Two sources of uncertainty independently modulate temporal expectancy

Matthias Grabenhorst[a,1], Laurence T. Maloney[b], David Poeppel[a,b,c,2], and Georgios Michalareas[a,2]

[a]Department of Neuroscience, Max-Planck-Institute for Empirical Aesthetics, Frankfurt 60322, Germany; [b]Department of Psychology, New York University, New York, NY 10003; and [c]Max-Planck–NYU Center for Language, Music, and Emotion, New York University, New York, NY 10003

Edited by Charles R. Gallistel, Rutgers, The State University of New Jersey, Piscataway, NJ, and approved March 15, 2021 (received for review September 14, 2020)

The environment is shaped by two sources of temporal uncertainty: the discrete probability of whether an event will occur and—if it does—the continuous probability of when it will happen. These two types of uncertainty are fundamental to every form of anticipatory behavior including learning, decision-making, and motor planning. It remains unknown how the brain models the two uncertainty parameters and how they interact in anticipation. It is commonly assumed that the discrete probability of whether an event will occur has a fixed effect on event expectancy over time. In contrast, we first demonstrate that this pattern is highly dynamic and monotonically increases across time. Intriguingly, this behavior is independent of the continuous probability of when an event will occur. The effect of this continuous probability on anticipation is commonly proposed to be driven by the hazard rate (HR) of events. We next show that the HR fails to account for behavior and propose a model of event expectancy based on the probability density function of events. Our results hold for both vision and audition, suggesting independence of the representation of the two uncertainties from sensory input modality. These findings enrich the understanding of fundamental anticipatory processes and have provocative implications for many aspects of behavior and its neural underpinnings.

anticipation | catch trial | probability learning | temporal–probabilistic inference | reaction time

A boxer circles her opponent, prepared to respond in fractions of a second, anticipating and blocking the next attack—if there is one. A commuter wonders for minutes when the train will arrive—knowing it might have been cancelled. A leopard lies in ambush for hours, ready to seize its prey at the water hole—if it comes to drink. A stockbroker follows the market over days and weeks trying to anticipate the right moment to sell—which may never come.

These examples illustrate, over a range of time scales, a fundamental challenge in anticipatory decision-making: How does the brain predict events which are distributed in time and which may or may not occur? This question pertains to two different types of uncertainty that frequently cooccur in real-world stochasticity, namely, a discrete type of uncertainty about whether an event will happen (Bernoulli) and a continuous type about when it will happen (continuous variation in probability density as a function of a continuously elapsing interval).

The discrete and perhaps most fundamental source of uncertainty is the probability P_O that an event E will occur at all. In case the event occurs, there is uncertainty about when exactly it will happen. This source of uncertainty can be summarized as a probability density function (PDF), γ_E, which is defined such that

$$Pr(E \in [a,b] | E) = \int_a^b \gamma_E(t)\, dt. \quad [1]$$

Here, t denotes time, and $Pr(E \in [a,b]|E)$ is the probability that the event E will happen within the time interval $[a,b]$, conditional on its occurrence.

Little experimental work investigates the interaction between these two types of uncertainty, P_O and γ_E, and the underlying mechanisms remain largely unknown. Importantly, both sources of uncertainty are frequent elements of experimental protocols, and their hypothesized impact on behavior varies remarkably across different fields of research.

In cognitive psychology and neuroscience, the continuous uncertainty, γ_E, is commonly investigated in variable–foreperiod designs, wherein stimulus appearance follows probability distributions across time (1–3). In more traditional switch designs, the foreperiods vary among a small number of time spans. Classic experiments demonstrated a monotonic decrease in reaction time (RT) in the case of a uniform foreperiod distribution (4) or RT modulation by interval variability (5). The discrete uncertainty, P_O, is implemented in these protocols by the use of catch trials (omission of a target stimulus, switch design) (4, 6) or the presentation of a catch stimulus ("no-go" trials) (7). The uncertainty of event occurrence is intended to reduce the frequency of early responses, to avoid automatic responses, or to maintain participants' alertness (4); however, potential effects of this manipulation are seldom addressed in data analysis.

In associative learning, the discrete uncertainty is addressed in models of behavior. Many Pavlovian and operant protocols investigate the effect on learning of events that are expected but fail to occur (8). Prominent models of behavior from this field, such as the Rescorla–Wagner model (9) from which many contemporary reinforcement learning models derive, treat such stimulus absence as an event in itself, a failure of reinforcement, which affects the associative learning strength (8). In trace conditioning, the interval between conditioned and unconditioned stimuli introduces the continuous uncertainty to these experimental designs (10).

In decision-making, both the discrete uncertainty (e.g., the probability of reward) (11, 12) and the continuous uncertainty (e.g.,

Significance

Biological organisms seek to reduce the uncertainty surrounding future events. Every possible event carries two distinct kinds of uncertainty: Will the event happen? And if it will happen, when exactly? It is commonly believed that the probability of whether an event occurs has a static effect on expectancy. However, here, we demonstrate that the effect is highly dynamic across time. We further show that the uncertainties about whether and when an event will occur independently shape anticipatory behavior. The results deepen our understanding of how the human brain interacts with the temporal structure of its environment.

Author contributions: M.G., D.P., and G.M. designed research; M.G. performed research; M.G. and G.M. analyzed data; and M.G., L.T.M., D.P., and G.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

[1]To whom correspondence may be addressed. Email: m.g@ae.mpg.de.

[2]D.P. and G.M. contributed equally to this work.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2019342118/-/DCSupplemental.

Published April 14, 2021.

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

a delayed reward) (13, 14) are considered critical to the calculation of value. Specifically, midbrain dopamine neurons adapt both to discrete (12) and continuous (14) uncertainty. Interestingly, in the reward literature, these two sources of uncertainty are typically not combined in single experiments, and conclusions about potential interactions cannot be drawn. Nonetheless, the effect of event uncertainty in reward anticipation (15, 16), and the certainty effects in decision-making (17, 18), invite the hypothesis that the impact of P_O on event anticipation might be dynamic rather than static (Fig. 1*A*).

In work on temporal anticipation, the deceptively simple manipulation of P_O is commonplace, through the use of catch trials, and it introduces uncertainty about the occurrence of events ($P_O < 1$) (2, 7, 19, 20). Likewise, the absence of catch trials in temporal anticipation (1, 21–23) results in event certainty ($P_O = 1$). In either case, the influence of P_O on anticipation introduces an unavoidable confound in experimental designs. However, the potential effects of P_O on adaptive behavior and its neural correlates are often not addressed or are assumed to be of fixed nature, uniform across time. Therefore, despite the large body of work on temporal anticipation (1–3, 7, 19–25), the specific nature of the effect of P_O on anticipatory processes remains unknown.

Unsurprisingly, very little is known about the interaction between the discrete and continuous uncertainties. One theoretically driven attempt to study the interaction proposed that the brain maintains a simultaneous representation of two complementary states: 1) the probability density of an event happening and 2) the probability density of an event not happening, thus combining both sources of uncertainty (26). This study was specifically tailored to associative learning, and the findings may not readily generalize to temporal anticipation.

The different examples from action, learning, and decision-making illustrate a common challenge: probability estimation of whether an event happens and when it happens. These two fundamental types of uncertainty and their interaction affect any task that has an anticipatory element. Hence, it is of broad significance to understand how the brain models each one of them and combines them.

Regarding the relation of RT to γ_E, it has typically been hypothesized that RT is inversely proportional to the hazard rate (HR) (1–4). The HR, $h(t)$, of an event happening at a given time t represents the probability density of this event at time t, given that it has not already happened, within the current time span (4).

$$h(t) = \frac{\gamma_E(t)}{\int_t^\infty \gamma_E(u)\, du}. \quad [2]$$

A sharp increase in the HR $h(t)$ at time t suggests the event is imminent. The widely adopted hypothesis that HR drives anticipation was recently challenged, and it was demonstrated that RTs are better captured by the reciprocal of event probability density in the context of a single, fixed P_O (27). Here, both hypotheses (Fig. 1*D*) are tested at various levels of P_O in order to evaluate whether the reciprocal PDF is indeed better than HR as a model of event anticipation (Fig. 1 *E* and *F*).

The reciprocal of probability is an interesting quantity in many aspects. Computationally, it is a simpler and more stable variable than the HR. When estimating event probability within a given time interval, the reciprocal of probability is equivalent to "1 in N" counting: If the probability of an event in this time interval is 0.1, then this can be directly computed by counting that 1 out of 10

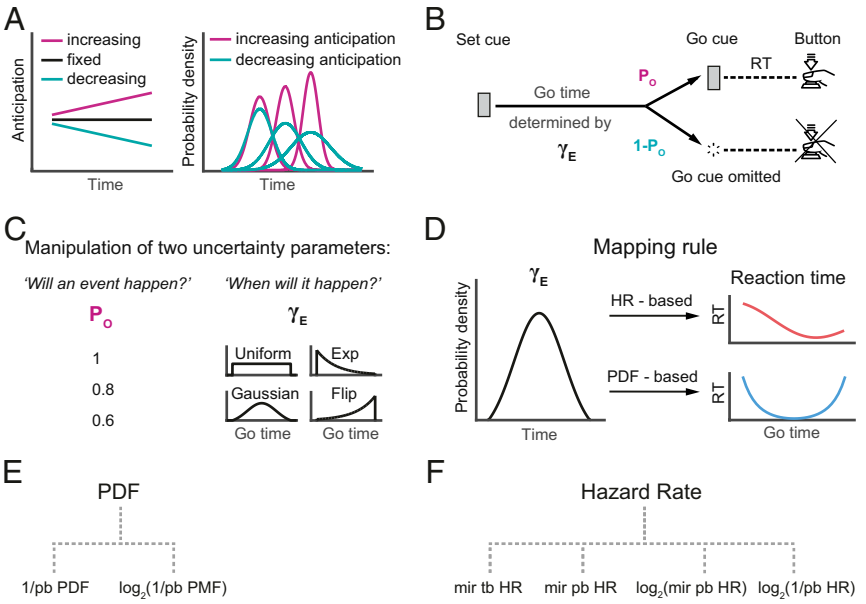

Fig. 1. Hypotheses and experimental design. (*A*) Possible dynamic anticipation effects of discrete uncertainty, P_O (*Left*). Uncertainty in time estimation, represented by dynamic Gaussian kernel, is hypothesized to be driven by P_O (i.e., to scale inversely with the anticipation of whether an event will or will not occur [*Right*]). In the fixed anticipation case (black line in *Left*), Gaussian kernel (not depicted) assumes a fixed SD. (*B*) In "set"–"go" trials, participants are asked to respond as fast as possible to the "go" cue but not to respond in case of its absence. Time between "go" cue onset and button press is RT. (*C*) Blocks of trials differed in overall event probability, which could assume one of three values: $P_O = 1$ ("go" cue on every trial), $P_O = 0.8$ ("go" cue in 80% of trials), and $P_O = 0.6$ ("go" cue in 60% of trials). Parameterization of P_O facilitates investigation of influence of event occurrence probability on event anticipation. Time between the onset of "set" and "go" cues, ("go" time), followed one of four PDFs, γ_E: uniform (mean = 0.92 s ± 0.31 s), Gaussian (mean = 0.9 s ± 0.21 s), exponential (mean = 0.66 s ± 0.27 s), or flipped exponential (mean = 1.18 s ± 0.27 s) (*SI Appendix*, Table S1). Each PDF was conditioned according to P_O in a given block. Block-wise presentation of trials: sensory modality (audition or vision), P_O and γ_E were fixed per block. (*D*) Hypothesized effect of continuous uncertainty, γ_E, on temporal expectancy. Mapping of γ_E to RT is presumed to be based on either HR model or PDF-based model (*SI Appendix, Supplementary Methods*). Note that in the case of symmetric γ_E (e.g., uniform or Gaussian), the PDF-based model predicts a symmetric RT curve. (*E*) Explanatory variables based on reciprocal probabilistically blurred ("pb") PDF (*SI Appendix, Supplementary Methods*). (*F*) Explanatory variables based on HR ("mir": mirrored, "tb": temporally blurred, and pb, *SI Appendix, Supplementary Methods*).

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

events occur in this interval. This simple operation is easily implementable in an elementary neural network (28). Another interesting aspect of the 1/PDF representation is that it is closely related to the concept of information, in the sense of Shannon: The log of the reciprocal of the probability is the surprisal of an event, which is the amount of information it conveys (29). In order to investigate the possibility that the brain estimates the amount of information of an anticipated event, we test models of RT based on Shannon surprisal. The question is not whether the brain can compute the logarithm of input information in order to guide behavior; the Weber–Fechner law (30) and Hick's law (31), as well as evidence from diverse fields such as language processing, in which reading times follow the logarithm of inverse word probability (32), provide evidence for this capacity. We ask whether there is logarithmic "mental scaling" when events are probabilistically distributed within a bounded time interval.

We note that behavioral results comprise a separable set of processes: 1) the perceptual processing of the input, 2) the representation of uncertainties, and 3) the translation of the representation into behavior. Therefore, our experiments were performed in two different modalities (vision and audition), with the aim to tease these processes apart.

The findings make two contributions to the understanding of temporal anticipation. First, we show that the probability of event occurrence, P_O, has a dynamic, monotonically increasing effect on RTs across the entire tested time span, irrespective of the type of event PDF, γ_E. Second, we demonstrate that models of RT based on the reciprocal probability or its log outperform models employing the HR, reinforcing the failure of the HR as a model of anticipation. Critically, our findings are consistent in vision and audition, suggesting that the representations of the two uncertainties, P_O and γ_E, are independent of the input modality.

Results

Design. The experiment investigates how two uncertainty parameters and their interaction affect event anticipation. In a simple set–go experiment, we independently parameterized both the discrete probability of event occurrence, P_O, and the continuous probability density of the event across time, γ_E (Fig. 1*B*). P_O is parameterized by using three levels of occurrence probability (Fig. 1 *C*, *Left*). γ_E is parameterized by using four different event PDFs (Fig. 1 *C*, *Right*). In the set–go task, participants responded as fast as possible to the "go" cue with a button press. A short RT required prediction of whether there will be a "go" cue as well as an estimate about when it will occur, linking the experiment to many everyday tasks that demand a rapid action based on accurate prediction of future events.

A total of 24 participants generated ∼5,700 RTs each. First, we examined the effect of P_O on average RT and its variance: An offset was observed in mean RT across the different levels of P_O, with the shortest RT at $P_O = 1$, where there is no uncertainty about event occurrence, and $P_O = 0.6$ yielding the longest RTs. These findings were consistent across all four different types of γ_E (Fig. 2*A*). A similar monotonic pattern with respect to P_O values, also consistent across all types of γ_E, was observed in the RT interquartile range (IQR) (Fig. 2*B*). Both findings show that RT becomes shorter and less variable as the probability of event occurrence, P_O, increases. These results also demonstrate that the average effect of P_O on RT is consistent, irrespective of the functional form of γ_E, suggesting independence between these two sources of uncertainty. This assumption of independence is statistically supported by the absence of any interactions in ANOVAs performed separately for median RT and IQR of RT (*SI Appendix*, Tables S3–S7).

As expected, γ_E determined the shape of RT curves across time in both vision (Fig. 2*C*) and audition (*SI Appendix*, Fig. S1). Specifically, the RT curves had an inverse relation to γ_E: in which the event PDF assumed high values, RT was short, and vice versa. This inverse relationship is clearly evident for the Gaussian, exponential, and flipped exponential γ_E cases. In the uniform γ_E case, in which probability density is fixed across time, the increase in RT toward the extrema of the "go" time range may result from the uncertainty in time estimation around the "go" time period. This feature of RT modulation will be addressed below. Nonetheless, the overall relation between γ_E and RT was qualitatively preserved for all examined levels of P_O as demonstrated in Fig. 2*C* for all four types of γ_E. The same plots also show that the effect of P_O on RT curves is not fixed but dynamic across time: Toward the end of the "go" time interval, the level of P_O strongly affects RT curves. Note, however, that the curves' offset between levels of P_O covers the entire "go" time range. In order to further examine the dynamics of this gradually changing offset of RT, we used the RT curve from the $P_O = 1$ condition as a reference. This reference RT curve was subtracted from the RTs of each of the other two levels of P_O. The resulting ΔRT curves demonstrate that the gradual effect of P_O increases monotonically across time (Fig. 2*D*).

Although the ΔRT curves appear to approximate linear functions across time, in some cases they show exponential profiles (e.g., in the Gaussian condition). For this reason, as a final step to descriptively assess whether the ΔRT curves are better captured by a linear (Eq. **3**) or an exponential (Eq. **4**) function of "go" time, we fitted two corresponding models.

$$f_{(t_{go})} = a\ t_{go} + c, \quad [3]$$

$$f_{(t_{go})} = a\ e^{b\ t_{go}} + c. \quad [4]$$

Overall, the exponential model (Fig. 2*D* and *SI Appendix*, Fig. S2) fit the data better than the linear one (*SI Appendix*, Fig. S3); median-adjusted $R^2 = 0.69$ for the linear model and 0.79 for the exponential model ($z = -3.31$, $P = 0.001$, Wilcoxon signed rank, *SI Appendix*, Fig. S4). The exponential model implies that, relative to the $P_O = 1$ condition, RT increases nonlinearly over the range of "go" time as P_O decreases to 0.8 and is even more pronounced when P_O decreases to 0.6.

To summarize, we report two results. First, the effect of P_O on temporal expectancy is dynamic and increases exponentially with time—in contrast to the widely held assumption in the literature that it is fixed across time. Second, the effect of P_O on RT was consistent across different forms of γ_E, indicating that both sources of uncertainty are treated as independent parameters of temporal–probabilistic information. The reported results held in all the respective conditions, suggesting independence of sensory input modality (*SI Appendix*, Fig. S5).

Modeling Effects of Continuous Event Uncertainty. We next investigated how the dynamic effects of γ_E and P_O can be combined in a model of RT. The aim was not to develop a process model of the mechanisms by which sensory evidence is accumulated (e.g., drift diffusion models). Rather, the aim was to develop a descriptive model that captures the combined effects of γ_E and P_O on RTs and whose residuals could provide insights about the mechanisms through which the brain models probabilities across time.

Regarding the effect of γ_E, the recently proposed PDF-based model (27) hypothesizes that RT is reciprocally related to the PDF of events, γ_E, and that the uncertainty in elapsed time estimation is largely modulated by event probability (probabilistic blurring, *SI Appendix*, *Supplemental Methods*). This probabilistic blurring uses a Gaussian kernel whose σ scales inversely with γ_E but not with elapsed time itself as suggested by the scalar variability of time estimation (33), which we here refer to as temporal blurring. Note that in both temporal and probabilistic blurring, the effect of the Gaussian kernel, hypothesized to represent the uncertainty in time estimation, is larger at the extrema of the "go" times range (t_{min}, t_{max}) because the Gaussians near the extrema of the range extend beyond the limits of

NEUROSCIENCE

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

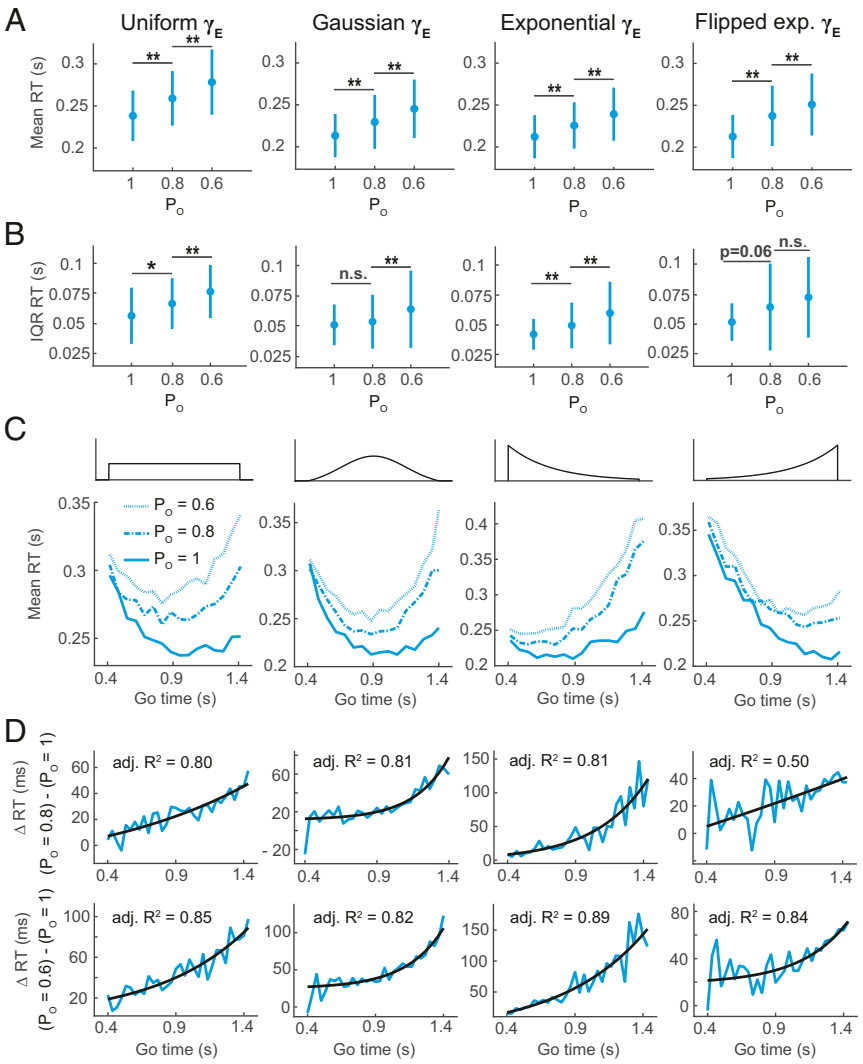

Fig. 2. Modulation of RT by both uncertainty parameters, P_O and γ_E (visual modality). (*A*) Mean RT increases as the probability of event occurrence, P_O, decreases, irrespective of the presented PDF, γ_E (planned contrasts, $^*P < 0.05$, $^{**}P < 0.01$, two-tailed Student's *t* test). (*B*) IQR of RT also increases with a decrease in P_O, resembling the pattern observed in mean RT (planned contrasts, $^*P < 0.05$, $^{**}P < 0.01$, two-tailed Student's *t* test). (*C*) Modulation of RT by P_O and γ_E. In Gaussian, exponential, and flipped exponential conditions, RT was inversely related to γ_E. In the uniform condition, in which the PDF is fixed over time, RT curves bent upwards toward the extrema of the time span. Note that in all γ_E conditions, additionally, P_O modulated RT over the entire range of "go" times. (*D*) ΔRT relative to the $P_O = 1$ condition. Top row: $\Delta RT = RT_{(PO = 0.8)} - RT_{(PO = 1)}$ and bottom row: $\Delta RT = RT_{(PO = 0.6)} - RT_{(PO = 1)}$. The monotonic increase in ΔRT over "go" time was consistent, irrespective of γ_E type. Note that the slopes of ΔRT in the bottom row ($P_O = 0.6$) were higher in all conditions compared to the ones in the top row ($P_O = 0.8$), indicating the dynamic effect of P_O scales with the magnitude of P_O. An exponential function of "go" time (Eq. **4**) captured the effect of P_O on ΔRT (black fit line). In *C*, RT curves were smoothed by reducing the "go" time step size from 32 to 64 ms. For the number of trials per condition, see *SI Appendix*, Table S2. For ANOVAs, see *SI Appendix*, Tables S3–S7. Error bars denote SD. $n = 66{,}279$ RTs.

"go" times (*SI Appendix*, *Supplemental Methods*). This modeling feature proposes that the brain potentially under- and overestimates the limits of the "go" times range. In the case of a uniform γ_E, temporal/probabilistic blurring therefore predicts asymmetric/symmetric U-shaped RT curves, despite the constant value of the uniform γ_E. This model prediction is related to the uncertainty in time estimation, and it differs conceptually from accounts of time perception based on optimality (34) or regression to the mean in a more general way (35). To capture the effect of γ_E, we fitted the prominent mirrored, temporally blurred HR model (1, 3) to RT; it did not capture the data well in any condition (Fig. 1*F* and *SI Appendix*, Fig. S6). The same analysis was repeated for the mirrored, probabilistically blurred HR model, which also did not capture the data (Fig. 1*F* and *SI Appendix*, Fig. S7). In contrast, the models based on the reciprocal, probabilistically blurred PDF fit the data well at all three levels of P_O and in all four γ_E conditions (*SI Appendix*, Fig. S8). This confirms the recent finding that the reciprocal PDF is superior to the HR as a model of RT in anticipation (27) and extends this result to different levels of P_O and the uniform distribution of events in time.

Reciprocal Probability versus Shannon Surprisal. There is a close relation between reciprocal probability and Shannon surprisal, defined as log(1/probability of event), inviting the question of whether the brain quantifies information about event probability by estimating Shannon surprisal. The computation of surprisal implies that the brain performs a logarithmic scaling on probability. This hypothesis was tested by comparing RT models based on the reciprocal probability and Shannon surprisal. As the HR can also be seen as probability scaled by the survival function (HR = PDF/[1 − CDF]), where CDF is the cumulative distribution function, the logarithm of both the mirrored probabilistically

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

blurred HR and the reciprocal, probabilistically blurred HR, was tested as a model of RT (Fig. 1*F*)—in comparison to the logarithm of the event PDF, $\log_2(1/P)$ (Fig. 1*E* and *SI Appendix*, *Supplemental Methods*). The two HR-based models failed to fit the data both qualitatively and quantitatively (*SI Appendix*, Figs. S9 and S10). The models based on the reciprocal probability (Fig. 3*D* and *SI Appendix*, Fig. S8) and on Shannon surprisal (Fig. 3*B* and *SI Appendix*, Fig. S11) captured the effect of γ_E on RT well (Fig. 3*C*), with comparable performance in terms of adjusted R^2 ($z = -1.5$, $P = 0.14$, Wilcoxon signed rank). Nonetheless, there are qualitative differences between the models. In the Gaussian PDF case, surprisal predicts a more linear increase/decrease in RT toward the flanks of the distribution (Fig. 3*B*) than the reciprocal PDF (Fig. 3*D*). Similar qualitative differences are also seen in the exponential and flipped exponential conditions (*SI Appendix*, Figs. S8 and S11; for a detailed overview of differences in model predictions, see plots of explanatory variables in Fig. 3*A* and *SI Appendix*, Fig. S12). Based on this evidence, one cannot conclude that either model is better. Consequently, we cannot confirm or reject the hypothesis that the brain is computing Shannon surprisal by logarithmically scaling reciprocal probability.

Modeling Effects of Discrete Event Uncertainty. Next, a model was developed of the dynamic effect of P_O on RT across time. This model augments the model of the effect of the continuous uncertainty on RT. For the latter, we chose the model based on the reciprocal PDF, as its performance is similar to the one based on surprisal but computationally simpler. The best fit of the reciprocal, probabilistically blurred PDF-based model to RT was for $P_O = 0.8$, providing an adequate model of RT modulation in all four γ_E conditions for both vision and audition (Fig. 3*D* and *SI Appendix*, Fig. S8). The better qualitative and quantitative suitability of the PDF-based model for $P_O = 0.8$ is demonstrated in the cases of the two symmetric γ_E types (Gaussian and uniform conditions), in which the RT patterns largely reflected the symmetry of the input event PDF, γ_E (*SI Appendix*, Fig. S13 and Table S8). The privileged level of $P_O = 0.8$ was further demonstrated by the higher average-adjusted R^2 across all γ_E cases (Fig. 3*E*) and by the fact that the model residuals' slope did not significantly differ from zero (Fig. 3 *F*, *Middle*). In contrast, at the other two levels of P_O, the model deviated from the data in a systematic way (Fig. 3*D*, arrows). At $P_O = 1$, the model underestimated RT at shorter "go" times and overestimated them at longer "go" times, as can be seen in the residual plots in Fig. 3 *F*, *Top*. At $P_O = 0.6$, the opposite pattern was observed (Fig. 3 *F*, *Bottom*). These systematic deviations between data and model across the levels of P_O are not surprising because the uncertainty parameter P_O was found to have an independent effect on RT over "go" time (Fig. 2*D*), and the reciprocal PDF model alone does not contain a component to account for it. In sum, the initial fits of the reciprocal PDF-based model capture basic features of RT modulation at all levels of P_O, although evidently an additional model component was needed to better account for the deviations between RT and model at $P_O = 1$ and $P_O = 0.6$.

We added the effect of P_O on RT to the PDF-based model to arrive at a combined model of RT that accounts for the independent effects of both γ_E and P_O. The fitted reciprocal event PDF model (Fig. 3*D* and *SI Appendix*, Fig. S8) was used as the component accounting for the effect of γ_E. To account for the effect of P_O, we built on the earlier exponential model of ΔRT (Fig. 2*D* and *SI Appendix*, Fig. S2) and used an exponential function, Φ_O, with respect to P_O and "go" times.

$$\Phi_O(P_O, t_{go}) = a\,((1 - P_O) - p)\,e^{(b|(1-P_O)-p|t_{go})} + c. \qquad [5]$$

This function allows both positive and negative exponential slopes, in accordance with the qualitative characteristics of the ΔRT curves (*SI Appendix*, Fig. S14). Note that p represents the function's pivot point, whose value was estimated from fits to ΔRT (*SI Appendix*, Tables S9 and S10 and *Supplemental Methods*). Here, instead of using the actual RTs of the reference condition ($P_O = 0.8$) for deriving ΔRT, we used the PDF-based model of RTs of the condition $P_O = 0.8$ and subtracted it from RT of the $P_O = 0.6$ and $P_O = 1$ conditions. These model-based residual curves were captured well by Φ_O (*SI Appendix*, Fig. S15). Finally, this Φ_O model was added to the event PDF-based model in order to derive the combined model of the effects of both γ_E and P_O. This combined, additive model fit the data well in both vision and audition (Fig. 4*A* and *SI Appendix*, Fig. S16), eliminating the tilting in the RT curves observed at $P_O = 1$ and $P_O = 0.6$ (Fig. 3*D*). At $P_O = 1$ and $P_O = 0.6$, the residuals' slopes of the combined model no longer differed from zero (compare with Fig. 3*F*), indicating an adequate modeling account of the effects of P_O on RT over the range of "go" times (Fig. 4*B*). The Φ_O function proved to be a beneficial model component, indicated by a significantly higher value of median-adjusted R^2: 0.67 for the model built on the reciprocal γ_E alone and 0.87 for the combined model, consisting of the fitted reciprocal γ_E and the added Φ_O function fitted to residuals (Fig. 4*C*).

We investigated whether between-participant RT variance affects the group-level fits of the combined model. Mixed-effects regression revealed that the largest part of the variance introduced at the single-participant level results from participants differing in their average RT (i.e., in offset) but not in RT curves' slope over "go" time (*SI Appendix*, *Supplemental Results 1* and Fig. S17). These findings support that the linear fits of the combined model to group-level data are adequate, and that the fits are not confounded by nonlinear between-participant differences in RT curves.

Modeling Cross-Modality Differences. The combined model of RT based on the function Φ_O and the reciprocal γ_E does not contain a component to account for potential effects of the sensory input modality. We found that median RT was shorter in auditory than in visual conditions (-17.5 ± 41.6 ms, mean ± SD, $P = 8.0 \times 10^{-12}$, $t_{(23)} = 7.13$, two-tailed Student's t test) and that IQR was also smaller in audition than in vision (-13.2 ± 33.8 ms, mean ± SD, $P = 1.8 \times 10^{-10}$, $t_{(23)} = -6.61$, two-tailed Student's t test); see *SI Appendix*, Table S3 for ANOVA. These findings agree with the literature on temporal discrimination, which suggests that audition is temporally more precise or highly resolved than vision (36–43). In light of this modality specificity in fundamental timing processes, the differences between audition and vision were further investigated. Building on 1) the possibly higher accuracy in temporal discrimination in audition and 2) the observed shorter RTs in audition (*SI Appendix*, Fig. S5*A*) compared to vision (Fig. 2*A*), audition was used as the reference condition. ΔRT curves were calculated by subtracting auditory RT from visual RT. In all experimental conditions, the curves show the highest values in ΔRT at short "go" times and thereafter monotonically decrease over time. This pattern indicates that the differences between the two modalities are dynamic over the examined range of "go" times and not fixed across time. Importantly, this pattern was observed in all four event distributions, γ_E, and at all three levels of P_O, suggesting a process independent of the two uncertainty parameters. Previous work identified similar differences in ΔRT between vision and audition and between somatosensation and audition that were also independent of γ_E (27). In this previous work, the ΔRT curves were captured by an exponential function of "go" time in three different event distributions (exponential, flipped exponential, and Gaussian) at a single level of $P_O = 0.9$. Therefore, in the present study, ΔRT curves were also modeled with the same exponential function of "go" time (*SI Appendix*, Eq. **S17**). This simple model captured ΔRT well in all conditions (*SI Appendix*, Fig. S18, black curves).

As a final step, the combined model was used to test the validity of the exponential function as a model of cross-modality ΔRT. If the exponential model of ΔRT indeed captured the

NEUROSCIENCE

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

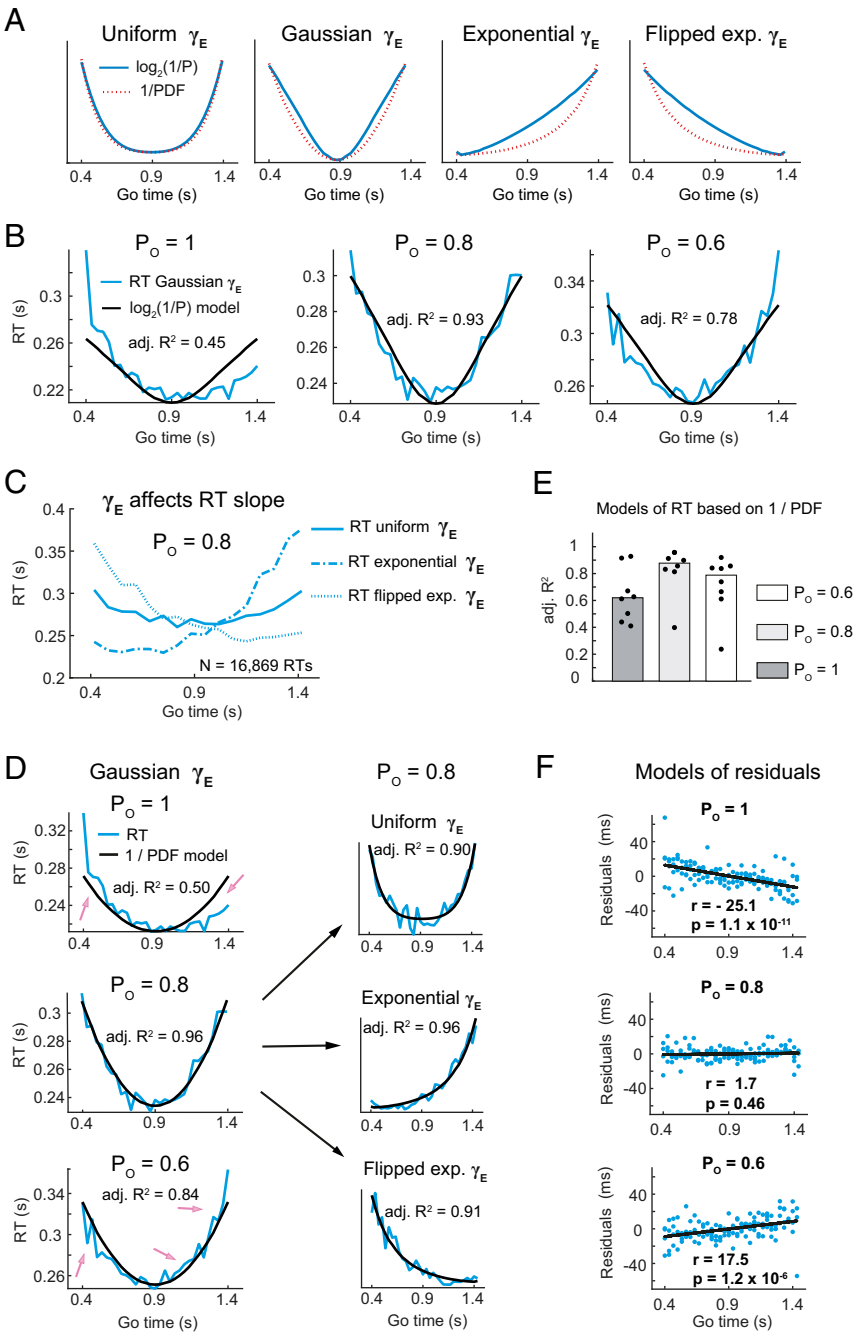

Fig. 3. Models based on γ_E fitted to visual RT. (*A*) Explanatory variables $\log_2$(1/P) (Shannon surprisal) and 1/PDF, both probabilistically blurred, predict similar RTs in the uniform case but differ in Gaussian, exponential, and flipped exponential cases, where $\log_2$(1/P) predicts more linear RT slopes. (*B*) Shannon surprisal-based model fitted to RT in Gaussian γ_E condition. (*C*) Condition-specific modulation of RT by γ_E over the entire "go" time span. Exemplary plot, P_O = 0.8. (*D*) Across levels of P_O, models based on the reciprocal, probabilistically blurred γ_E (*SI Appendix, Supplementary Methods*) capture key aspects of RT modulation. Arrows indicate systematic deviations between model and data due to the skewed shape of RT curves at P_O = 1 and, less pronounced, at P_O = 0.6. At P_O = 0.8, the model fitted the data particularly well in all γ_E conditions (black arrows). (*E*) Comparison of goodness of fit of the reciprocal, probabilistically blurred γ_E model across levels of P_O. Median-adjusted R^2, n = 8 per condition. (*F*) Residuals from fitted PDF-based model. At P_O = 1, negative slope indicates that model is skewed to left relative to data. At P_O = 0.6, positive slope indicates that model is skewed to right relative to data. No significant slope at P_O = 0.8 (linear regression) (visual conditions). In *A*, RT curves were smoothed by reducing "go" time step size from 32 to 64 ms.

difference in RT between audition and vision, then adding this model to the combined model of auditory RT from the previous section should capture visual RT in all conditions. Indeed, this combined, cross-modality model accounted well for the visual RT curves (*SI Appendix*, Fig. S19). Taken together, the exponential function as a model of cross-modality ΔRT implies that the brain's efforts to model its temporal environment based on estimation of two uncertainty parameters are affected by the processing difference between audition and vision.

Analysis of RT Distributions. As a final step, the RT distributions were analyzed to describe possible computations underlying the generation of responses. We fitted the data with an exponential–Gaussian PDF. This model is a convolution of exponential and

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

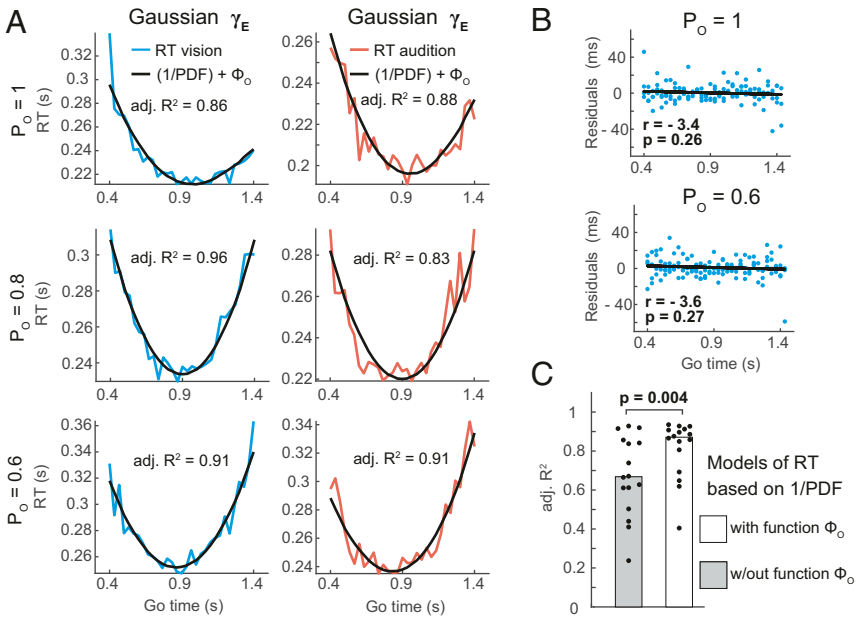

Fig. 4. Combined model based on the function Φ_O and the reciprocal γ_E captures RT modulation. (*A*) In vision and audition and across levels of P_O, RT modulation is captured by a combined model based on the reciprocal event PDF, γ_E, and the exponential function Φ_O (Gaussian γ_E condition, see *SI Appendix*, Fig. S16 for all other γ_E conditions). (*B*) Residuals from combined model (vision, all four γ_E conditions). At both $P_O = 1$ and $P_O = 0.6$, there is no significant slope indicating that adding the Φ_O function as a model component eliminated the skew observed in Fig. 3*F* (linear regression). (*C*) Model component Φ_O significantly improved goodness of fit as assessed by median-adjusted R^2 (Wilcoxon z = −2.90, n = 16).

Gaussian PDFs, and it proposes that RT can be decomposed into the sum of peripheral Gaussian process and a more central decision process that is hypothesized to be exponential (4, 44). The ex-Gaussian model provided an excellent fit to the data on both group-level (adjusted $R^2 = 0.99$, *SI Appendix*, Fig. S20) and single-participant levels (adjusted $R^2 = 0.84 \pm 0.073$, mean of all single subject fits across all conditions ± SD). The model's three parameter estimates, Gaussian μ and σ and exponential τ, were investigated over the three levels of P_O. In all γ_E conditions, and in both vision (*SI Appendix*, Fig. S21) and audition (*SI Appendix*, Fig. S22), Gaussian μ resembles the pattern observed in median RT, and the exponential parameter τ resembles the pattern observed in $IQR_{(RT)}$. Gaussian SD σ is smaller in magnitude than τ and shows no clear modulation by P_O, indicating that most of the variance in the RT distributions introduced by P_O is captured by the exponential part of the model and not by the Gaussian. In sum, the offset in RT across levels of P_O was captured by the Gaussian component, and the variance in RT was captured mostly by the exponential, giving interesting predictions for what neural activity patterns could be sought in neurophysiological recordings of the brain's responses in anticipatory behavior.

The results were validated by further control analyses. These included split data analyses investigating the stability of behavior within condition and adaptation to new experimental conditions: (*SI Appendix*, Figs. S23–S27 and Table S11) an analysis investigating the effect of early responses on fits of the combined model (*SI Appendix*, Fig. S28), an analysis investigating the potential effects of catch trials on consecutive trials (*SI Appendix*, Fig. S29), and an analysis of the potential effects of trial number on RT (*SI Appendix*, *Supplemental Results 2*).

Discussion

We investigated the influence of two uncertainty parameters: the probability of event occurrence, P_O, and the event PDF, γ_E, on temporal expectancy. RT was sensitive to both sources of uncertainty. The first finding was that the effect of P_O is not uniform across time but is dynamic. The second finding was that the dynamic effect of P_O on RT is qualitatively the same (monotonically increasing with time), irrespective of the type of event distribution γ_E. This evidence suggests that these two sources of uncertainty are processed independently. The distinct influences of P_O and γ_E on temporal expectancy were captured by a combined, additive model relating both uncertainty parameters to RT. This model summarized anticipatory behavior—in vision and audition—indicating that modality-independent, fundamental computations underlie the processing of both uncertainty parameters.

Discreteness and Catch Trials. The first parameter, P_O, representing the uncertainty of whether an event will occur at all, is typically exploited in experimental designs by the use of catch trials (i.e., trials in which no event occurs). In the vast literature of such experiments, the effect of the catch trial percentage on temporal expectancy is typically not addressed. Instead, it is implicitly assumed to be a static factor, uniform across time. We demonstrate that this assumption is wrong, and that the effect of catch trials on RT is dynamic across time. This finding has wide implications for how RT patterns should be interpreted.

The dynamic effect of P_O on RT across time was similar, irrespective of the event PDF type (uniform, Gaussian, exponential, or flipped exponential). It implies independence between the neurocomputational processes involved in the estimation of these two different sources of uncertainty.

We modeled the dynamic effect of P_O on RT with an exponential function, Φ_O (Eq. **5**). The effect of P_O on RT was small at short "go" times and increased exponentially for longer ones. The RT modulation by P_O can be intuitively summarized as the following: The lower the probability of event occurrence, P_O, the steeper the exponential increase across time. The function Φ_O captures this RT dynamic, which suggests that the brain's representation of event occurrence probability is indeed a dynamic variable in stochastic space that, similar to a PDF, unfolds in time. In sum, our results were consistent across audition and vision and

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

across four different event PDFs, γ_E, indicating that Φ_O describes a modality-independent, canonical process in temporal expectancy.

We designed our experiment to be very basic, without any decision (choice) or reward-based manipulations, in order to identify in their simplest form the foundations of event anticipation. The discovery that the discrete uncertainty of occurrence has a continuous, monotonically increasing effect on event expectancy—regardless of the continuous uncertainty—has fundamental implications for different brain functions that involve temporal anticipation, including learning, decision-making, motor planning, and perceptual processing.

We converged on describing the functional form of the monotonically increasing effect of P_O on RT with an exponential function. Note that the CDF and the HR both contain information that an event has not already happened and therefore—theoretically—these variables could also drive the effect of P_O. However, in the four event PDF conditions the HR is not even consistently monotonic (*SI Appendix*, Fig. S30), and consequently, the variable cannot account for the data. The CDF is monotonically increasing but, depending on the PDF, this monotonic increase differs in form: It is close to linear in the uniform case, sigmoidal for the Gaussian, and nonlinear in a convex (exponential) or concave way (flipped exponential) (*SI Appendix*, Fig. S30). Therefore, the CDF is not the driving parameter used by the brain to model P_O. The obvious conclusion is that the CDF and HR are not related to P_O, which indicates that the effect of P_O is independent of the distribution of events in time.

What other kind of process could account for such exponential behavior? One possibility could be that the effect of P_O on RT reflects a rate process. These are typically encountered in temporal discounting, in which the brain discounts value across time with a specific rate.

Two popular types of models used to describe such temporal discounting are exponential and hyperbolic models (13, 14). The exponential discounting models are used in economics, and in their simplest form, the value is discounted by an exponential function of the form Ae^{-rt}, where A is the value at time t = 0 and r is the discount rate (45). The hyperbolic temporal discounting models are most popular in behavioral psychology and neuroscience, and in their simplest form, the value is discounted across time by the equation $A/(1+rt)$, where A is the value at $t = 0$ and r is the discount rate across time t. We speculate that the exponential-like effect of P_O on RT can be the result of a similar exponential or hyperbolic rate process. The main difference is that this process is not decreasing but increasing as it corresponds to an exponential increase of RT. The interpretation behind either of these nonlinear functions would be that the brain reduces resources allocated to the task as time elapses, and this happens in a more pronounced way the higher the probability of an event not occurring at all. Such a reduction of resources could be seen as driven by a dynamic estimate of value. Of course, our experiments do not include a reward or punishment, and any connection to temporal discounting remains speculative.

The mechanistic implementation of the effect of P_O on RT may be related to the scalar variability, which proposes that the uncertainty in elapsed time estimation increases in direct proportion to elapsed time itself (33). This increasing uncertainty is typically represented by Gaussians with means centered on the anticipated event time point and SDs proportional to the means (10). A similar representation can be formulated for the continuous, monotonically increasing effect of P_O across time, captured by the function Φ_O. The dynamic effects of Φ_O can be reinterpreted with a Gaussian kernel whose SD scales with Φ_O (*SI Appendix*, Fig. S31). This leads to a Gaussian kernel centered at each "go" time, which can be hypothesized to reflect the allocation of attention over time (*SI Appendix*, *Discussion 1*). Contrary to the currently widespread view that inferential processes themselves—in our case, the brain's efforts to model event stochasticity based on sensory information—do not incorporate a cost function (46), we suggest that, based on the simple hypothesis of attention, Φ_O may reflect principles of economy governing neural resources. It is also tempting to interpret the effect of P_O on RT to be an attentional phenomenon because—through top-down modulation—attention has been suggested to influence early sensory processing (19, 47). This may have indirect effects on fundamental computations such as the estimation of elapsed time, which is an example of a source of endogenous uncertainty for which humans rapidly form accurate representations (48). The deployment of attention based on event probability may further influence later processing stages in the cortical hierarchy. Candidate processes include the modulation of motor system preparation based on event expectancy (49, 50). The resulting dynamics in the readiness to respond might balance the benefits of fast responses with the costs of false alarms, linking the concept of a dynamic state of expectancy, captured by Φ_O, to known features of behavior under risk (51). These hypotheses about potential mechanistic underpinnings of the effect of *Po* on anticipation require targeted experimental designs that are beyond the scope of this paper.

Continuity and Event PDFs. The second uncertainty parameter manipulated in the experiment was the type of event PDF, γ_E. The brain's representation of γ_E has received much emphasis in previous research and the mirrored, temporally blurred HR emerged as a prominent model of RT. Technically, the HR model scales the PDF, γ_E, by the reciprocal of its survival function (1 − CDF), in which CDF is the corresponding cumulative distribution function: $HR = \frac{PDF}{1-CDF}$. When there is event certainty ($P_O = 1$), the CDF approaches 1 over time. When catch trials are used, the CDF approaches P_O over time. It is clear that as time elapses and the CDF increases toward P_O that the higher the P_O, the steeper the rise of the HR.

In previous work, we have shown that the claim of the HR as a canonical model of RT does not hold (27). Instead, a computationally simpler model based on the reciprocal event PDF, γ_E, outperforms the HR (in an experimental design with a single fixed percentage of catch trials, i.e., a single level of P_O). The current experiments put both the HR and the reciprocal PDF models to the test under different values of P_O. The results clearly confirm that the reciprocal PDF is the better model across all levels of P_O (*SI Appendix*, *Discussion 2*).

Regarding the relationship between the effects of P_O and γ_E on RT, no obvious dependence was observed. The event PDF γ_E did not seem to affect the computations driven by P_O. This raises the question of how the two uncertainties are combined and represented. It was proposed that the brain can hold two parallel states, described by one PDF reflecting γ_E and one PDF reflecting the complementary probability of the event not having occurred until a time point t (26). The shapes of both PDFs are identical, but the latter PDF is inverted. In our case, it seems that the brain indeed computes two variables that change over time. To investigate this hypothesis further, however, would require targeted experiments, including the analysis of neural data.

The impact of our experimental manipulations on behavior may be investigated at different levels of analysis (52). In similar tasks involving speeded choice, the implementation of the computations involved is often described with process models, of which the drift-diffusion model (DDM) (53) is a prominent example. DDMs are widely employed in the context of two alternative, forced-choice tasks and may be extended to cover three and even more choices (54). However, DDMs' complexity and the assumptions on which they are based, like the central hypothesis that RT is a function of signal-to-noise ratio, have also led to criticism of this class of models (55). We aimed to avoid the assumptions that these process models require. Consequently, our experimental task does not

Downloaded at Palestinian Territory, occupied on December 31, 2021
www.manaraa.com

contain an obvious decision component: The act of pressing the button is solely contingent on the appearance of an easily perceptible "go" cue. Our modeling approach was guided by computational parsimony and focused on the mapping between stochastic input to RT output to generate hypotheses about the specific computations involved.

Behavioral experiments investigating only one sensory modality may fail to differentiate between central processing and modality-specific computations. We identified a processing difference between audition and vision: The dynamic modulation of cross-modality ΔRT was modeled with an exponential function (27). This model provided a good qualitative account of ΔRT, irrespective of γ_E and P_O, which invites the hypothesis of a process distinct from the estimation of the two uncertainty parameters themselves. Since the range of "go" times was not parameterized, we could not further investigate this cross-modality difference at very short "go" times. However, previous research on time estimation informs about tasks without temporal uncertainty. It is, for example, well known that the brain can synchronize to auditory metronomes of much shorter interonset intervals than visual metronomes (56). The underlying mechanisms are not well understood (57). The temporal resolution of the auditory system is argued to be higher than in vision (58), which would attribute the differences in processing we observed to more peripheral sensory contingencies. Activity in the auditory system also has been proposed to be directly related to the motor system (59–61), much more so than in vision (62), highlighting audition's complex relation to more central computations also found in temporal prediction (63). Taken together, modality specificity is an important aspect of temporal–probabilistic inference, requiring an account of both central and peripheral processing components.

Our findings illuminate how the brain models uncertainty over time. First, we show that the probability of event occurrence modulates temporal expectation dynamically across time. Second, we present compelling evidence that the two sources of uncertainty affect temporal expectancy independently, which generates the hypothesis that this behavior may be driven by independent neuronal systems. Although other sources of uncertainty may be relevant in the fundamental cognitive task of event prediction, our results aid the identification of neural correlates of predictive processes in time.

Materials and Methods

Ethics Statement. The experiments were approved by the Ethics Council of the Max Planck Society. Written informed consent was given by all participants before the experiment.

Subjects. A total of 24 human participants (15 female), aged 19 to 33 y (mean 26 y), completed the experiments. They were right-handed and had normal or corrected-to-normal vision and reported no hearing impairment and no history of neurological disorder. Participants were naive to the purpose of the experiment. They received €10 per hour for participating.

Task and Procedure. In visual and auditory blocks of trials, participants performed a simple set–go task in which a "set" cue was followed by a "go" cue. The time span between the onset of both cues (the "go" time) was a random variable that was drawn from one of four PDFs, γ_E, (uniform, Gaussian, exponential, or flipped exponential). Participants were asked to press a button as fast as possible with their right index finger in response to the "go" cue onset. In case the trial did not feature a "go" cue (a catch trial), participants were instructed to not press the button. They were asked to foveate a central black fixation dot during the entire experimental block and restrict eye blinks to the time after their response (i.e., during the intertrial interval [ITI]). After each button press, a small black circle appeared for 0.2 s around the central fixation dot, indicating the end of the trial.

The experiment consisted of four separate sessions each taking place at the same time of the day on four consecutive days. The probability of "go" cue occurrence, P_O, was manipulated in the experiment. In one third of the experimental blocks, there were no catch trials (i.e., every trial featured a "go" cue [$P_O = 1$]), in another third, the probability of a catch trial was 0.2 ($P_O = 0.8$), and in the remaining third of blocks, the probability of a catch trial was 0.4 ($P_O = 0.6$). In the catch trials, a small black circle appeared 1.9 s after "set" cue onset, indicating again the end of the trial. Within each single session, all three P_O levels were presented. The event PDF, γ_E, was fixed within single sessions (session #1: uniform, session #2: Gaussian, session #3: exponential, and session #4: flipped exponential). A session consisted of six blocks per sensory modality and lasted ∼2.5 h. A single block was comprised of 120 trials (0% catch trials, $P_O = 1$), 150 trials (20% catch trials, $P_O = 0.8$), or 200 trials (40% catch trials, $P_O = 0.6$). A short training block was run before the first block of each sensory modality on all days to familiarize participants with the task.

All stimuli were generated using MatLab (the MathWorks) and the Psychophysics Toolbox (PTB-3) (64) on a Fujitsu Celsius M730 computer running Windows 7 (64 bit). The experiment took place in a dimly lit soundproof booth. Participants wore headphones and positioned their heads on a forehead-and-chin rest (Head Support Tower, SR Research Ltd.) at a fixed distance of 60 cm relative to the computer monitor. An eye tracker (Eyelink DM-890, SR Research Ltd.) recorded participants' eye movements at a sampling frequency of 500 Hz for fixation control.

Visual Stimuli. The "set" cue consisted of two checkerboard patterns, which were presented simultaneously. One was positioned to the left of a central black fixation dot and the other on the opposite side. The "go" cue consisted of two checkerboard patterns the same location but with the black–white pattern reversed. Each checkerboard subtended 6.5 × 6.5° of visual angle and consisted of 7 × 7 black and white squares of equal size. The center of each checkerboard was positioned at a horizontal distance of 8.7° of visual angle and at a vertical distance of 0° from the center of the central fixation dot. "Set" and "go" stimuli were each presented for 50 ms on a BenQ XL2420-B monitor (resolution 1,920 × 1,080, refresh rate 144 Hz), which was set to a gray background.

Auditory Stimuli. Two white noise bursts (50 ms duration, 8 ms cosine ramp, onset and offset) served as "set" and "go" cues. The stimuli were presented diotically at the same volume level for all subjects (∼60 dB SPL) using an RME Fireface UCX interface and electrodynamic headphones (Beyerdynamic DT 770 PRO) driven by a headphone amp (Lake People GT-109).

Temporal Probabilities. The "go" time was a random variable drawn from one of four PDFs, γ_E, (Fig. 1*C*) that was fixed during each of the four experimental sessions. The distributions were constructed so that the each one could be arranged in five bins over time, with each bin containing a similar number of trials. As the probability distributions only contained integer values (i.e., the number of trials for each "go" time), it was not possible to identify PDFs with exactly the same number of trials in each quintile. This criterion was relaxed so that each quintile should have the same number of trials ± 2.5% as the neighboring quintiles.

Uniform "go" time distribution.

$$\gamma_E(x) = \frac{1}{b-a} \text{ for } x \in [a, b]. \quad [6]$$

Gaussian "go" time distribution.

$$\gamma_E(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}}. \quad [7]$$

The Gaussian distribution with parameters $\mu = 0.9$ and $\sigma = 0.25$ was truncated at the flanks, giving the distribution a spread of two SDs around the mean.

Exponential "go" time distribution. A parametric search identified a Weibull distribution with parameters $k = 1$ and $l = 0.33$ to accord with the requirements for a "go" time distribution outlined above.

$$\gamma_E(x) = \frac{k}{l}\left(\frac{t}{l}\right)^{k-1} e^{-\left(\frac{t}{l}\right)^k}. \quad [8]$$

The shape parameter k =1 reduces the Weibull to an exponential distribution:

$$\gamma_E(x) = \frac{1}{l} e^{-\left(\frac{t}{l}\right)}. \quad [9]$$

Truncation at the flanks gave the distribution a spread of 1 s to accord with the other distributions.

Flipped exponential "go" time distribution. The exponential "go" time distribution was mirrored around the mean "go" time to arrive at the flipped exponential distribution.

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

The *x*-axis of the four probability distributions was discretized with a step size of 33 ms, resulting in discrete approximations of the continuous functions. This discretization is arguably not perceivable in the context of the "set"–"go" design, rendering the distributions continuous to the brain. The *y*-axis was also discretized as the probability distributions described the number of trials at each discrete "go" time point. All distributions were offset by 0.4 s, resulting in a range of "go" times from 0.4 to 1.43 s for uniform, exponential, and flipped exponential conditions and 0.4 to 1.4 s for the Gaussian condition.

Randomization in experimental conditions. Within each of the four distributions, the order of "go" times was randomized, allowing for no more than two consecutive trials with the same "go" time to minimize sequential effects. The ITI was randomly drawn from a uniform distribution with a range of 1.4 to 2.4 s. To control for order effects, the conditions (sensory modalities, probability distributions, and level of catch trials) were arranged in a Latin square design, based on which modality, distribution, and catch trial percentage were shuffled and balanced across subjects. Within each probability distribution condition, the percentage of catch trials changed after two blocks without notification. Per block, 120 "go" cue trials were presented. Each participant was exposed to 240 "go" cue trials per sensory modality, per "go" time distribution, and per catch trial percentage, resulting in 1,440 trials per session for each subject (5,760 trials for all four sessions) and a total of 138,240 "go" cue trials for all subjects (42,240 catch trials removed).

Data selection. Trials in which visual fixation was not maintained within a radius of 5° visual angle around the central fixation point for more than 0.3 s during the "go" time were discarded for data analysis (n = 1,121 trials). Based on common practices in the literature, anticipatory responses and early guesses (4) were removed by a lower bound of RT of 0.05 s (n = 3,368). Likewise, to eliminate RTs that were unreasonably long for the employed simple RT task (4), RTs longer than 1.05 s were removed (n = 251 trials). After removal of nonfixation trials and trials whose RT was outside the defined cutoff, 133,500 trials remained for analysis. The histogram of all analyzed RTs indicates that removal of RTs did not truncate substantial parts of their distribution (*SI Appendix*, Fig. S20).

Data Availability. Anonymized RT data have been deposited in Edmond (https://edmond.mpdl.mpg.de/imeji/collection/2OqgP9U__VOl4CJT?fq=collection%3D2OqgP9U__VOl4CJT). All other study data are included in the manuscript and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Niels Hein, Claudia Lehr, Charlette Diercks, Gerald Hock, and Cornelius Abel for help with data acquisition and technical support.

1. P. Janssen, M. N. Shadlen, A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* **8**, 234–241 (2005). Correction in: *Nat. Neurosci.* **9**, 396 (2005).
2. J.-M. Schoffelen, R. Oostenveld, P. Fries, Neuronal coherence as a mechanism of effective corticospinal interaction. *Science* **308**, 111–113 (2005).
3. A. C. Nobre, F. van Ede, Anticipated moments: Temporal structure in attention. *Nat. Rev. Neurosci.* **19**, 34–48 (2018).
4. R. D. Luce, *Response Times: Their Role in Inferring Elementary Mental Organization* (Oxford University Press, 1986).
5. R. M. Church, D. M. Lacourse, J. D. Crystal, Temporal search as a function of the variability of interfood intervals. *J. Exp. Psychol. Anim. Behav. Process.* **24**, 291–315 (1998).
6. P. Niemi, R. Näätänen, Foreperiod and simple reaction time. *Psychol. Bull.* **89**, 133–162 (1981).
7. A. M. Cravo, G. Rohenkohl, V. Wyart, A. C. Nobre, Endogenous modulation of low frequency oscillations by temporal expectations. *J. Neurophysiol.* **106**, 2964–2972 (2011).
8. R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA, 1998).
9. R. A. Rescorla, A. R. Wagner, "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement" in *Classical Conditioning II: Current Research and Theory*, W. F. Prokasy, A. H. Black, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64–99.
10. C. R. Gallistel, J. Gibbon, Time, rate, and conditioning. *Psychol. Rev.* **107**, 289–344 (2000).
11. S. W. Kennerley, J. D. Wallis, Evaluating choices by single neurons in the frontal lobe: Outcome value encoded across multiple decision variables. *Eur. J. Neurosci.* **29**, 2061–2073 (2009).
12. P. N. Tobler, C. D. Fiorillo, W. Schultz, Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
13. J. W. Kable, P. W. Glimcher, The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* **10**, 1625–1633 (2007).
14. S. Kobayashi, W. Schultz, Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* **28**, 7837–7846 (2008).
15. C. K. Starkweather, B. M. Babayan, N. Uchida, S. J. Gershman, Dopamine reward prediction errors reflect hidden-state inference across time. *Nat. Neurosci.* **20**, 581–589 (2017).
16. C. D. Fiorillo, P. N. Tobler, W. Schultz, Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
17. A. Pouget, J. Drugowitsch, A. Kepecs, Confidence and certainty: Distinct probabilistic quantities for different goals. *Nat. Neurosci.* **19**, 366–374 (2016).
18. T. D. Hanks, C. Summerfield, Perceptual decision making in Rodents, monkeys, and humans. *Neuron* **93**, 15–31 (2017).
19. G. M. Ghose, J. H. R. Maunsell, Attentional modulation in visual cortex depends on task timing. *Nature* **419**, 616–620 (2002).
20. G. Notaro, W. van Zoest, M. Altman, D. Melcher, U. Hasson, Predictions as a window into learning: Anticipatory fixation offsets carry more information about environmental statistics than reactive stimulus-responses. *J. Vis.* **19**, 8 (2019).
21. A. Oswal, M. Ogden, R. H. S. Carpenter, The time course of stimulus expectation in a saccadic decision task. *J. Neurophysiol.* **97**, 2722–2730 (2007).
22. J. Sharma *et al.*, Spatial attention and temporal expectation under timed uncertainty predictably modulate neuronal responses in monkey V1. *Cereb. Cortex* **25**, 2894–2906 (2015).
23. S. Vangkilde, A. Petersen, C. Bundesen, Temporal expectancy in the context of a theory of visual attention. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20130054 (2013).
24. C. de Hemptinne, S. Nozaradan, Q. Duvivier, P. Lefèvre, M. Missal, How do primates anticipate uncertain future events? *J. Neurosci.* **27**, 4334–4341 (2007).
25. Y. Tsunoda, S. Kakei, Reaction time changes with the hazard rate for a behaviorally relevant event when monkeys perform a delayed wrist movement task. *Neurosci. Lett.* **433**, 152–157 (2008).
26. C. R. Gallistel, J. T. Wilkes, Minimum description length model selection in associative learning. *Curr. Opin. Behav. Sci.* **11**, 8–13 (2016).
27. M. Grabenhorst, G. Michalareas, L. T. Maloney, D. Poeppel, The anticipation of events in time. *Nat. Commun.* **10**, 5802 (2019).
28. C. H. Papadimitriou, S. S. Vempala, D. Mitropolsky, M. Collins, W. Maass, Brain computation by assemblies of neurons. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14464–14472 (2020).
29. C. E. Shannon, A mathematical theory of communication. *Bell Sys. Tech. J.* **27**, 379–423 (1948).
30. S. Dehaene, The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends Cogn. Sci.* **7**, 145–147 (2003).
31. W. E. Hick, On the rate of gain of information. *Q. J. Exp. Psychol.* **4**, 11–26 (1952).
32. N. J. Smith, R. Levy, The effect of word predictability on reading time is logarithmic. *Cognition* **128**, 302–319 (2013).
33. J. Gibbon, Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* **84**, 279–325 (1977).
34. M. Jazayeri, M. N. Shadlen, Temporal context calibrates interval timing. *Nat. Neurosci.* **13**, 1020–1026 (2010).
35. K. Vierordt, *Der Zeitsinn nach Versuchen* (Laupp, Tübingen, Germany, 1868).
36. S. Grondin, R. Rousseau, Judging the relative duration of multimodal short empty time intervals. *Percept. Psychophys.* **49**, 245–256 (1991).
37. R. Rousseau, J. Poirier, L. Lemyre, Duration discrimination of empty time intervals marked by intermodal pulses. *Percept. Psychophys.* **34**, 541–548 (1983).
38. E. A. Williams, E. M. Yüksel, A. J. Stewart, L. A. Jones, Modality differences in timing and the filled-duration illusion: Testing the pacemaker rate explanation. *Atten. Percept. Psychophys.* **81**, 823–845 (2019).
39. W. Zarco, H. Merchant, L. Prado, J. C. Mendez, Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* **102**, 3191–3202 (2009).
40. S. Grondin, Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* **72**, 561–582 (2010).
41. S. Grondin, "Sensory modalities and temporal processing" in *Time and Mind II: Information Processing Perspectives*, H. Helfrich, Ed. (Hogrefe & Huber, Göttingen, 2003), pp. 61–77.
42. J. A. Grahn, See what I hear? Beat perception in auditory and visual rhythms. *Exp. Brain Res.* **220**, 51–61 (2012).
43. M. J. Hove, M. T. Fairhurst, S. A. Kotz, P. E. Keller, Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *Neuroimage* **67**, 313–321 (2013).
44. R. H. Hohle, Inferred components of reaction times as functions of foreperiod duration. *J. Exp. Psychol.* **69**, 382–386 (1965).
45. L. Green, J. Myerson, E. McFadden, Rate of temporal discounting decreases with amount of reward. *Mem. Cognit.* **25**, 715–723 (1997).
46. A. Pouget, J. M. Beck, W. J. Ma, P. E. Latham, Probabilistic brains: Knowns and unknowns. *Nat. Neurosci.* **16**, 1170–1178 (2013).
47. J. W. Bisley, M. E. Goldberg, Attention, intention, and priority in the parietal lobe. *Annu. Rev. Neurosci.* **33**, 1–21 (2010).
48. F. Balci, D. Freestone, C. R. Gallistel, Risk assessment in man and mouse. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2459–2463 (2009). Correction in: *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11424 (2009).
49. S. Roux, W. A. Mackay, A. Riehle, The pre-movement component of motor cortical local field potentials reflects the level of expectancy. *Behav. Brain Res.* **169**, 335–351 (2006).
50. S. Bestmann *et al.*, Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr. Biol.* **18**, 775–780 (2008).

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com

51. W. Schultz *et al*., Explicit neural signals reflecting reward uncertainty. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3801–3811 (2008).
52. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (W. H. Freeman and Company, New York, NY, 1982).
53. R. Ratcliff, G. McKoon, The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
54. A. Roxin, Drift-diffusion models for multiple-alternative forced-choice decision making. *J. Math. Neurosci.* **9**, 5 (2019).
55. P. Sun, M. S. Landy, A two-stage process model of sensory discrimination: An alternative to drift-diffusion. *J. Neurosci.* **36**, 11259–11274 (2016).
56. B. H. Repp, Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *J. Mot. Behav.* **35**, 355–370 (2003).
57. B. H. Repp, Sensorimotor synchronization: A review of the tapping literature. *Psychon. Bull. Rev.* **12**, 969–992 (2005).
58. A. O. Holcombe, Seeing slow and seeing fast: Two limits on perception. *Trends Cogn. Sci.* **13**, 216–221 (2009).
59. J. L. Chen, V. B. Penhune, R. J. Zatorre, Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* **18**, 2844–2854 (2008).
60. T. Fujioka, L. J. Trainor, E. W. Large, B. Ross, Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. *J. Neurosci.* **32**, 1791–1802 (2012).
61. J. A. Grahn, J. B. Rowe, Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* **29**, 7540–7548 (2009).
62. M. H. Thaut, G. P. Kenyon, M. L. Schauer, G. C. McIntosh, The connection between rhythmicity and brain function. *IEEE Eng. Med. Biol. Mag.* **18**, 101–108 (1999).
63. J. M. Rimmele, B. Morillon, D. Poeppel, L. H. Arnal, Proactive sensing of periodic and aperiodic auditory patterns. *Trends Cogn. Sci.* **22**, 870–882 (2018).
64. D. H. Brainard, The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).

Downloaded at Palestinian Territory, occupied on December 31, 2021

www.manaraa.com