T,i_T}) d\tau}{\Sigma_T \Sigma_{i_T} q_{T,i_T} \int_t^\infty \varphi_t(\tau, \tau_{T,i_T}) d\tau}. \quad (3)$$

Let $p_{\text{end}}(T|t) = p(T|t, t_{\text{end}} = t)$ for continuous decision experiments and $p_{\text{end}}(T|t) = p(T|t, t_{\text{end}} \geq t)$ for one-time decision experiments. The expected value of a specific choice (predicting harvest or predicting cancellation, denoted C) at time t is thus

$$EV(C|t) = \sum_{T} r(C|\mathrm{T}) p_{\mathrm{end}}(\mathrm{T}|t), \tag{4}$$

where $r(C|T)$ denotes the payoff of choice C given trial type T (see Figure 1A or see Task Procedure above for the payoff matrix). The ideal observer will choose the option that maximizes expected value.

Statistical analysis

For the simple response task, the mean and standard error of RT were calculated in logarithmic scale and presented in linear scale. For the real-time gambling task, predicting harvest and cancellation at each specific time point were respectively coded as status values 0 and 1. The probability of predicting cancellation at a specific time point was defined as the mean status value across trials for all trials that ended after the time point.

Linear mixed-effects model (LMEM). We used LMEMs with Satterthwaite's approximate F test and Wald test for group-level statistical conclusions, where the random effects of LMEMs followed keep-it-maximal rule (Barr et al., 2013). To improve convergence and avoid singular fits, we identified and trimmed redundant random effects based on principal component analysis and the fitting of zero-correlation models. RTs were transformed to log scale to improve normality. Decisions (predicting harvest or cancellation) were treated as binomially distributed random variables when entered as dependent variable in generalized LMEMs. All LMEMs and tests were implemented with R packages *lmer* and *car*.

We applied Benjamini-Hochberg correction to multiple comparison p values to control the false discovery rate.

Results

Real-time behavioral measures of temporal expectation

During the open cage stage of each trial, an ideal observer would keep updating her probabilistic estimate for the incoming event and change her choice accordingly (see Methods). Consider the ideal observer's preference in the continuous decision experiments under the early-cancelled context (Figure 2B, green curve) as an illustrating example. At the beginning of the trial, the ideal observer prefers predicting harvest, because the first event on the time line is harvest at 1 s (Figure 1B). But when it passes 1.5 s and no event occurs yet, the ideal observer starts to prefer predicting cancellation, which corresponds to the next event (2.5 s) on the time line. That is, the temporal expectation of the ideal observer at each moment determines her choice in a real-time manner. In other words, if we assume human participants also seek to maximize expected reward, we can use their choices in the real-time gambling task as real-time measures of temporal expectation.

A comparison of participants' choices with those of the ideal observer allows us to see in what aspects participants' temporal expectation agrees with and in what aspects it

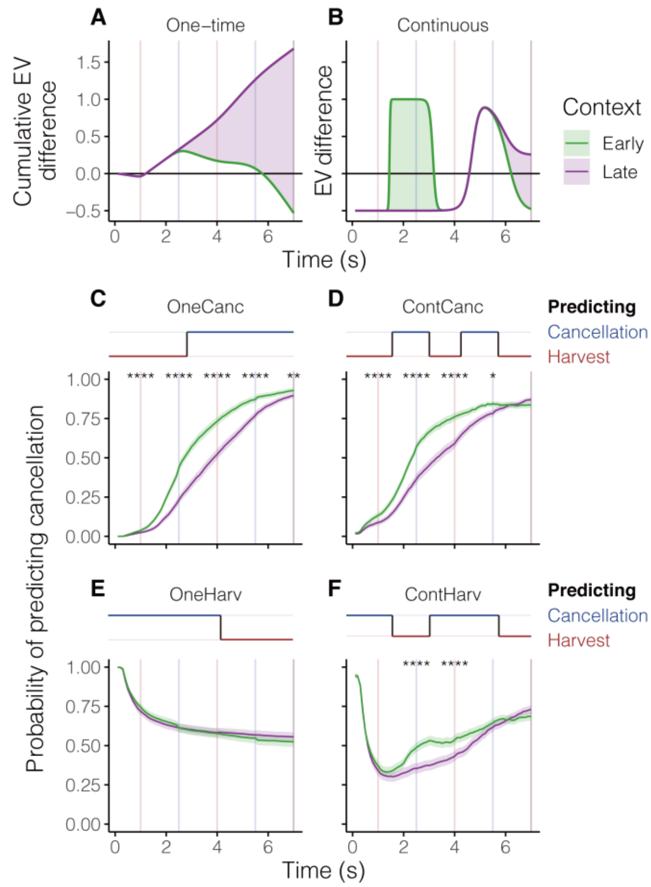

Figure 2: Temporal dynamics of gambling for an ideal observer and participants. (Cumulative) difference of expected value between two decision options in one-time gamble tasks (**A**) and continuous gamble tasks (**B**). See Ideal observer analysis for detail. Participants linear mixed-effects model. Green denotes early-cancelled context; purple denotes late-cancelled context. Shades indicates difference between two contexts. The color of the shades at each time poing denotes which context has higher (cumulative) EV difference favoring cancellation. (**C**-**F**) Probability of predicting cancellation was the average of real-time prediction at each time point (0, predicting harvest; 1, predicting cancellation). Upper panels, examples of status value change during a trial. The discrepancy between the EV difference curves and participants' prediction curves indicates that the participants seem to have inaccurate representation of temporal distributions. Red vertical lines indicates the event latencies of harvest trials (1, 4, 7 s). Blue vertical lines denotes the latencies of cancellation trials (2.5, 5.5, 7 s). Asterisks at five event latencies denote the significance of difference between contexts in LMEM analysis. ****, $p < 0.0001$; **, $p < 0.01$; *, $p < 0.05$. P values were corrected for multiple comparison.

deviates from the ground truth. For continuous decision experiments, an ideal observer should predict cancellation when the expected value (EV) difference is greater than 0 and predict harvest when the difference is less than 0. The ideal observer's preferences for the two temporal contexts differ around 2.5 s and 7 s (indicated by colored shadings in Figure 2B), but are almost identical for the rest of the time. In contrast, on one hand, participants' choices showed significant across-context differences at 2.5 s in the correct direction (Figure 2D & F, significance marked by asterisks). On the other hand, their choices also showed significant differences at 4 s and lack of significant differences at 7 s, deviating from those of the ideal observer who has perfect knowledge of the generative temporal distributions.

A remarkable bias in participants' choices was their increasing probability of predicting cancellation up to the end of the trial (Figure 2D & F; For participants' choice after 1s, ContCanc, $\chi^2(1) = 222.29$, $p < 0.0001$; ContHarv, $\chi^2(1) = 76.88$, $p < 0.0001$) despite that cancellation would be impossible after 5.5 s in the early-cancelled context.

In one-time decision experiments, participants made at most one key press per trial and we would not know whether their prediction ever changed after the key press. However, when we described participants' choices on each trial as a step function, as if they stuck to the choice at the time of key press, the resulting curves (Figure 2C & E) surprisingly resembled their counterparts in continuous decision experiments (Figure 2D & F) and disagreed with the ideal observer's preferences in one-time decision experiments. It suggests that even when having no opportunity to choose again, participants did not consider all possible events in the future but only focused on the next moment.

RTs of speeded responses

We then examined the RTs of the speeded response task, the more conventional part of our dual-task experiments, and performed LMEM analysis on RTs to test the effects of event latency and temporal context, separately for the four experiments. The RTs observed in our experiments (Figure 3A) resembled previous simple-response studies with varied event latency (Luce, 1986), no matter in the mean response time or in the decrease of RT with increasing event latency (1, 4, 7 s), which was significant in three out of four experiments (OneCanc, $F(1, 83.83) = 68.38$, $p < 0.0001$; OneHarv, $F(1, 83.94) = 101.48$, $p < 0.0001$; ContCanc, $F(1, 82.01) = 0.42$, $p = 0.52$; ContHarv, $F(1, 80.93) = 25.50$, $p < 0.0001$).

The RTs were also significantly influenced by temporal context (early-cancelled vs. late-cancelled), according to the main effect of temporal context (OneCanc, $F(1, 83.22) = 25.61$, $p < 0.0001$; OneHarv, $F(1, 83.83) = 1.87$, $p = 0.175$; ContCanc, $F(1, 81.77) = 6.58$, $p = 0.012$; ContHarv, $F(1, 81.02) = 6.58$, $p = 0.012$) or the interaction between event latency and temporal context (OneCanc, $F(1, 83.70) = 3.43$, $p = 0.067$; OneHarv, $F(1, 11206.72) = 11.49$, $p < 0.001$; ContCanc, $F(1, 81.74) = 1.17$, $p = 0.28$; ContHarv, $F(1, 80.89) = 1.57$, $p = 0.21$). That RTs were smaller in the early-cancelled than in the late-cancelled context might be due to the shorter average trial duration of the former than the latter, which is consistent with the predictions of the temporal

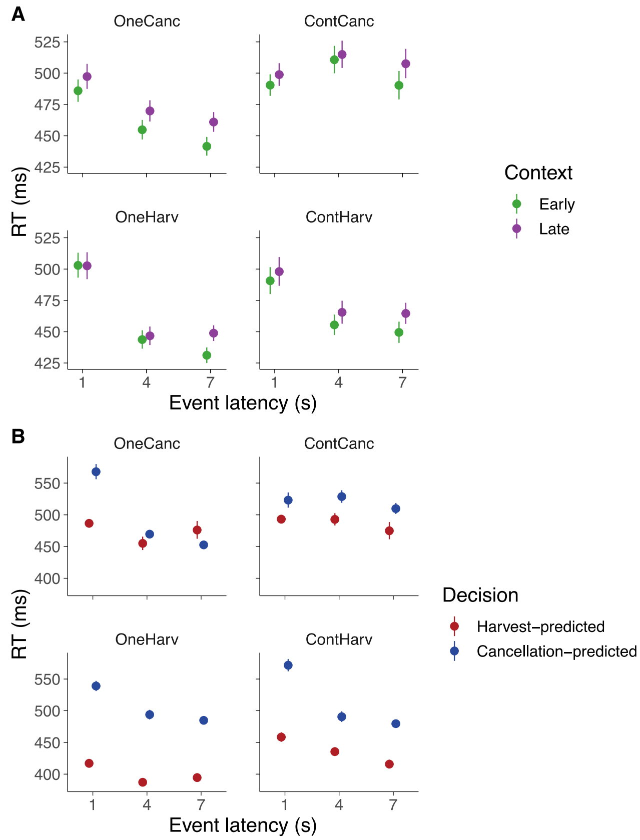

Figure 3: RT patterns under different temporal contexts (**A**) and predictions (**B**). (**A**) Response time decreased with event latency and differed between context, which was coherent with previous studies (e.g., Niemi & Näätänen, 1981). Green dots denote RT under early-cancelled context; purple dots denote late-cancelled context. Error bars denote standard error of mean. (**B**) Participants responded faster when they predicted harvest trials than when they predicted cancellation trials. The means and standard errors were calculated in logarithmic scale and plotted in linear scale. Red dots denote RT when participants predict a harvest trial; blue dots denote RT when participants predict a cancellation trial. Some error bars are too short to be seen.

discounting theory of motor control (Shadmehr, de Xivry, Xu-Wilson, Shih, 2010).

Evidence for shared temporal expectation across real-time gambling and speeded response

Participants confronted the same time structure in the dual tasks of real-time gambling and speeded response. Did the temporal expectation revealed by the gambling task also influence participants' response time on the same trial?

We split harvest trials into two categories, harvest-predicted vs. cancellation-predicted, depending on whether harvest or cancellation had been predicted in gambling. For continuous decision experiments, this classification was based on the decision immediately before worm onset.

We found that the RTs of cancellation-predicted trials were significantly longer than those of harvest-predicted trials (Figure 3B), for all four experiments (main effect of decision, OneCanc, $F(1, 69.17) = 28.86$, $p < 0.0001$; OneHarv, $F(1, 80.80) = 441.14$, $p < 0.0001$; ContCanc, $F(1, 82.42) = 26.89$, $p < 0.0001$; ContHarv, $F(1, 78.88) = 230.26$, $p < 0.0001$). The effect size of this effect was striking: in some experiments, cancellation-predicted RTs were longer than harvest-predicted RTs by 150 ms, up to 1/3 of the latter.

Furthermore, the robustness of this effect in the four experiments with different decision types and action mappings allowed us to exclude many trivial explanations. For example, if the effect had been a motor artefact due to facilitation or inhibition of previous key press to later motor responses, the direction of the effect should have been opposite for experiments where action mappings were opposite (i.e., OneHarv and ContHarv vs. OneCanc and ContCanc).

Though it is unknown whether the effect reflects an influence from gambling to speeded response, or the reverse, or a common cause for both, we may safely conclude that temporal expectation is shared, at least partly, across the two tasks.

Discussion

In the present study, we introduce a real-time gambling task to provide real-time behavioral measures for temporal expectation, through which we identify patterned deviations of participants' temporal expectation from ground truth. We also combine real-time gambling with a traditional simple response task in the same trial to investigate whether participants' temporal expectations are task-specific or shared across tasks. We find converging evidence in four experiments for shared temporal expectation across the two tasks.

The real-time gambling and speeded response tasks can be considered as representative tasks respectively of higher- and lower-level cognition that involve temporal expectation. Though temporal expectation has been extensively studied in both higher-level cognition, such as value-based decisions (Frederick et al., 2002; McGuire & Kable, 2013), and lower-level cognition, such as perceptual and motor tasks (Cui et al.,2009; Jepma et al., 2012), these two lines of research in the literature almost never cross, with a few exceptions (Shadmehr et al., 2010; Wang et al., 2020). It is probably because temporal expectations in higher- and lower-level cognition look so different: The former is often explicit, in a time scale ranging from minutes to years, while the latter is more implicit, ranging from milliseconds to seconds. It is even hardly realized that the two should be studied in a general framework.

However, there are reasons to believe the two are actually not so different as they look to be. On one hand, task-specific mechanisms proposed for low-level tasks to substitute for an abstract representation of temporal expectation, such as motor adaptation, can only explain some but not all temporal context effects in simple response tasks (Los & Agter,2005). On the other hand, even animal conditioning studies suggest a more sophisticated representation of temporal expectation than that could be approximated by simple rules (Starkweather et al., 2017).

Indeed, here we find that temporal expectation may be shared across higher- and lower-level cognition. It should be noted that shared and task-specific temporal expectations are not necessarily mutually exclusive, given that timing is implemented in the brain by a main core interacting with multiple task-specific brain regions (Merchant et al., 2013). What behaviors are guided by shared and what by task-specific temporal expectations, how temporal expectations learned in one task may generalize to a second task, and whether the biases we observed in the real-time gambling task apply to all tasks, are all interesting questions for future research.

Our results also shed light on how well people can learn temporal distributions from experience. Previous studies on how temporal expectation influences perceptual and motor processes often (implicitly) assumed that the participants have accurate representations of temporal distributions (Niemi & Näätänen, 1981; Rohenkohl et al., 2012), though there was little empirical evidence for this assumption. Instead, the considerable biases participants showed in our real-time gambling task suggest that people may not able to acquire accurate representations of temporal distributions, at least not through hundreds of trials.

Meanwhile, the real-time gambling task we develop allow direct and continuous measures of participants' temporal expectations. It can be further combined with pupillometry, microsaccades, and neuroimaging measures to deepen our understanding of human cognition.

Acknowledgments

This work was partly supported by grant 31871101 from National Natural Science Foundation of China and funding from Peking-Tsinghua Center for Life Sciences to H.Z.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ, 8*, e9414.

Burr, D., Tozzi, A., & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience*, *10*(4), 423–425.

Cui, X., Stetson, C., Montague, P. R., & Eagleman, D. M. (2009). Ready…Go: Amplitude of the fMRI Signal Encodes Expectation of Cue Arrival Time. *PLOS Biology*, *7*(8), e1000167.

Dragoi, V., Staddon, J. E. R., Palmer, R. G., & Buhusi, C. V. (2003). Interval timing as an emergent learning property. *Psychological Review*, *110*(1), 126–144.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, *40*(2), 351–401.

Ghose, G. M., & Maunsell, J. H. R. (2002). Attentional modulation in visual cortex depends on task timing. *Nature*, *419*(6907), 616–620.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026.

Jepma, M., Wagenmakers, E.-J., & Nieuwenhuis, S. (2012). Temporal expectation and information processing: A model-based analysis. *Cognition*, *122*(3), 426–441.

Los, S. A., & Agter, F. (2005). Reweighting sequential effects across different distributions of foreperiods: Segregating elementary contributions to nonspecific preparation. *Perception & Psychophysics*, *67*(7), 1161–1170.

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review*, *104*(2), 241–265.

McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, *120*(2), 395–410.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural Basis of the Perception and Estimation of Time. *Annual Review of Neuroscience*, *36*(1), 313–336.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162.

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience*, *32*(24), 8424–8428.

Shadmehr, R., Orban de Xivry, J. J., Xu-Wilson, M., & Shih, T.-Y. (2010). Temporal Discounting of Reward and the Cost of Time in Motor Control. *Journal of Neuroscience*, *30*(31), 10507–10516.

Starkweather, C. K., Babayan, B. M., Uchida, N., & Gershman, S. J. (2017). Dopamine reward prediction errors reflect hidden-state inference across time. *Nature Neuroscience*, *20*(4), 581–589.

Wang, M., Huang, Y., Luo, H., & Zhang, H. (2020). Sustained Visual Priming Effects Can Emerge from Attentional Oscillation and Temporal Expectation. *Journal of Neuroscience*, *40*(18), 3657–3674.