



OPEN

Change point detection with multiple alternatives reveals parallel evaluation of the same stream of evidence along distinct timescales

Alexa Booras^{1,4}, Tanner Stevenson^{1,4}, Connor N. McCormack¹, Marie E. Rhoads^{1,2} & Timothy D. Hanks^{1,3}✉

In order to behave appropriately in a rapidly changing world, individuals must be able to detect when changes occur in that environment. However, at any given moment, there are a multitude of potential changes of behavioral significance that could occur. Here we investigate how knowledge about the space of possible changes affects human change point detection. We used a stochastic auditory change point detection task that allowed model-free and model-based characterization of the decision process people employ. We found that subjects can simultaneously apply distinct timescales of evidence evaluation to the same stream of evidence when there are multiple types of changes possible. Informative cues that specified the nature of the change led to improved accuracy for change point detection through mechanisms involving both the timescales of evidence evaluation and adjustments of decision bounds. These results establish three important capacities of information processing for decision making that any proposed neural mechanism of evidence evaluation must be able to support: the ability to simultaneously employ multiple timescales of evidence evaluation, the ability to rapidly adjust those timescales, and the ability to modify the amount of information required to make a decision in the context of flexible timescales.

Decision making often involves evaluation of information from the environment to form a response. In order to behave appropriately in a rapidly changing world, individuals must be able to detect when changes occur in that environment. This process is referred to as change point detection¹. In real world situations, there are many possible aspects of the environment that could change, but there are often constraints on what changes are possible. For instance, when tracking a fly in the air, the fly could change direction along three dimensions, but when tracking an ant on the ground, the ant could change direction along only two dimensions. We sought to address how knowledge about the space of possible changes affects human change point detection.

Sequential sampling models that have been applied to a wide range of decision making tasks provide a framework to describe the process of change point detection. In these models, an observer makes a choice based on a decision variable that combines information from different samples of evidence gathered over time, a process we refer to as evidence evaluation. Evidence evaluation can take a variety of forms, and the optimal form depends on task demands and environmental conditions. For example, in a stable environment, all samples of evidence can be treated equally, and evidence can be perfectly integrated over a long timescale to maximize performance²⁻⁷. In contrast, in an unstable environment, samples from the past are less likely to reflect the current state of the world, and evidence should be evaluated over a briefer timescale to avoid integrating information that is no longer relevant⁸⁻¹⁰. Change point detection is a specific example that involves an unstable environment and shorter timescales of evidence evaluation¹⁰⁻¹².

Knowledge about the space and type of possible changes may affect evidence evaluation for change point detection. Different types of changes may have distinct statistical properties that demand different optimal

¹Center for Neuroscience, University of California Davis, Davis, CA, USA. ²Department of Neuroscience, University of California Los Angeles, Los Angeles, CA, USA. ³Department of Neurology, University of California Davis, Sacramento, CA, USA. ⁴These authors contributed equally: Alexa Booras and Tanner Stevenson. ✉email: thanks@ucdavis.edu

timescales of evidence evaluation. For example, detecting the approach of a speedy predator benefits from a shorter timescale of evidence evaluation than detecting the approach of a slow and stealthy predator. Thus, information about the nature of the change—for example being in a location with only speedy predators or only slow, stealthy predators—has the potential to improve change detection performance through optimization of the timescale of evidence evaluation.

Knowledge about the space and type of possible changes may also improve change detection performance through mechanisms typically associated with selective attention^{13–15}. In the sequential sampling framework, each sample of evidence that is evaluated carries a signal about the environment that is potentially corrupted by signal-independent variability known as noise^{2,3}. That noise may involve environmental factors, but it may also be related to neural processing^{16–19}. To the extent the brain accounts for this noise, knowledge about the possible changes may affect the signal to noise ratio of the evidence evaluated for change point detection. Numerous studies using a variety of tasks have described improvements in performance with attention that are attributable to increases in the signal to noise ratio that enhance perceptual sensitivity of attended stimuli^{20–22}. Thus, knowledge about the space of possible changes may alter evidence evaluation by changing perceptual sensitivity for change point detection.

A third possibility is that knowledge about the space and type of possible changes may alter decision commitment policies. In the sequential sampling framework, a bound on the decision variable determines the timing of decision commitment^{23–26}. For change point detection, a lower decision bound results in fewer misses at the expense of more false alarms, and a higher decision bound results in fewer false alarms at the expense of more misses¹¹. Thus, the decision bound determines the tradeoff between false alarms and misses. There are many important considerations for setting the decision bound for change point detection. Among those is the number of potential changes that are simultaneously being monitored in any given situation. If each potential change carries a fixed probability of false detection, then the greater the number of potential changes simultaneously under consideration, the greater the probability of a false alarm detection of one of those potential changes. Therefore, higher decision bounds may be required in situations with a greater number of potential changes that need to be considered. Following this logic, if fewer potential changes are under consideration, false alarm rates can be kept in check even with relatively lower decision bounds with the benefit of also reducing the miss rate. Thus, knowledge about the space of possible changes may affect the setting of decision bounds for change point detection.

Previous research suggests that both perceptual and decision components of processing can be influenced by cues about possible changes in the environment²⁷. As a particularly relevant example for our study, Sridharan and colleagues showed dissociations of the effects of spatial cueing between sensory and decision components of processing for “yes/no” change detection in the visual domain, where people have to report if a change has occurred between two comparison stimuli. In particular, they found that sensory effects of cueing were spatially localized while decision effects were not²⁸. In order to distinguish these possibilities, they used a multidimensional signal detection model that relies on a comparison of two samples of evidence²⁹, as is appropriate for the “yes/no” detection task they studied. However, it is not known how cues about the space and type of possible changes influence change point detection, which involves evaluation of multiple sequential samples of evidence rather than a comparison of two.

Here, we sought to address how knowledge about the space of possible changes affects evidence evaluation and decision bounds for an auditory change detection task we have developed. In this task, subjects must rapidly report when a change occurs in an auditory stimulus that is unpredictable in time. The change in the stimulus could be of two different varieties, and we provided cues that either informed the subject about the type of change that would occur or were uninformative about the type of change. We compared performance when the cue was informative versus uninformative using model-free psychophysical reverse correlation analyses combined with sequential sampling models to characterize behavior. We found that subjects can simultaneously apply distinct timescales of evidence evaluation to the same stream of evidence. Furthermore, accuracy for detection improved with informative cues through mechanisms involving these timescales of evidence evaluation and adjustments of decision bounds. These results establish important capacities of information processing for decision making that any proposed neural mechanism of evidence evaluation must be able to support.

Methods

Subjects. We recruited 15 subjects to perform this experiment. Six subjects elected to discontinue the experiment, and we analyzed the data of the remaining nine subjects. All subjects were undergraduate students from UC Davis over the age of 18. None had any prior knowledge of the research motivations or task design prior to data collection. Subjects were compensated with \$10 Amazon gift cards for every one hour session, regardless of their performance on the task. The study procedures were approved by the UC Davis Institutional Review Board and all experiments were performed in accordance with relevant guidelines and regulations including that all subjects provided informed consent.

Subjects performed one session of the experiment per day, with each session consisting of three blocks lasting 15 min and separated by short breaks. The first session was used as a training session, and the data was not analyzed. Subjects completed either three or four non-training sessions, with three subjects electing not to return for the final session.

Change detection task. Subjects were asked to perform an auditory change detection task. They were seated in an isolated room and given headphones to listen to the stimulus. The testing apparatus consisted of three ports with LEDs that indicated whether a port could be used, and IR sensors that detected when a finger was inserted into a port. The stimuli were generated with Psychtoolbox and programmed in MATLAB. The task

| Subject | Informed increase | Informed decrease | Uninformed increase | Uninformed decrease |
|----------|-------------------|-------------------|---------------------|---------------------|
| 1 | 379 | 374 | 405 | 390 |
| 2 | 555 | 559 | 556 | 532 |
| 3 | 451 | 428 | 475 | 463 |
| 4 | 541 | 561 | 544 | 600 |
| 5 | 510 | 526 | 511 | 533 |
| 6 | 551 | 545 | 511 | 510 |
| 7 | 390 | 400 | 419 | 427 |
| 8 | 345 | 364 | 372 | 380 |
| 9 | 415 | 410 | 440 | 417 |
| Combined | 4137 | 4167 | 4233 | 4252 |

Table 1. Trial condition counts for each subject and totals for combined subject analysis.

was controlled through BPod which recorded the behavioral responses in real time. Auditory feedback was generated by the PC and played through the headphones.

Subjects were instructed to report a change in the underlying rate of a series of clicks randomly generated by a Poisson process. The change in rate could be either an increase or a decrease. The underlying rate was initially 60 Hz. For all trials the frequency changed with equal probability by ± 10 , ± 30 , or ± 50 Hz. We intentionally kept the lowest click rate above 10 Hz to maintain a flutter percept such that individual clicks could not be counted. However, in order to have psychometric functions spanning a range of performance levels, our resulting stimuli span a range of flutter and potentially some forms of pitch percepts³⁰. There was always a minimum of 1 s before any change, after which the change time was drawn from an exponential distribution with a mean of 3.5 s and truncated at a maximum of 7.5 s. The exponential distribution results in a flat hazard rate of the stimulus change between the 1-s pre-change period and the 7.5-s maximum, so the probability of the change occurring did not vary throughout most of the trial. Before each trial, a cue on the computer screen displayed a word indicating what type of trial would occur. The cues were either informative, stating “increase” or “decrease”, or uninformative, stating “either”. The trial type was chosen at random, with a 50% chance of the trial being informed. For uninformed trials, the change direction was chosen at random with a 50% chance of the trial being an increase. This resulted in an equal amount of informed and uninformed trials, and increase and decrease trials (Table 1).

At the start of each trial, an illuminated LED in the center port of the apparatus indicated the trial was ready to begin, and the trial cue appeared on screen. Subjects inserted their finger into the port to start the trial, triggering the onset of the auditory stimulus. Finger pokes were detected by the finger obstructing an infrared light across the port. Subjects were instructed to keep their finger in the port until they detected the change in click rate, at which point they were instructed to remove their finger immediately which ended the stimulus. If subjects did not remove their finger within a 0.8 s response window after the change, the stimulus would end automatically. If the subject removed their finger within the response window, side LEDs illuminated indicating that the subject should report whether they thought the change was an increase or decrease in click rate. If the subject responded to the change within the response window and correctly reported the direction of change, this trial was a “hit”. Trials in which subjects responded within the response window but reported the incorrect direction were few in quantity and were therefore not analyzed. There were two primary types of errors: premature responses, or ‘false alarms’, and failures to respond, or ‘misses’. In the case of false alarms, subjects were instructed to report the direction that they actually responded to, even if this was different than what the cue indicated. In the case of misses, the side LED corresponding to the correct direction of change briefly illuminated at the end of the trial. Immediately following the direction report, feedback was given via a high or low pitched auditory tone to indicate success or failure on that trial respectively. There were two additional types of errors that signified lapses in attention to the task—responses opposite the informative cue or responses within 0.75 s of the trial start—that were excluded from analysis.

Psychometric functions. Psychometric functions were computed for individual subjects and for pooled subject data. Hit rates, false alarm rates, and reaction times were calculated for each of the four trial categories: informed increase, informed decrease, uninformed increase, and uninformed decrease. Hit rates were calculated for each change delta in each category as the ratio of ‘hits’ to the number of trials where a change occurred; therefore false alarms were excluded from the calculation. 95% confidence intervals were calculated using the MATLAB *binofit* function. To quantify performance changes between uninformed and informed conditions the change delta associated with a 50% hit rate was estimated by fitting a sigmoid curve to the hit rates at each change delta in each trial category. Confidence intervals were constructed for this threshold change delta with a bootstrapping method wherein thresholds were computed from hit rate curves constructed from data matching the trial count resampled 1000 times with replacement. The 95% confidence intervals were estimated as the 5th and 95th percentiles of the bootstrapped thresholds.

Informed false alarm rates were calculated as the ratio of informed false alarms to the total number of trials in each informed condition. False alarms with responses opposite the cue were excluded from the calculation. Uninformed false alarm rates in each direction were calculated as the ratio of false alarms in one direction to the total number of uninformed trials minus the false alarms in the other direction. This exclusion attempts to

account for the possibility that a false alarm in one direction masks the potential for a later false alarm in the other direction. 95% confidence intervals were calculated using the MATLAB *binofit* function. Confidence intervals for false alarm rate differences between conditions were calculated using a bootstrapping method where rate differences were calculated from data matching the trial count resampled 1000 times with replacement. The 95% confidence intervals were estimated as the 5th and 95th percentiles of the bootstrapped differences.

Reaction times were calculated as the time from the click rate change to the subject's response. The 95% confidence intervals were estimated as the 5th and 95th percentiles of 1000 bootstrapped means.

Psychophysical reverse correlations. We utilized psychophysical reverse correlations (RC) of false alarms to investigate how subjects evaluate evidence in different trial conditions^{11,31–33}. RC 'kernels' were produced by aligning false alarm trial stimuli to the time of the response, computing instantaneous click rates by convolving the click times with a half Gaussian filter ($\sigma=0.075$ s), and averaging over all trials grouped by cue and response direction. False alarms with responses opposite the cue were excluded from the calculation. To quantitatively compare the RC kernels, we computed the average change in click frequencies from baseline for each category of false alarms. This was done by counting the number of clicks in the 1 s window preceding the false alarm and subtracting the baseline rate of 60 Hz. The 95% confidence intervals were estimated as the 5th and 95th percentiles of 1000 bootstrapped values. We also estimated the timescale of evidence evaluation by computing the start of the RC kernel in each trial type. Kernel starts and confidence intervals were estimated by fitting a piecewise linear function from the baseline to the kernel start and from the kernel start to the kernel maximum (or minimum).

Statistical significance. Hypothesis testing was performed using a bootstrap method to avoid assumptions about the underlying distributions³⁴. Test statistics were computed from 1000 bootstrap data sets sampled with replacement from a pseudo null distribution that contain the same number of trials as the original data set. For 50% hit rate threshold and false alarm rate differences between informed and uninformed conditions, a one-sided p-value was calculated from the number of bootstrapped differences that were opposite in sign to the average differences for increasing and decreasing stimuli. For all other metrics, t-statistics were calculated based on the null hypothesis that the metrics came from the same underlying distribution, so bootstrap data sets were sampled from a pseudo null distribution containing all values from both data sets. Two-sided p-values were computed from these statistics from the number of bootstrap statistics whose absolute values were greater than the absolute value of the test statistic computed from the unmodified data set.

Model based analysis. To generate a better algorithmic understanding of the differences in decision making processes between trial conditions, we fit sequential sampling behavioral models separately to informed increase, informed decrease, and uninformed trials from individual and combined subject data. In general, the model estimates a response time for each trial by convolving the clicks with an exponential filter to produce a decision variable (DV), determining when the DV crosses a bound, and adding a non-decision time (NDT) delay to account for sensory and motor delays (Fig. 5). To account for behavioral variability, we added noise to this process and computed response time (RT) distributions. Variability was added to the NDT to properly account for variability in perceptual and motor delays. In addition, lack of NDT variability in the model would result in zero likelihood for any RT shorter than the mean NDT and therefore force the mean NDT parameter to be shorter than the fastest measured RT. We fit our model to the behavioral data by maximizing the likelihood that the subjects' response times could be produced by our model parameters: exponential filter width (τ), bounds, mean NDT (μ_{NDT}), NDT variability (σ_{NDT}), and process noise ($\sigma_{process}$). Models for the informed conditions consisted of a single process with a single high or low bound, whereas the uninformed model consisted of two processes: one to detect increases and the other to detect decreases. We also compared two different versions of the uninformed model: one with separate filter widths per process and another with a shared filter width. We found NDT variability did not differ between trial conditions and therefore models for all three conditions were fit simultaneously sharing the same NDT variance parameter, which slightly reduced the overall model complexity.

The model uses the click stimulus to drive the dynamics of a decision variable that is governed by an Ornstein–Uhlenbeck (OU) process. This is equivalent to a drift–diffusion process where clicks act as a time-varying drift term whose contributions to the decision variable decay over time exponentially. The exponential form of this decay has the benefit of conferring the Markov property to the decision variable, therefore rendering its current state independent from all prior states. Other functional forms of this decay could be employed in theory, such as a half Gaussian filter, but in that case, the DV trajectory would depend on its prior states. For example, consider applying a half Gaussian filter during a period without clicks. In this scenario, the future DV trajectory would differ depending on the exact times of prior clicks, even when starting from the same value, because the contribution of more recent clicks will decay at a slower rate than that of clicks further in the past due to the shape of the Gaussian. This time-varying decay rate significantly complicates the behavioral model; therefore we deemed the OU process most amenable to the robust modeling approach we employed.

The OU process is described with the following input³⁵:

$$da = -(a/\tau + c)dt + \sigma_{process} dW \quad (1)$$

where a is the DV (in Hz), c is the DV input within each interval, $\sigma_{process}$ represents noise at the level of the DV, and dW is white noise. For our purposes, c was calculated as:

$$c = C/(dt * \tau) \quad (2)$$

where C is the number of clicks in each time interval. Dividing the click frequency in the time interval (C/dt) by the filter width scales the DV units to Hz, which allows for direct parameter comparison across trial conditions with different filter widths. Finally, because previous work with Poisson click-based auditory decision tasks revealed click input depression³⁵, we also applied a version of the model that incorporated input depression. With no depression, each click contributes 1 unit to C as described above. With depression, the click magnitude C is multiplied by a scale factor after each click that then recovers towards the unadapted value of 1 with a time constant set by another term. For this procedure, we used the maximal depression terms found in previous human subjects in a Poisson clicks task: a depression scale factor of 0.45 and time constant of 0.02 s³⁵.

To calculate RT distributions for each trial, we used a discretized solution to the Fokker–Planck equation corresponding to Eq. (1) with a timestep (dt) of 0.02 s. With this solution, we propagated a discretized probability distribution of DVs for each trial starting from DV = 0 at $t = 0$ ³⁵. This propagation was performed in two segments: first without bounds for 0.8 s, then with upper and/or lower ‘sticky’ bounds depending on the trial type. Informed increase trials used an upper bound, informed decrease trials used a lower bound, and uninformed trials used both an upper and lower bound. Sticky bounds were implemented so that any portion of the DV distribution that crossed these bounds would persist there, and thus, the proportion of the DV distribution that had crossed a bound would accumulate over time. The first DV propagation segment is designed to correspond to the period of time right after stimulus onset where subjects know there is no change and do not need to respond. This is especially important when implementing lower sticky bounds to establish a baseline DV estimate that is higher than the lower bound. The time of 0.8 s was chosen as a reasonable minimum for the subjects’ estimate of the change detection task’s baseline period (fewer than 1% of all trials had responses in this time period, with the vast majority being within 0.1 s corresponding to failures to completely obstruct the infrared light used to detect finger pokes). We used the proportion of the DV distribution that crossed each bound within each time bin as the bound crossing probability distribution, and then computed RT distributions by convolving the bound crossing distribution with a Gaussian NDT filter parameterized by μ_{NDT} and σ_{NDT} . Uninformed trials had separate RT distributions for responding to increases and decreases that were computed as joint distributions of responding to changes in one direction over the other.

The likelihood of each trial was calculated from the RT distribution by summing the distribution values in a 0.2 s window around the subject’s response time. Uninformed trials used the RT distribution corresponding to the reported direction of the response. For ‘miss’ trials where there is no response time, the trial likelihood was calculated as the remainder of the DV distribution that did not cross any bounds before the end of the trial. Optimal parameters were found by maximizing the cumulative log likelihood over all trials using both Bayesian Adaptive Direct Search (BADS)³⁶ and Variational Bayesian Monte Carlo (VBMC)^{37,38} optimization algorithms. These algorithms were chosen because they were both designed to solve difficult optimization problems with non-analytical likelihood calculations, and VBMC allows for confidence interval estimation. BADS was first used to find a region of the parameter space with a high likelihood. The returned parameters were then used as the starting point for VBMC, which simultaneously estimates posterior distributions of the model parameters and a lower bound of the log model evidence (ELBO). VBMC requires a prior over the parameters, and since we made no assumptions about the parameter distributions, we used a uniform prior over the range of possible parameter values. Optimal values were estimated as the mean of 100,000 samples randomly drawn from each approximated posterior distribution. 95% confidence intervals were estimated as the 5th and 95th quantiles of the same random samples. Hypothesis testing between analogous parameters across trial conditions was performed by a two-tailed t-test calculated from a t-statistic on the parameter differences. ELBO values are analogous to the Bayes factor and were used for model selection. Multiple model fits were run with varying starting parameter values to ensure the optimal parameters were consistently found.

Results

Task structure. In this experiment, nine subjects performed an auditory change detection task (Fig. 1; see methods for details). They listened to a stream of auditory clicks generated by a 60 Hz Poisson process while holding their index finger in a response port. Subjects were told to remove their finger when they detected either an increase or decrease in the generative click rate. The timing and magnitude of the change were randomly chosen. At the start of each trial a cue either informed subjects of the direction of the change by stating “increase” or “decrease”, or the cue was uninformative and stated “either”. At the end of each trial with a reported change, subjects then reported the direction of perceived change.

Task performance. All the subjects learned the change detection task quickly and their performance plateaued after a pre-experimental practice session. Subject performance depended on the informed and uninformed task conditions (Fig. 2). Similar to results from related experiments^{11,12,39}, the hit rate of the informed trials for increasing stimuli was lowest at the smallest stimulus delta (10 Hz), and increased until peaking at the largest delta (50 Hz), with a psychometric threshold (50% hit rate) at a delta of 30.4 Hz (CI: 29.2 to 31.5). This psychometric function shows that subjects were attending to the stimuli. Performance on informed trials followed a similar trajectory for the decreasing stimuli. Hit rates were the lowest at –10 Hz and the highest at –50 Hz. The psychometric threshold for these trials was –27.7 Hz (CI: –28.6 to –26.9).

Performance was noticeably different in the uninformed condition. While hit rates followed the same pattern of increasing with the absolute magnitude of the stimulus change, hit rates were lower than the informed condition for all stimuli, on average. Uninformed trials with increasing stimuli had a psychometric threshold of 38.6 Hz (CI: 37.4 to 39.7), which was significantly higher than informed trials ($p < 0.001$). Uninformed trials

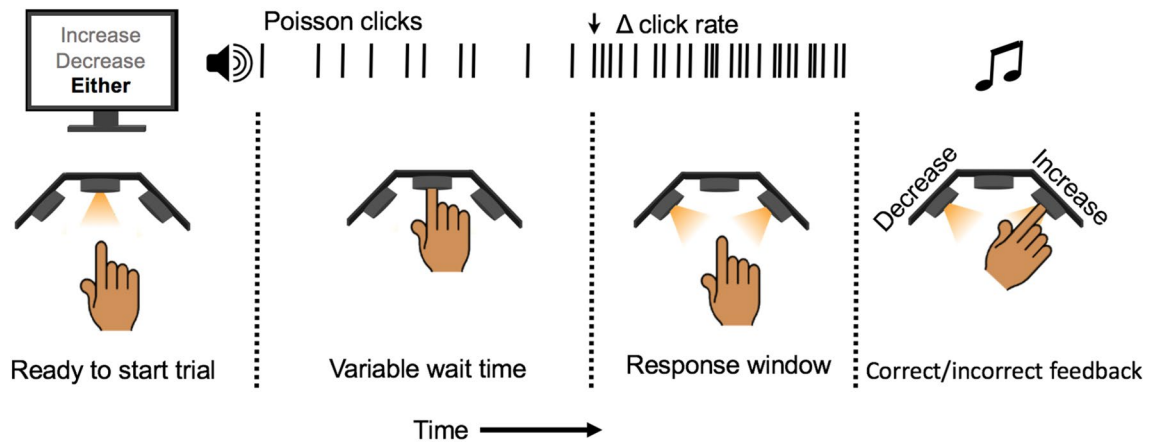


Figure 1. Schematic of change detection task. An illuminated LED indicates the trial is ready to begin. The computer screen displays a word indicating what type of trial will occur. The cues are either informative, stating “increase” or “decrease”, or uninformative, stating “either”. Subjects insert their finger into the port, triggering the onset of a stochastic auditory stimulus and keep their finger in the port until they detect the change in click rate upon which they remove their finger. Then they insert their finger into either the left port to report a decrease or the right port to report an increase.

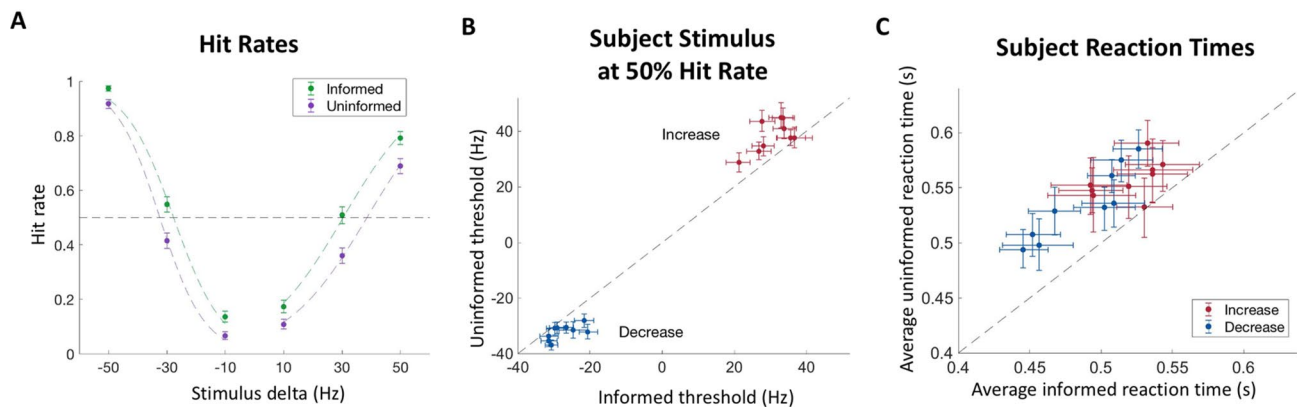


Figure 2. Hit rates and reaction times. (A) Performance plotted as a function of the change in click rate for each trial condition. Colored dashed lines indicate fits by logistic curves and the dashed line indicates a 50% hit rate psychometric threshold. Data is combined from all subjects. (B) The psychometric threshold for each subject in the informed and uninformed conditions of both increasing and decreasing stimuli trials. (C) Average hit rate reaction times of each subject for increasing and decreasing stimuli in the informed versus uninformed conditions. All error bars indicate 95% confidence intervals. $N=9$ subjects.

with decreasing stimuli had a psychometric threshold of -32.4 Hz (CI: -33.1 to -31.7), which was significantly larger in magnitude than informed trials ($p < 0.001$). These differences were consistent across subjects for both the increasing and decreasing stimuli (Fig. 2B, Supplementary Table S1).

We also measured the reaction times of hit trials. Consistently, reaction times were longer in the uninformed condition than the informed condition, with subjects responding more quickly to changes when the cue was informative. The average reaction time on uninformed trials was 0.54 s (CI: 0.54 – 0.55), and on informed trials, it was 0.50 s (CI: 0.49 – 0.51), with consistent differences of this type across subjects regardless of whether it was detection of increasing or decreasing stimuli (Fig. 2C, Supplementary Table S2).

Finally, we examined the false alarm rate for each trial type. The average rate was 18.9% (CI: 18.91 – 18.94) and 13.3% (CI: 13.31 – 13.33) for the informed condition, and 11.2% (CI: 11.18 – 11.20) and 5.7% (CI: 5.70 – 5.72) for the uninformed condition, for increasing and decreasing stimulus trials respectively (Fig. 3A). When plotting the difference between informed and uninformed false alarm rates in Fig. 3B, we found that in both increasing and decreasing trial groups, the difference was positive (μ : 0.077 , CI: 0.063 – 0.092 and μ : 0.076 , CI: 0.066 – 0.087 , respectively). In other words, the false alarm rates were higher in the informed condition than the uninformed condition. These differences were consistent across subjects for both the increasing and decreasing stimuli (Fig. 3C, Supplementary Table S3).

Model-free analysis. To further assess the differences between trial conditions, we computed the psychophysical reverse correlation (RC) kernels for false alarms of each trial type. The RC kernels were calculated by convolving the click times with a causal half Gaussian filter and aligning the result to the time of the response

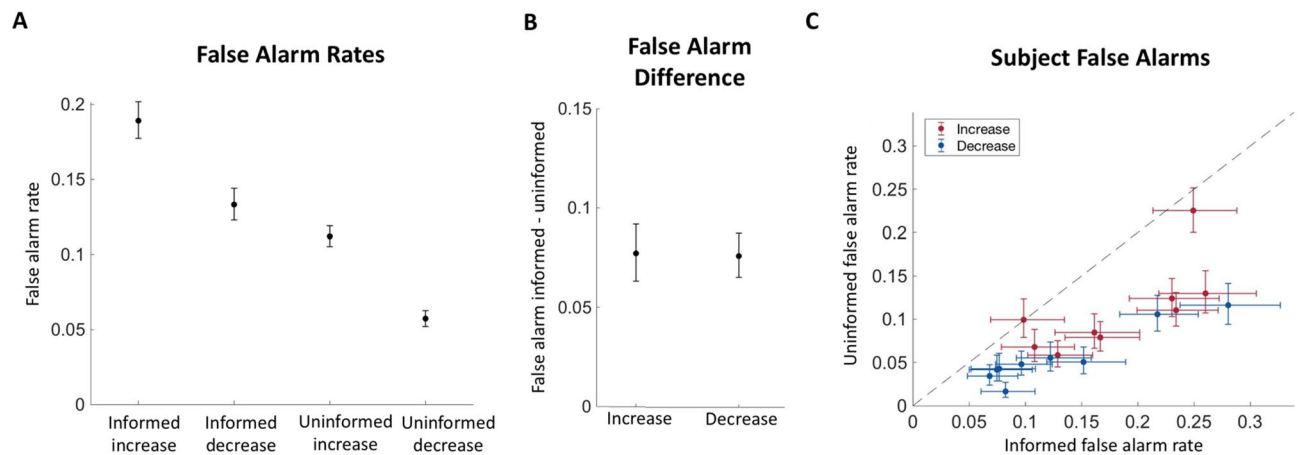


Figure 3. False alarm rates. (A) False alarm rate for each trial condition. Data is combined from all subjects. (B) Difference between informed and uninformed false alarm rates for combined subject data. (C) Informed and uninformed false alarm rates for each subject. All error bars indicate 95% confidence intervals. $N=9$ subjects.

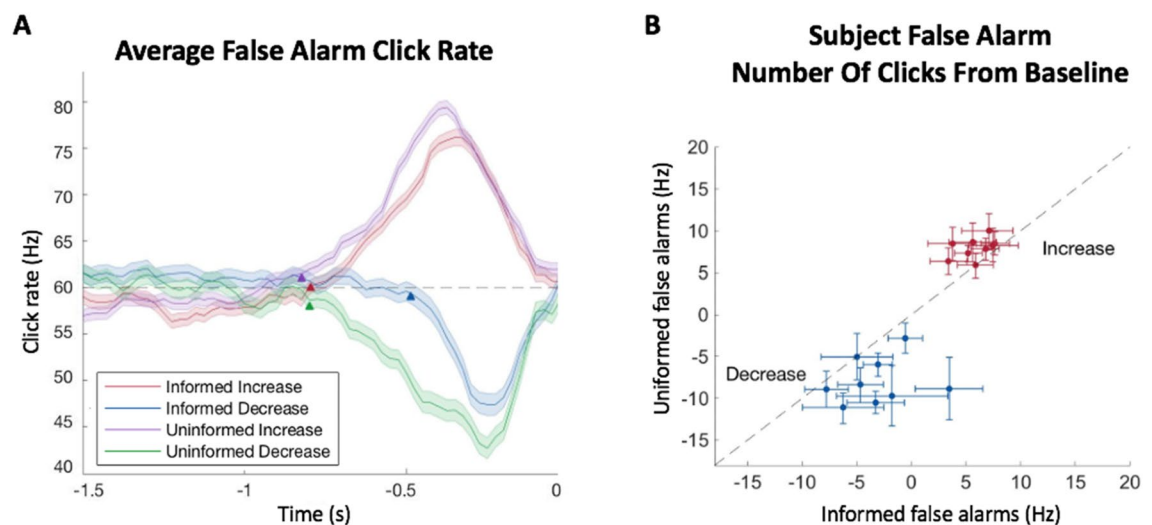


Figure 4. False alarm RC kernels. (A) Detection kernels of different trial types for pooled subject data. Shaded regions indicate the standard error of the mean. Arrows indicate the kernel start point. Dashed line indicates baseline click rate. Data is combined from all subjects. (B) The average deviation in click rate from the baseline preceding a false alarm for increasing and decreasing trials. Error bars indicate 95% confidence intervals. $N=9$ subjects.

for each false alarm trial. This allowed us to reconstruct the average stimulus that preceded a false alarm in order to discern the influence of the stimulus on the false alarm decision. Segments where the RC kernel varies from the baseline click rate indicate the average periods and stimulus magnitudes that influenced the subsequent choice. In our data, increasing false alarms were preceded by an increase in click rate and decreasing false alarms were preceded by a decrease in click rate (Fig. 4A). The RC kernels also suggest there was a larger change in the stimulus to cause a false alarm in uninformed conditions than in informed conditions, for both increases and decreases. To confirm this difference across subjects, we compared average click rate deviations from the baseline rate over one second preceding a false alarm for each condition (Fig. 4B). For increasing false alarms, uninformed trials had a larger increase from the baseline stimulus than informed trials. Similarly for decreasing false alarms, uninformed trials had a larger decrease from the baseline than informed trials (Supplementary Table S4).

The starting points of the RC kernels serve as an estimate for the timescale of evidence evaluation. For example, an earlier starting point indicates a longer timescale. We found that the starting points of the RC kernels for increasing trials was slightly shorter in the informed condition at -0.80 s (CI: -0.81 to -0.77) than the uninformed condition at -0.82 s (CI: -0.85 to -0.79). In decreasing trials, the start time of the informed condition was significantly shorter at -0.49 s (CI: -0.50 to -0.47 , $p < 0.05$) than the uninformed condition at -0.80 s (CI: -0.83 to -0.77). In sum, the RC kernels suggested that subjects may adjust both decision bounds and decision filters in the different conditions.

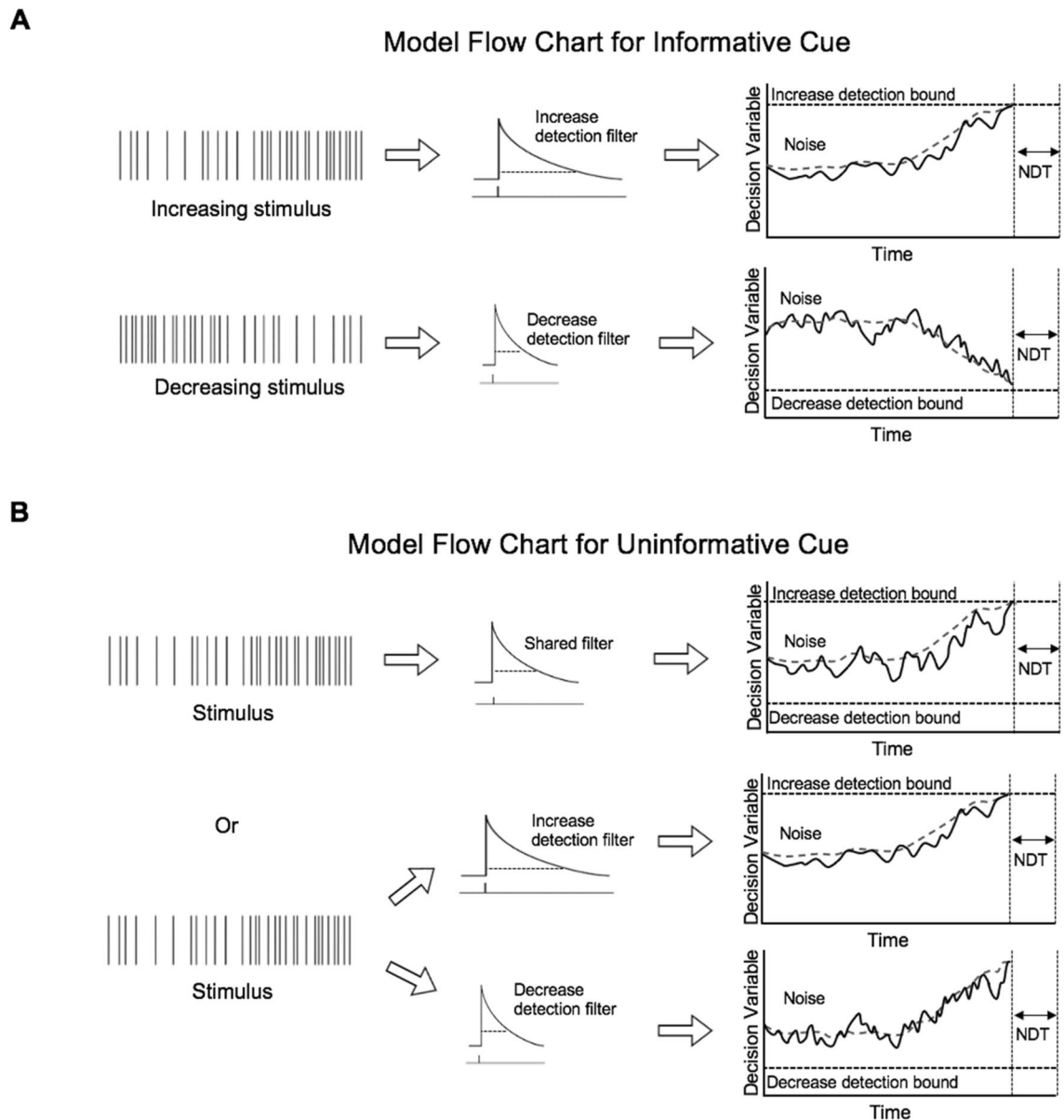


Figure 5. Model schematic. **(A)** Model schematic for informative trials. A decision filter is convolved with the stimulus to generate the decision variable. When the decision variable crosses the detection bound, a choice is made. **(B)** Model schematic for uninformative trials. The stimulus may be either increasing or decreasing. Either a shared filter can be used regardless of increasing or decreasing stimuli (top) or two different filter widths can be used simultaneously (bottom).

Model-based analysis. To further characterize the behavioral adjustments suggested by our model-free analyses, we fit subject responses to a sequential sampling behavioral model that parametrizes both decision filters and bounds. In general, these models predict a response time for a given stimulus by estimating the value of a decision variable over time and responding when the decision variable crosses a detection bound. Specifically, our model estimates the decision variable by convolving stimulus clicks with an exponential detection filter, and computes the response time by adding a non-decision time component to the bound crossing time to account for sensory and motor delays in responding. Thus, our model includes five free parameters: filter width, noise amplitude, bound, non-decision time, and non-decision time variability (see methods). In total, three separate models were fit to trials grouped by cue (informed increase, informed decrease, and uninformed) from both individual and combined subject data. Informed models contained a single process to detect either increasing or decreasing stimuli depending on the cue (Fig. 5A), whereas uninformed models contained two separate processes to detect both increasing or decreasing stimuli (Fig. 5B). Each model was fit to the behavioral data by maximizing the likelihood that the subjects' response times could be produced by the model parameters.

We used this modeling framework to better understand how subjects alter their decision making process when they are informed versus uninformed about the direction of stimulus change. To this end, we asked whether behavior in uninformed conditions was better explained by a model with separate or shared detection filters

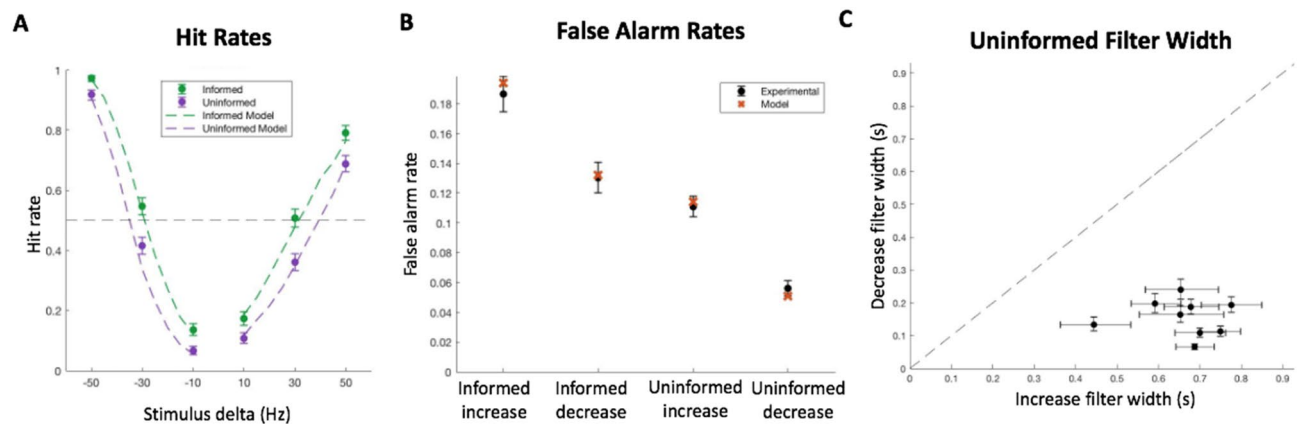


Figure 6. Model fit. (A) Experimental combined subject data with performance plotted as a function of change in click rates compared to the estimated model performance. (B) Experimental false alarm rates for combined subject data in each trial condition compared to estimated model false alarm rates. (C) Unshared filter width parameter pairs in uninformed models fit to individual subject data. All error bars indicate 95% confidence intervals. $N=9$ subjects.

(Fig. 5B). To best characterize the difference between these two model configurations, we fit the behavioral data in the uninformed condition to two separate single-bound models: one where the models had independent filter widths (λ) and another where they shared a single filter width. By comparing the two conditions in this way, we were able to isolate the effects of constraining the filter widths because any differences between the two fits are not explainable by the other model parameters.

Model fit improvement was characterized by comparing log model evidence lower bound (ELBO) values between each fit. ELBO values are analogous to Bayes factors and naturally penalize overfitting, which allows us to compare each model even though they contain different numbers of parameters^{36,37}. The ELBO metric favored the model with unshared filter width parameters for both the combined behavioral data (ELBO difference = 341) and for all nine subjects (mean ELBO difference = 41; Range: 5–110). Similar results were found using Bayesian Information Criterion and Akaike Information Criteria for model comparison. We tested a variety of model formulations where the mean non-decision time (NDT), the NDT variance, and the noise term were additionally shared either between the two models for the uninformed condition or between models for all conditions. These variants yielded results consistent with those reported here, which are produced from the formulation with only the NDT variance shared across all models.

In addition to improved ELBO values, the model with separate filter widths produced psychometrics that matched the experimental data very well without explicitly fitting the model to these metrics (Fig. 6A, B, Supplementary Figures S5–S7). This further suggests that subjects used separate decision making processes that operated simultaneously with distinct timescales of evidence evaluation. Indeed, fits on individual subject data consistently revealed shorter filter widths for detecting decrease changes than those for detecting increase changes (Fig. 6C).

To ensure that the longer filter width to detect increases does not simply arise from input depression during faster click rates, we additionally incorporated input depression into our model using the maximal depression terms found in a prior human psychophysics auditory Poisson clicks task: a depression scale factor of 0.45 and time constant of 20 ms³⁵. We found that with depression, the best fit filter widths were even wider for increases than decreases compared to the model without depression (Supplementary Figure S8). Therefore, input depression cannot account for the differences in filter widths that we found. In total, all of our model-based analyses support our conclusion that behavior in the uninformed condition was better explained by a model configuration with separate decision filter widths, suggesting that subjects simultaneously used distinct timescales of evidence evaluation to make their decisions.

To better understand how subjects adjusted their decision process when cued to the nature of the change, we compared how the four major model parameters (filter width, bound, noise amplitude and non-decision time) varied between informed and uninformed conditions for each change detection process (increase or decrease). For increasing stimuli, we found that the primary change in the decision process could be described as a change in the detection bound. The model fit to combined subject data showed a significant decrease in the bound when subjects were informed (83.9 Hz versus 78.3 Hz, $p < 0.001$) (Fig. 7A, Supplementary Table S9), which is consistent with our model-free analysis. Additionally, models fit to individual subject data revealed the same trend in eight of the nine subjects (Fig. 7B). Although other parameters seemed to vary slightly between conditions when fit to the combined subject data (such as filter width, Fig. 7A), these differences were not as consistent across subjects (Fig. 7C). For decreasing stimuli, we found that the change in the decision process was more complex, as the filter width, noise, and bound all seemed to change by a substantial amount (Fig. 7D). In contrast to the increasing stimuli, the change in the filter width was consistent across subjects for decreasing stimuli, but the change in bound was not (Fig. 7E, F). Further comparing differences between increase and decrease model parameters, there seems to be a robustly different set of optimal values depending on the change direction, which can be most strongly seen in the filter width being substantially longer for increasing stimuli. This suggests that optimizing the ability to detect changes in different directions requires different parameter tunings. With this

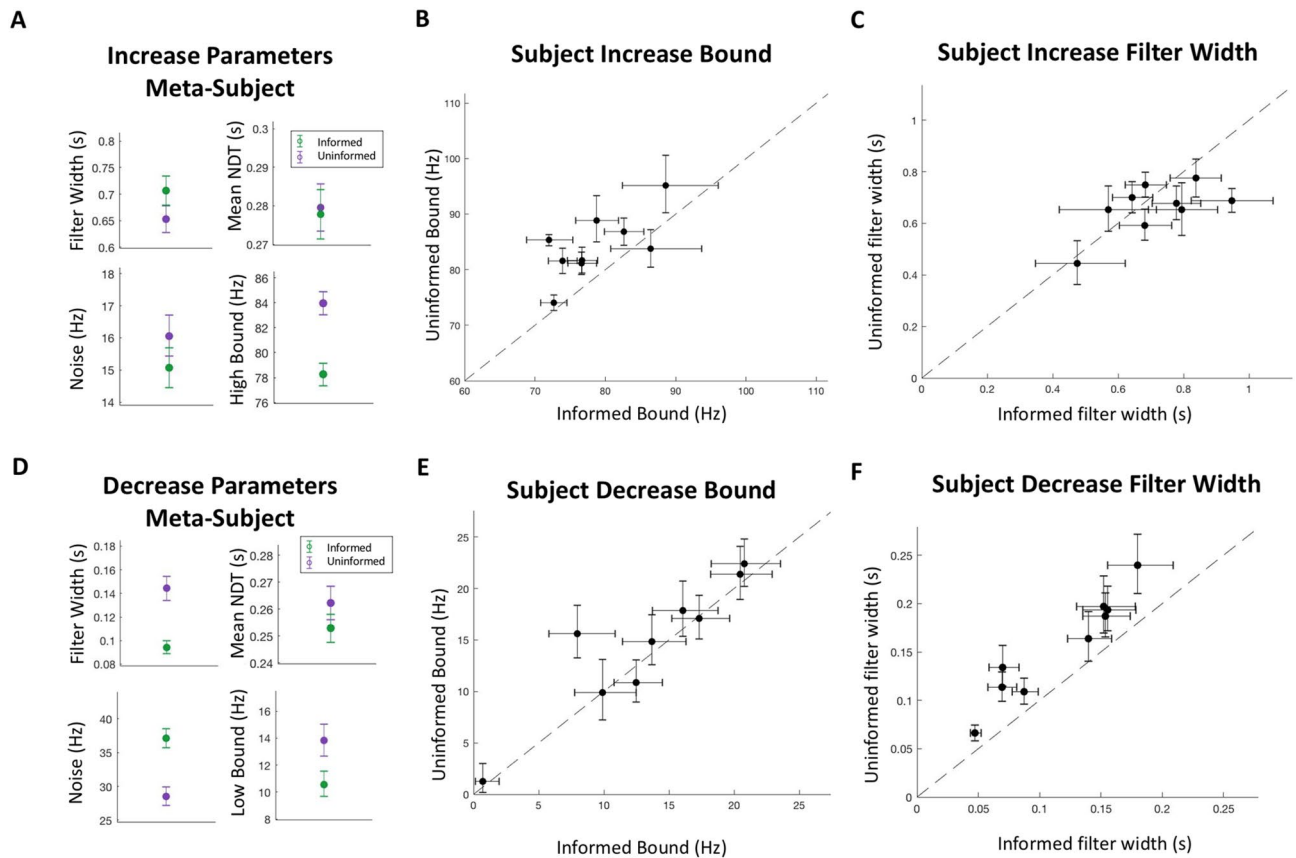


Figure 7. Model parameters (A) Increase change detection parameters for the informed and uninformed conditions fit to combined subject data. (B) Increase bound values between informed and uninformed conditions from models fit to individual subject data. (C) Increase detection filter widths between informed and uninformed conditions from models fit to individual subject data. (D) Decrease change detection parameters for the informed and uninformed conditions fit to combined subject data. (E) Decrease bound values between informed and uninformed conditions from models fit to individual subject data. (F) Decrease detection filter widths between informed and uninformed conditions from models fit to individual subject data. All error bars indicate 95% confidence intervals. $N = 9$ subjects.

understanding, we found that uninformed decrease change detection parameters were generally pulled away from their optimal informed values toward their optimal increase detection values, which can be seen especially strongly by the increased filter width (compare Fig. 7D with Fig. 7A).

To further ensure that any parameter differences were not due to the stochastic nature of the model fitting process, we performed a parameter recovery exercise wherein we fit our model to synthetic data generated from a range of model parameters, and then compared these ‘true’ values to the resulting best-fit values. We found that the parameter values for filter width, noise, and bound were recovered with high precision and little bias for all models across conditions (Supplementary Figure S10). For mean non-decision time, there was a slight trend for recovered parameters to overshoot known values by 15–20 ms, however this did not adversely affect the recovery of the other parameters, and it was a consistent bias across conditions.

Discussion

We used a stochastic auditory change point detection task to investigate how knowledge about the space of possible changes affects human change point detection. In our task, changes to the generative click rate in either direction occurred at unpredictable times, and subjects received either an uninformative or an informative cue about the direction of the impending change. Through model-free and model-based analysis we found two primary conclusions: subjects use knowledge about the dimension of the change to narrow their ‘decision criteria’, and subjects are able to process the same sensory information over multiple distinct timescales.

From the behavioral results, subjects were able to discern changes at a greater rate and in a shorter amount of time when they were informed of the direction of the change (Fig. 2A, C). However, this improved performance came at the cost of an increase in false alarms (Fig. 3). When comparing the false alarm trials between conditions with psychophysical reverse correlation analyses, we found that subjects responded to smaller deviations from the baseline click rate when they were informed about the direction of the change (Fig. 4). This suggests that subjects narrowed their ‘decision criteria’ when they were able to reduce the number of attended dimensions in the stimulus stream. By reduced ‘decision criteria’, we imply that subjects required less evidence to commit to their decision, which meant subjects were able to react faster and were more likely to respond to changes when

they did happen. Further supporting this claim, a reduced evidence threshold would make subjects more likely to falsely respond to stimulus noise, which we also found to be the case. The question still remained, however, how this reduced bound can be implemented. Intuitively, we can imagine three main non-mutually exclusive implementations that could account for this behavior: improving the ability to remember past stimuli to compare with current stimuli to better discern changes, reducing the amount of evidence required to suggest that the new stimuli are different from past stimuli, and increasing the ability to filter out stimulus noise while making this comparison. To better understand how these implementations might explain our behavioral results, we developed a model that incorporates terms for all of these scenarios (filter width, bound, and noise, respectively) and fit it to our data.

When comparing best-fit model parameters between informed and uninformed conditions, we generally found what looked to be different implementations depending on the direction of the change. One consistency between directions is that our models suggest subjects were not able to improve the signal-to-noise ratio of the underlying evidence, as characterized by our noise term. Rather than being able to directly increase perceptual sensitivity, subjects seemed to improve their change detection capacity with a mixture of adjustments to their decision bounds and timescales of evidence evaluation. This is consistent with other studies suggesting an important role of changes in the decision criterion for improvements in performance that occur with attentional cues^{28,29}. This is important because improvements in accuracy with cuing are often ascribed to attentional processes sometimes thought to improve perceptual sensitivity^{20,21}. We do not rule out that improved signal to noise can be important for change detection behavior, but it is interesting that we find no evidence for it, even with the cuing manipulation that we perform. An implementation that was unique for cues informing increases was that subjects reduced their decision bound, consistent with the inferences drawn from our model-free analyses. This suggests that subjects reduce the amount of evidence required to commit to the decision when there are fewer alternatives to detect and is reminiscent of the reductions in decision bounds found in discrimination tasks with fewer alternatives to consider⁴⁰. Supporting this claim, neural correlates that are potentially consistent with higher decision thresholds when considering multiple alternatives⁴¹ have been reported in parietal cortex^{40,42,43}, frontal cortex^{44,45}, and subcortical regions involved in action selection^{46,47}. However, none of those studies involved change point detection, so it would be interesting to explore whether similar neural mechanisms may be involved in our task. For decreasing changes, the model suggested that all three decision criteria parameters were adjusted when subjects were informed. From our results, we are not able to conclude that any one change was more important than the other, but in general each of the parameters moved in a direction toward their best fit values for cued increase trials when subjects were uninformed. This result seems to suggest that there may be some resource sharing between the two processes when both change directions were attended to simultaneously.

When comparing parameters between informed conditions, we found that there were distinct differences in the best fit values, especially in regard to filter width. We speculate that this difference may relate to an important asymmetry for detecting increase changes versus decrease changes in our task, as can be seen in the behavioral psychometrics. In terms of our model, the decision variable moves closer to the increase decision bound in a stepwise manner with each click, and may therefore be more susceptible to noise and explain higher false alarm rates for detecting increases. In contrast, the decision variable smoothly decays toward the decrease decision bound in the absence of clicks and may therefore be more robust to noise and explain lower false alarm rates for detecting increases. This asymmetry can best be seen in the model's filter width terms. In order to respond quickly to decreases, it makes sense that the decision value should decay toward the decrease bound faster than when detecting increases, which we found to be the case in our model fits. Additionally, a longer filter width smooths the decision value's upward trajectory thereby moderating the influence of stimulus noise, such as short bursts of clicks. With this understanding of the differences between increase and decrease detection processes, we investigated whether the data from the uninformed condition could be best described by two detection processes with a shared filter width, or by two completely separate detection processes, each with their own timescale of evidence evaluation. We found that the latter best explained our data, which suggests subjects were able to evaluate the same stimuli over multiple timescales simultaneously.

The idea of distinct processes for increase and decrease change detection is supported by prior work that suggests behavioral performance on decision making tasks with multiple alternatives is best explained by parallel decision processes⁴¹. The additional finding that each process employs distinct timescales of evidence evaluation on the same sensory stimulus closely mirrors results from a previous experiment in our lab, where confidence in a change point detection had a longer timescale of evidence evaluation than the timescale required to detect the change¹². Our results also fit nicely with prior studies, from our lab and others, that show subjects can flexibly adjust their timescales of evidence evaluation on the same sensory stimuli based on task demands. For example, in the visual and auditory domains, humans adjust their timescale of evidence evaluation to match the distribution of signal durations and timing they experience^{10,39,48–50}. Humans can also adjust their timescales of evidence evaluation when classifying the location of visual stimuli, using a shorter timescale in more volatile environments and a longer timescale in more stable environments⁹. Likewise, both humans and rats can adjust their timescales of evidence evaluation in a similar manner when discriminating the current state of a changing sensory stimulus, using a shorter timescale when the stimulus changes more often and a longer timescale when it changes less often^{8,9}. A distinct difference between the prior studies and ours is that all of these previous findings derived from tasks that involved a timescale of evidence evaluation that varied over blocks of trials, giving subjects time to adapt to a new optimal evaluation period. In contrast, we demonstrate that subjects were able to rapidly change their timescales on a trial-by-trial basis, and additionally, were able to simultaneously process the same sensory stream with multiple timescales when needed.

Recent work has shown that even in situations where sensory evidence for multiple decisions can be acquired in parallel along distinct stimulus dimensions, there can exist a bottleneck that limits updating of decision variables to one at a time⁵¹. In that work, decision variables were found to be held in two distinct buffers with

time-multiplexed updating of each in turn. Our results do not rule out the possibility of a similar bottleneck for our detection task with buffers for decision variables for increases and decreases that each involve different timescales of evidence evaluation. In describing the processing as “simultaneous”, we therefore do not rule out high frequency time-multiplexed components of the processing, and our experiments were not designed to address this question. It is also possible that detection of increases and decreases involves decision variables based on different sources of evidence. For example, periods of silence may inform decrease detections while clicks themselves may inform increase detections.

Related to this, we do not know if perceptions for lower and higher click rates in our task are supported by overlapping or distinct processing channels. Periodic sounds within the frequency range of the click rates of our task have percepts that range from flutter to those having a character of pitch³⁰, which may imply different processing channels that depend on frequency. Regardless of the source of sensory evidence for the detection decisions, our results suggest the capacity to simultaneously evaluate evidence along multiple timescales.

This capacity provides insight into potential underlying neural mechanisms. It seems to rule out mechanisms that allow only a single timescale of evaluation for a particular source or stream of evidence, even if that timescale is tunable. Instead, any proposed neural mechanism must be able to support at least two distinct timescales. One potential mechanism would involve distinct neural populations with separately tunable timescales that process the same stream of evidence in parallel. It is also possible that interactions between distinct decision variables for increase and decrease detection may play a role in controlling their timescales. For example, previous work has demonstrated the importance of inhibitory interactions between leaky competing accumulators in determining timescales of evidence evaluation^{7,48,52}. Another contrasting mechanism involves a single population of neurons that contains a heterogeneity of timescales^{53–57}, which is suggested by a number of network architectures^{58,59}. One advantage of these architectures is that they readily support evidence evaluation across multiple timescales without information loss or re-tuning of the networks involved in the evaluation process. However, our results do not provide clear support for either of these implementations over the other.

Overall, our data supports the idea that parallel evaluation of the same stream of evidence can occur along distinct timescales for change detection with multiple alternatives. Furthermore, cued attention narrows the decision space and allows observers to make changes in decision processes that improve performance. These results establish three important capacities of information processing for decision making that any proposed neural mechanism of evidence evaluation must be able to support: the ability to simultaneously employ multiple timescales of evidence evaluation, the ability to rapidly adjust those timescales, and the ability to modify the amount of information required to make a decision in the context of flexible timescales. Future work will be needed to bridge the gap from these capacities to the neural mechanisms that support them.

Received: 30 November 2020; Accepted: 8 June 2021

Published online: 23 June 2021

References

- Wilson, R. C., Nassar, M. R. & Gold, J. I. Bayesian online learning of the hazard rate in change-point problems. *Neural Comput.* **22**, 2452–2476 (2010).
- Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- Brody, C. D. & Hanks, T. D. Neural underpinnings of the evidence accumulator. *Curr. Opin. Neurobiol.* **37**, 149–157 (2016).
- Liu, A. S. K., Tsunada, J., Gold, J. I. & Cohen, Y. E. Temporal integration of auditory information is invariant to temporal grouping cues. *eNeuro* <https://doi.org/10.1523/ENEURO.0077-14.2015> (2015).
- Ratcliff, R. & Smith, P. L. A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* **111**, 333–367 (2004).
- Ratcliff, R. & McKoon, G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
- Usher, M. & McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592 (2001).
- Piet, A. T., El Hady, A. & Brody, C. D. Rats adopt the optimal timescale for evidence integration in a dynamic environment. *Nat. Commun.* **9**, 4265 (2018).
- Glaze, C. M., Kable, J. W. & Gold, J. I. Normative evidence accumulation in unpredictable environments. *Elife* **4**, e08825 (2015).
- Ossmy, O. *et al.* The timescale of perceptual evidence integration can be adapted to the environment. *Curr. Biol.* **23**, 981–986 (2013).
- Johnson, B., Verma, R., Sun, M. & Hanks, T. D. Characterization of decision commitment rule alterations during an auditory change detection task. *J. Neurophysiol.* **118**, 2526–2536 (2017).
- Ganupuru, P., Goldring, A. B., Harun, R. & Hanks, T. D. Flexibility of timescales of evidence evaluation for decision making. *Curr. Biol.* **29**, 2091–2097.e4 (2019).
- Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn. Psychol.* **12**, 97–136 (1980).
- Treisman, A. M. Strategies and models of selective attention. *Psychol. Rev.* **76**, 282–299 (1969).
- Carrasco, M. Visual attention: the past 25 years. *Vision Res.* **51**, 1484–1525 (2011).
- Tolhurst, D. J., Movshon, J. A. & Dean, A. F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vis. Res.* **23**, 775–785 (1983).
- Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J. & Cohen, M. R. Learning and attention reveal a general relationship between population activity and behavior. *Science* **359**, 463–465 (2018).
- Stevens, C. F. Neurotransmitter release at central synapses. *Neuron* **40**, 381–388 (2003).
- Churchland, A. K. *et al.* Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
- Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Ann. Rev. Neurosci.* **18**, 193–222 (1995).
- Moore, T. & Zirnsak, M. Neural mechanisms of selective visual attention. *Ann. Rev. Psychol.* **68**, 47–72 (2017).
- Downer, J. D., Rapone, B., Verhein, J., O’Connor, K. N. & Sutter, M. L. Feature-selective attention adaptively shifts noise correlations in primary auditory cortex. *J. Neurosci.* **37**, 5378–5392 (2017).
- Hanks, T., Kiani, R. & Shadlen, M. N. A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *Elife* **3**, e02260 (2014).
- Heitz, R. P. & Schall, J. D. Neural mechanisms of speed-accuracy tradeoff. *Neuron* **76**, 616–628 (2012).

25. Thura, D. & Cisek, P. Modulation of premotor and primary motor cortical activity during volitional adjustments of speed-accuracy trade-offs. *J. Neurosci.* **36**, 938–956 (2016).
26. Hanks, T. D. & Summerfield, C. Perceptual decision making in rodents, monkeys, and humans. *Neuron* **93**, 15–31 (2017).
27. Smith, P. L. & Ratcliff, R. An integrated theory of attention and decision making in visual signal detection. *Psychol. Rev.* **116**, 283–317 (2009).
28. Banerjee, S., Grover, S., Ganesh, S. & Sridharan, D. Sensory and decisional components of endogenous attention are dissociable. *J. Neurophysiol.* **122**, 1538–1554 (2019).
29. Sridharan, D., Steinmetz, N. A., Moore, T. & Knudsen, E. I. Distinguishing bias from sensitivity effects in multialternative detection tasks. *J. Vis.* **14**, 16–16 (2014).
30. Nourski, K. V. & Brugge, J. F. Representation of temporal sound features in the human auditory cortex. *Rev. Neurosci.* **22**, 187–203 (2011).
31. Zylberberg, A., Barttfeld, P. & Sigman, M. The construction of confidence in a perceptual decision. *Front. Integr. Neurosci.* **6**, 79 (2012).
32. Okazawa, G., Sha, L., Purcell, B. A. & Kiani, R. Psychophysical reverse correlation reflects both sensory and decision-making processes. *Nat. Commun.* **9**, 3479 (2018).
33. Churchland, A. K. & Kiani, R. Three challenges for connecting model to mechanism in decision-making. *Curr. Opin. Behav. Sci.* **11**, 74–80 (2016).
34. Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (CRC Press, 1994).
35. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
36. Acerbi, L. & Ma, W. J. Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. *Adv. Neural Inf. Process. Syst.* (2017).
37. Acerbi, L. Variational Bayesian Monte Carlo with noisy likelihoods. *Adv. Neural Inf. Process. Syst.* **33** (2020).
38. Acerbi, L. Variational Bayesian Monte Carlo. In *Advances in Neural Information Processing Systems* Vol. 31 (eds Bengio, S. et al.) 8213–8223 (Curran Associates Inc., 2018).
39. Harun, R. et al. Timescales of evidence evaluation for decision making and associated confidence judgments are adapted to task demands. *Front. Neurosci.* **14**, 826 (2020).
40. Churchland, A. K., Kiani, R. & Shadlen, M. N. Decision-making with multiple alternatives. *Nat. Neurosci.* **11**, 693–702 (2008).
41. Churchland, A. K. & Ditterich, J. New advances in understanding decisions among multiple alternatives. *Curr. Opin. Neurobiol.* **22**, 920–926 (2012).
42. Balan, P. F., Oristaglio, J., Schneider, D. M. & Gottlieb, J. Neuronal correlates of the set-size effect in monkey lateral intraparietal area. *PLoS Biol.* **6**, e158 (2008).
43. Louie, K., Gratton, L. E. & Glimcher, P. W. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* **31**, 10627–10639 (2011).
44. Lee, K.-M. & Keller, E. L. Neural activity in the frontal eye fields modulated by the number of alternatives in target choice. *J. Neurosci.* **28**, 2242–2251 (2008).
45. Cohen, J. Y., Heitz, R. P., Woodman, G. F. & Schall, J. D. Neural basis of the set-size effect in frontal eye field: timing of attention during visual search. *J. Neurophysiol.* **101**, 1699–1704 (2009).
46. Basso, M. A. & Wurtz, R. H. Modulation of neuronal activity by target uncertainty. *Nature* **389**, 66–69 (1997).
47. Basso, M. A. & Wurtz, R. H. Modulation of neuronal activity in superior colliculus by changes in target probability. *J. Neurosci.* **18**, 7519–7534 (1998).
48. Tsetsos, K., Gao, J., McClelland, J. L. & Usher, M. Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. *Front. Neurosci.* **6**, 79 (2012).
49. Bronfman, Z. Z., Brezis, N. & Usher, M. Non-monotonic temporal-weighting indicates a dynamically modulated evidence-integration mechanism. *PLoS Comput. Biol.* **12**, e1004667 (2016).
50. Levi, A. J., Yates, J. L., Huk, A. C. & Katz, L. N. Strategic and dynamic temporal weighting for perceptual decisions in humans and macaques. *eNeuro* **5** (2018).
51. Kang, Y. H. et al. Multiple decisions about one object involve parallel sensory acquisition but time-multiplexed evidence incorporation. *Elife* **10**, e63721 (2021).
52. Bogacz, R., Usher, M., Zhang, J. & McClelland, J. L. Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 1655–1670 (2007).
53. Cavanagh, S. E., Wallis, J. D., Kennerley, S. W. & Hunt, L. T. Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice. *Elife* **5**, e18937 (2016).
54. Murray, J. D. et al. A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
55. Bernacchia, A., Seo, H., Lee, D. & Wang, X.-J. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* **14**, 366–372 (2011).
56. Scott, B. B. et al. Fronto-parietal cortical circuits encode accumulated evidence with a diversity of timescales. *Neuron* **95**, 385–398. e5 (2017).
57. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* **548**, 92–96 (2017).
58. Goldman, M. S. Memory without feedback in a neural network. *Neuron* **61**, 621–634 (2009).
59. Chaudhuri, R., Bernacchia, A. & Wang, X.-J. A diversity of localized timescales in network activity. *Elife* **3**, e01239 (2014).

Acknowledgements

We thank Preetham Ganupuru and Adam Goldring for helpful feedback and comments throughout all phases of this study.

Author contributions

All authors were involved in conception and design of the experiments. A.B., T.S., C.M., and M.R. were involved with data collection. A.B. and T.S. led the data analysis with input from T.D.H. A.B. drafted the manuscript with assistance from T.S. and T.D.H. All authors reviewed the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-92470-y>.

Correspondence and requests for materials should be addressed to T.D.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021