Sex-specific speed-accuracy tradeoffs shape neural processing of acoustic signals

Jan Clemens[1], Bernhard Ronacher[2], Michael Reichert[3]

Affiliations:
1 European Neuroscience Institute Göttingen – A Joint Initiative of the University Medical Center Göttingen and the Max-Planck Society, Grisebachstrasse 5, Göttingen, 37077, Germany
2 Behavioral Physiology Group, Department of Biology, Humboldt-Universität zu Berlin, Germany
3 Department of Integrative Biology, Oklahoma State University, Stillwater OK USA

Correspondence: clemensjan@gmail.com

Funding:
JC: DFG CL 596/1-1 (32951824), SPP 2205 (430158535)
BR: DFG Ro 547/12-1
MR: US National Science Foundation International Research Fellowship Program (IRFP 1158968)

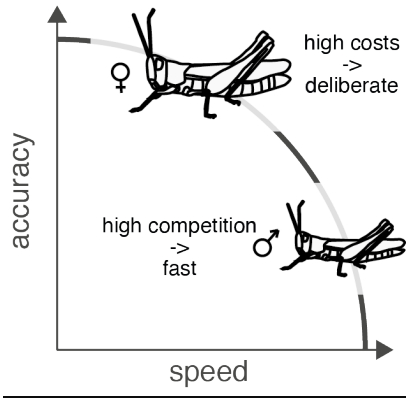
Acknowledgments:
Michael Rumpold assisted with the experiments, and members of the Behavioral Physiology Group helped with field collection and husbandry.

Contributions:
Designed experiments and collected the behavioral data – BR, MR
Fitted and analyzed the model – JC
Wrote the manuscript – JC, BR, MR

Abstract

Speed-accuracy tradeoffs – being fast at the risk of being wrong – are fundamental to many decisions and natural selection is expected to resolve these tradeoffs according to the costs and benefits of behavior. We here test the prediction that females and males should integrate information from courtship signals differently because they experience different payoffs along the speed-accuracy continuum. We fitted a neural model of decision making to behavioral data to determine the parameters of temporal integration of acoustic directional information used by male grasshoppers to locate receptive females. The model revealed that males had a low threshold for initiating a turning response, yet a large integration time constant enabled them to continue to gather information when cues were weak. This contrasts with parameters estimated for females of the same species when evaluating potential mates, in which thresholds for response were much higher and behavior was strongly influenced by unattractive stimuli. Our results reveal differences in neural integration consistent with the sex-specific costs of mate search: Males often face competition and need to be fast, females often pay high error costs and need to be deliberate.

Graphical abstract



Introduction

Sensory information is inherently noisy. Repeated sampling and integration of information over time reduces noise, and is a ubiquitous strategy in information processing and decision making (Chittka et al., 2009; Gold and Shadlen, 2007; Heitz, 2014). The neural algorithm by which ecologically-relevant sensory information is integrated is expected to be shaped by a fundamental tradeoff between the speed of decision making and the accuracy of the inferred sensory input (Gold and Shadlen, 2007; Ratcliff and Mckoon, 2008; Reddi and Carpenter, 2000). Studies of animals trained in artificial decision-making tasks show that the resolution of the speed-accuracy tradeoff depends on the relative costs of delaying the decision to accumulate more information compared to the costs of making an error based on insufficient information, the signal-to-noise ratio, and the stability of the sensory information over time (Constantinople et al., 2019; Davidson and El-Hady, 2019; Palmer et al., 2005; Piet et al., 2018; Tajima et al., 2016; Zhang et al., 2018; Zylberberg et al., 2016). However, all of these factors vary in natural environments, and among individuals, sexes and species (Chittka et al., 2009; Sponberg et al., 2015; Stöckl et al., 2017). While speed-accuracy tradeoffs and changes thereof with condition or experience have been reported for naturalistic tasks (Bonachea and Ryan, 2011; Chittka et al., 2003; Franks et al., 2003; Ings and Chittka, 2008; Wang et al., 2015), there exists little direct evidence that natural integration processes are shaped by selection, for instance from comparisons across groups expected to face different costs (Chittka et al., 2009; El-Hady et al., 2019). The processing of signals related to mate choice presents a clear instance in which selection likely favors different resolutions of the speed-accuracy tradeoff in the two sexes, which in turn are expected to result in sex differences in temporal integration: Integration processes in males should facilitate the fast decisions required for successful competition, while integration in females should be slower, but enable more accurate decisions. The existence of sex-specific circuits in the nervous system suggests that temporal integration could indeed be implemented in a sex-specific manner (Dickson, 2008; Yang and Shah, 2014), but whether the characteristics of temporal integration differ between males and females in a natural task is unclear. Here we use a drift-diffusion model (Ratcliff and Mckoon, 2008) to characterize how the nervous system accumulates sensory cues and triggers decisions in mate searching. Based on the expected costs and benefits of different integration strategies under sexual selection theory, we test predictions for how integration may differ between males and females evaluating acoustic signals of the opposite sex.

The grasshopper *Chorthippus biguttulus* provides an excellent model for studies of temporal integration because both males and females produce and respond to acoustic signals during mate searching (Fig. 1A) (Helversen, 1972), but integration and decision making strategies are expected to differ between the sexes because they are subject to different selection pressures (Helversen and Helversen, 1994). Males produce calling songs to find females; receptive females are stationary but respond with songs that facilitate mate localization by the male (Helversen, 1997). According to sexual selection theory, females are expected to be highly selective in their responses (Kuijper et al., 2012). A drift-diffusion model for temporal integration based on female response behavior indicated that this selectivity is caused by slow integration

of information across the entire calling song of a male, with a high threshold for response and very high negative weighting of unattractive song components (Clemens et al., 2017; 2014). Here, we extend the modelling approach to male mate localization, which is predicted to have very different integration characteristics. Males face high levels of competition to rapidly localize responsive females in a crowded and noisy environment (Romer, 2013), and likely pay lower costs for mating attempts with unsuitable females (Kriegbaum, 1989; Kriegbaum and Helversen, 1992). Therefore, males should be selected to respond quickly to female songs, and we predict that this will be reflected in how they integrate information to arrive at decisions, namely that they will weight sensory information highly (or, equivalently, have a lower response threshold). Furthermore, although males may therefore be more likely to respond with only partial information from the signal, we predict that integration times should nevertheless be long (i.e. at least as long as typical female songs) so that males can maximize the chances of integrating sufficient directional information from quiet female signals in noisy environments.

To test these predictions, we used a two-speaker playback design to measure male localization of artificial female songs with conflicting directional cues. We then applied a drift-diffusion model to the behavioral data to determine the parameters of temporal integration in males. The drift-diffusion model corresponded very well with males' decisions, and the model's parameters matched our predictions of long temporal integration times and a low threshold for response, which contrasts with the parameters determined for female behavior using the same model. This reveals sex-specific differences in the neural processing of sexual signals consistent with predictions from sexual selection theory.

RESULTS

To infer the dynamics of sensory integration in males, we generated artificial female songs that varied in the total number of syllables and in directional cues (see Fig. S1 for all stimulus patterns). Each syllable of the song provided one of three types of directional cue: timing cues (syllables from one speaker led the other by 4 ms), level cues (syllables only broadcast from one speaker, which results in approximately an 8 dB level difference; (Helversen and Rheinlaender, 1988; Wolf, 1986)), or no directional cues (simultaneous syllables of equal amplitude from both speakers). Both timing and level cues elicit orientation responses in male grasshoppers. Short (3-5 syllables) and long (8-12 syllables) songs were used to probe decision making in the presence of limited and abundant information, respectively. The number and location of the different directional cues within the song was systematically varied (Fig. S1), which was critical for calibrating the model parameters (Stine et al., 2020). For instance, responses to songs with serially conflicting directional information in which syllables at the beginning of the song indicated a female in the direction of one speaker and those at the end of the song indicated a female in the opposite direction reveal over how many syllables males integrate and when decisions are fixed. The stimuli were played to mature males using a two-speaker playback device in which speakers were placed on either side of the male at an angle of 90 degrees from the midline (Fig 1B). Males exhibit a stereotyped turning response in the direction of a female song (lateralization; (Helversen, 1997)) and the male's response to ten

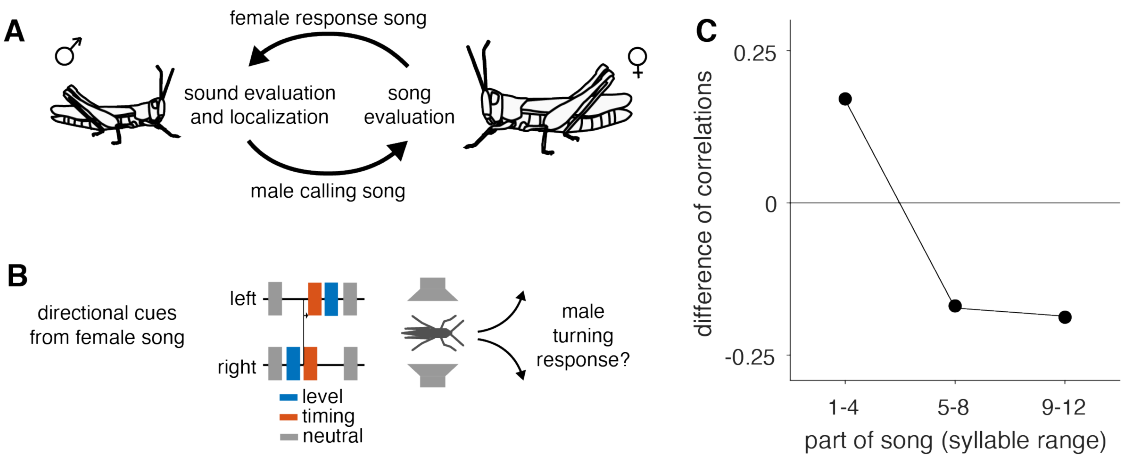


Figure 1 – Performance of a simple averaging model compared to male behavior
A Bidirectional acoustic communication during mate search in the grasshopper *Chorthippus biguttulus*.
B Schematic of the paradigm – two speakers were placed on either side of the male and artificial female song is played, and the direction of the male turning response is scored. Directional cues are provided by each syllable and arise from level differences (sound on one speaker only, blue boxes) or timing differences (sound on one speaker delayed by 4ms, red boxes). Both cue types are known to elicit turning responses in males.
C Difference in the correlation of different thirds of the 12 syllable songs observed in behavior and estimated from the stimulus statistics. The beginning tends to be more, the middle and end less influential on behavior than expected from the stimulus statistics. See Fig. S2 for details and N stimuli.

presentations of the stimulus was quantified as the proportion of turns towards the reference speaker out of all turns made (see Methods and (Reichert and Ronacher, 2019)).

Noisy integration to threshold explains turning behavior.
Looking at the correlation of the cues in different parts of the song with behavior we find that the beginning of the song influences behavior more than expected (Fig. 1C, S2). This suggests an integration process in males that does not always consider information from the full female song but instead fixes decisions rapidly and dynamically with the available sensory evidence (Reichert and Ronacher, 2019). To account for this finding, we fitted a model of dynamical, noisy cue integration to threshold – a so-called drift-diffusion model (Fig. 2A). In the model, the cues from each syllable are weighted and assigned a sign based on the direction they indicate. The weighted cues are then integrated with an integration timescale τ, which determines the "leakiness" of integration, with a value of infinity corresponding to perfect integration with no forgotten information, and smaller values corresponding to forgetting of information that came before that time interval. Noise σ is added to the integrated sensory information from each syllable, and the decision is fixed when a decision threshold of either $+\theta$ or $-\theta$ is crossed, indicating the decision to turn towards or away from the reference speaker, respectively. If the threshold is not crossed before the end of the song, the decision is made based on the sign of the integrated information at the end of the song. An urgency gain parameter was included to account for the possibility that sensory weights increase or decrease with time (Cisek et al.,

2009; Tajima et al., 2016); increased urgency may be expected for males that need to localize receptive females quickly upon receipt of evidence that one is present. We fitted the model using a genetic algorithm and evaluated performance using cross-validation (see Methods).

Model comparison (Akaike Information Criterion) revealed that the simplest model that explained our data is a drift-diffusion model with an infinite integration time in which timing and level cues had different weights, but their weights did not change over time (i.e. an urgency gain of zero) (Table 1, Fig. 2B, C, S3). We consider this as the best fit model in discussions below. A threshold-less model that simply averaged directional cues with identical weights for both cue types across the entire song performed worse (Fig. 2C) as did a drift-diffusion model variant with identical weights for both cue types (Table 1). These simpler models performed well on average (Table 1) because for many stimuli in our dataset, the average cue still predicted the behavior well. However, the performance gap between these models and our best fit model was much higher for stimuli with conflicting or mixed cues, for which correct weighting and stopping of integration after threshold-crossing were crucial to predict males' behavior (Fig. 2C). Adding even more complexity to the model with the addition of leaky integration or an urgency parameter did not improve performance (Table 1). Model parameters were very similar for all of the fitted variants of the drift-diffusion model, indicating that our results are robust to changes in model complexity.

Model variant	r^2	**AIC**	σ	θ	w_{timing}	w_{level}	τ	$\gamma_{urgency}$
Simple averaging single cue	0.75	138	0	∞	1	1	∞	0
Perfect integration, single cue	0.80	107	1.98 ± 0.05	5.81 ± 0.25	1	1	∞	0
Perfect integration, two cues	0.86	49	2.25 ± 0.06	7.14 ± 0.29	1	1.65 ± 0.04	∞	0
Leaky integration, two cues	0.86	53	1.89 ± 0.06	5.88 ± 0.31	1	1.67 ± 0.05	24.58 ± 4.81	0
Perfect integration, two cues, urgency	0.85	60	2.32 ± 0.12	7.11 ± 0.77	1	1.68 ± 0.07	∞	0.026 ± 1.01

Table 1 - Model comparison and parameters.
All values indicate the mean ($\pm$ standard deviation) of parameter estimates over the 81 cross-validation runs. r^2 - cross-validated coefficient of determination, AIC - Akaike information criterion (smaller is better), σ - noise level, θ - decision threshold, w_{timing} & w_{level} - weight for timing and level cues (w_{timing} fixed to 1 for all models), τ - integration time constant in units of syllables (fixed to ∞ for perfect integration), $\gamma_{urgency}$ - urgency gain (fixed at 0 for models without urgency). See also Fig. S3.
Models are displayed in order of increasing complexity. The simple averaging model (first row) corresponds to a perfect, noiseless integrator without a threshold. The simplest model explaining the data best (minimal AIC score) is a drift-diffusion model with perfect integration, different weights for the two cue types, and no urgency (green shading). Simpler models (yellow shading) perform worse, in particular for stimuli with conflicting or mixed cues (Fig. 2C). More complex models (red shading) do not perform better than the best fit model.

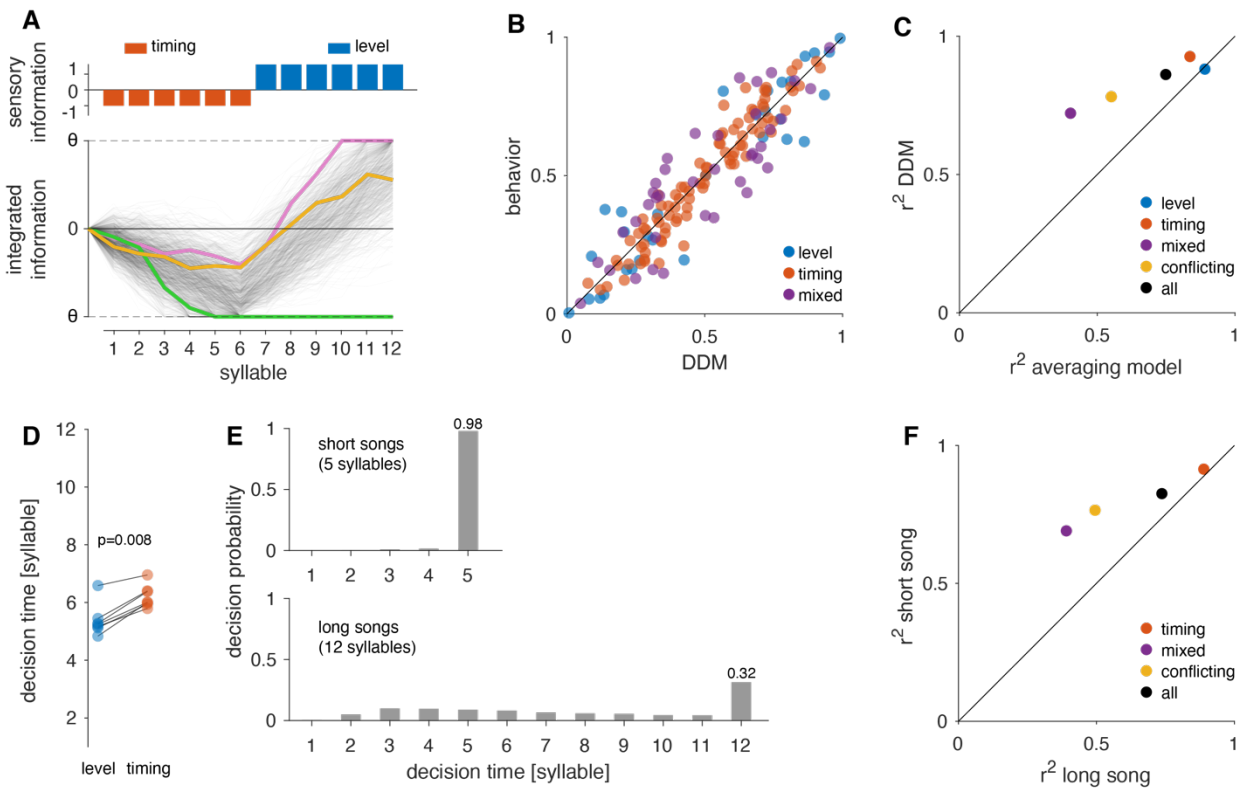


Figure 2 – A drift-diffusion model (DDM) reproduces the behavior well and reveals dynamics of temporal integration.

A DDM responses for a 12-syllable stimulus (top) with the first 6 syllables containing timing cues (red) away from the reference speaker (-), and 6 syllables with level cues (blue) towards the reference speaker (+). Each cue type is assigned a weight (height of bars, top). Stimulus information is integrated noisily and a decision towards the reference (+) or opposite (-) speaker is fixed when the decision threshold θ is crossed. Thin black lines indicate 1000 runs with independent noise realizations. Colored lines highlight example runs that cross the negative threshold (green), the positive threshold (purple), or no threshold (orange), in which case the decision is determined by the sign of the evidence at song end (+).

B Model versus behavior for songs (color indicates cue composition of the songs). Diagonal corresponds to perfect match. All points are close to that line (r^2=0.86).

C R^2 between model predictions and behavioral data for the best fitting model (DDM) compared to that of a simple averaging model for different data subsets (see legend). The best fitting model outperforms the simple averaging model in particular for stimuli with mixed (purple) and conflicting cues (yellow).

D Mean decision syllable for seven stimuli with matching patterns (lines) but level (blue) or timing (red) cues. Consistent with their higher weight in the model, level cues drive decisions by about 1 syllable earlier (p=0.008, left-sided sign test). See Fig. S4 for the patterns and decision time distributions for each of the stimuli.

E Decision times for short songs with 5 syllables (top, N=66 stimuli) and long songs with 12 syllables (bottom, N=83 stimuli). For most long songs, integration reaches threshold before song end. For nearly all short songs, integration fails to cross threshold. Numbers in the last bar indicate the probability proportion of not reaching the threshold for the two stimulus sets.

F Correlation of behavior with the average directional cue over the full song for short and long songs. The failure of threshold crossing before song end for short songs (E, top) leads to integration over the full song and a higher correlation with the average directional cue. No “level” stimulus set for this analysis since our data set did not contain such stimuli with 5 syllables (cf. C, Fig. S1).

Males integrate directional cues with long memory, cue-specific weights, low thresholds, and high noise

The best fit model indicated that males can integrate directional cues over the whole song, and even in models with a leaky integration, the time constant was estimated at 24 syllables (Table 1), which is twice as long as both the longest song in our dataset (12 syllables; 1.186 s) and a typical female song (12-15 syllables, mean±SD = 1.18±0.23 s; (Helversen and Helversen, 1975)). Thus, sensory information from the whole song has the potential to influence the localization response. The decision threshold θ of the best fit model had a value of 7.14. Level and timing cues were weighted differently, with the level cue outweighing the timing cue by a factor of 1.65. The minimum number of syllables required to cross the threshold (θ/w) was therefore 8 for timing and 5 for level cues, meaning that the stronger level cues drove faster decisions (Fig. 2D, S4). The low threshold resulted in decisions usually being fixed before the end of long 12-syllable songs (Fig. 2E), consistent with our finding that the beginning of the song is more and the end of the song is less influential than expected for observed male turning responses (Fig. 1C, S2). For the short songs in our dataset, sensory information was insufficient to drive decisions by crossing the threshold in our model, and the turning direction was determined by the value of the integrated information at the song end (Fig. 2E). This is consistent with the observation that average directional cue over the full song is more predictive of behavior for the short songs, than for the long songs (Fig. 2F). The noise level σ of the best fit model was 2.25; thus, the signal-to-noise ratios (w/σ) were 0.44 for timing cues and 0.73 for level cues. Localization cues provided by single syllables are therefore relatively noisy, and integration is indeed necessary to infer sound direction reliably.

Discussion

Our drift-diffusion model of temporal integration demonstrates that sexual selection has shaped the neural processing of acoustic stimuli to favor speedy decisions in males, in contrast to the slower, but more accurate decisions in females (Fig. 2A). The model accurately reproduced the males' localization behaviors (Fig. 2B, C) and the model parameters describe an integration process that is consistent with the pressures facing males to rapidly localize a stationary, singing female in a noisy environment (Fig. 2D-F, Table 1). The same modeling technique was previously used on females of the same species evaluating songs of potential mates (Clemens et al., 2014), and the differences in model parameters correspond with expectations of sexual selection theory that females should have a higher threshold for response and strongly avoid unattractive signal characteristics (Table S1). This is a rare demonstration of variation in temporal integration strategies associated with ecologically relevant and natural behaviors.

Decisions in males are fast for strong cues and accurate for weak cues

A drift-diffusion model of noisy dynamic integration to threshold explained male localization behavior well (Fig. 2B, C, Table 1) and constitutes a physiologically plausible integration mechanism that facilitates rapid orientation towards females by males. Speed and accuracy are both important in male localization of females, but speed is particularly critical for males to reach the female before rival competitors. We found that males had a low threshold for response: the average time to decision inferred from the model was much less than the duration of the

standard female song stimulus used in this study (Fig. 2D, E). Thus, when the evidence is strong, males can decide quickly. This corresponds with the behavior of males in localization experiments, in which they frequently turn towards a song before it ends (Helversen and Rheinlaender, 1988). The best fit model had higher weights for level cues than for timing cues (Table S1, Fig. S3B), and decisions were therefore faster with level cues than with timing cues (Fig. 2D). This does not imply that level cues always predominate over timing cues; instead the difference likely arose because of the specific values chosen for each cue: the unilateral level cues (resulting in an approximately 8 dB ILD) were expected to provide stronger directional information than the bilateral timing cues (4 ms ITD) (Helversen and Rheinlaender, 1988). Importantly, this finding implies that cues are weighted by their strength, such that decision making is accelerated when evidence is strong.

Although we expected males' decisions to be biased towards speed, accuracy is also important because mistakes in localization could cause males to move out of hearing range of the female and prevent them from finding one another (Michelsen and Rohrseitz, 1997; Reichert, 2015). The signal-to-noise ratio for a single syllable was low, and males therefore did integrate over multiple syllables, although they usually reached threshold and made a decision before they heard all of the syllables in the full song (Fig. 2D, E. Nevertheless, the model indicated that males had the capacity to integrate over a much longer time period if directional cues were weak and the threshold was not reached. The integration time constant was approximately 25 syllables, or more than twice the duration of an average female song (Table 1). Thus, when directional cues were too weak for a speedy decision, males could integrate additional sensory information, which should improve signal-to-noise ratios and ultimately lateralization accuracy (Romer, 2013). Long integration times are maladaptive when the information being integrated changes more rapidly than the integration time constant, leading to erroneous decisions (Dunlap et al., 2009; Piet et al., 2018). However, in this system long integration may have few costs because the information evaluated by males in the female song, her position, is constant prior to the turning decision because females remain stationary while singing. The integration dynamics in males therefore resolve the speed-accuracy tradeoff by allowing for flexibility in decision making: Sensory information is able to drive fast responses when it is strong, but long integration times allow accurate localization of the female in case of weak cues.

Integration of courtship signals is tuned to sex-specific costs

Our finding that male *C. biguttulus* have a low threshold for response contrasts with the results from previous studies using a similar behavioral and modeling paradigm to characterize temporal integration in females of the same species (Clemens et al., 2014; 2017). Females were tested with songs consisting of a mixture of attractive and unattractive syllables. There was a large difference between males and females in how they weighted sensory information. In females, positive cues were weak and on their own could not reach the threshold by the end of the song; in other words, females do not commit to a positive decision before the end of the song. However, negative cues (i.e., unattractive song syllables) had a much stronger weight and even a few unattractive syllables could reach the threshold for not responding. In contrast, in

males, we found that clear directional information had a strong weight and is capable of driving responses before the end of a typical female song. Both sexes had integration times that were longer than the duration of typical songs, but in females this likely serves less to enhance the signal-to-noise ratio (as we argue is the case for males), but rather to ensure the detection of unattractive elements at any point in the song, preventing them from initiating courtship with a low-quality or heterospecific male.

There are some differences in the behavioral paradigms because females were tested for a response to songs with both positive and negative information on male attractiveness, while males were tested using only attractive syllables but with varied directional cues. Pattern and directional information are extracted from the song in parallel pathways and the pattern decision then gates turning (Helversen and Helversen, 1995). Thus, turning in males in this study reflects both the attractiveness of the song syllable and the quality of directional cues. Despite these differences, both the female decision to respond and the male decision to turn signal readiness of each sex to further escalate the courtship interaction, and therefore the integration differences between males and females reflect differences in the costs and benefits of decision-making strategies affecting each sex. Future studies examining integration of stimuli with unattractive pattern information in males would further elucidate sex differences in temporal processing in this species.

Our model reveals that changes to the parameters of a common drift-diffusion threshold mechanism can replicate the behavior of both sexes. All model parameters map to biophysical properties of decision making neurons and circuits (Brody and Hanks, 2016; O'Connell et al., 2018): Sensory weights could correspond to the number and strength of synapses to an integrating neuron; The integration time constant could correspond for instance to the membrane time constant, the kinetics of intracellular calcium, or to factors that determine the dynamics of a recurrent network (Deutsch et al., 2020; Zhang et al., 2019); The decision threshold could correspond to a spiking threshold determined by the density of sodium channels at the spike initiation zone or controlled by neuromodulators (Bargmann, 2012; Marder and Thirumalai, 2002; Zhang et al., 2019). Sexual selection can act on these parameters to produce the sex-specific integration of sensory information in grasshoppers. Our results therefore point the way towards a study of the evolution of sensory processing mechanisms in realistic ecological contexts and natural behaviors.

Methods

Behavioral experiments

Behavioral data were taken from a previous study of male lateralization in response to conflicting directional cues (Reichert and Ronacher, 2019). Briefly, sexually receptive adult males of the species *Chorthippus biguttulus* were placed between two speakers that broadcast an artificial female song stimulus (Fig. 1B) in which certain syllables were manipulated to have timing or level differences between the speakers (see below). Stimuli were broadcast at 60 dB SPL. Males were presented with ten repetitions of each stimulus (most males were tested with more than one stimulus). Most stimuli contained 12 syllables to mimic a typical female song, but we also tested shortened stimuli with 8, 5 or 3 syllables to better characterize the dynamics of integration. We include data for male responses from a total of 81 stimuli (Fig. S1).

Male lateralization behavior was quantified as follows. First, the response to each stimulus repetition for each male was quantified as the proportion of turns directed towards the stimulus channel designated as the reference out of the total number of turns towards either stimulus ("0" if the male jumped away from the reference speaker, "1" when the male jumped towards the reference speaker). For some stimuli, males responded to the stimulus but jumped forward instead of towards one of the speakers and we scored these responses as "0.5". We then averaged the responses across all males tested with that stimulus (N=11-23 males tested per stimulus, median 20 males).

Most stimuli produced responses >0.5 due to the stimulus design and to balance the data, we exploited the symmetry of the assay and created mirror versions of the stimuli, for which the response was taken as one minus the response to the original stimulus. We did not mirror stimuli that had identical patterns with their mirrored version or with other stimuli in the set. This resulted in a total of 157 unique stimuli used for all analyses and for fitting the data.

Stimulus design

We generated stimuli with level differences by silencing some syllables on one speaker channel. This effectively results in an 8 dB interaural level difference (Helversen and Rheinlaender, 1988; Wolf, 1986) . Stimuli with timing differences had some syllables on one channel leading those on the other by 4 ms. Any other syllables were presented simultaneously at equal amplitude from both channels ("neutral"). We systematically varied the number and location within the song of these directional cues to generate stimuli with different amounts of directional information, and in some cases with conflicting directional information (e.g. in first part of the song directional cues favored the reference speaker, while in later parts of the song, directional cues favored the opposite speaker). See Fig. S1 for a list of all stimulus patterns used in this study. Combining this stimulus design of time-varying and sequentially conflicting cues with a neural model of decision making (see below) allowed us to infer the sensory weights and thresholds and the decision times in the absence of reaction time data, which were difficult to obtain for the natural behavior.

Comparison of correlations between the stimulus and the behavior

We assessed the correlation (squared Pearson's r) of the males' turning response with different parts of the 12 syllable stimuli, by dividing each stimulus in thirds (syllables 1-4, 5-8, 9-12), calculating the average directional cue for each third, and correlated that average with the males' turning response (Fig. S2B). To account for stimulus-intrinsic correlations, arising from regularity in the stimulus sequences (Fig. S1), we also correlated the average cue for each third with the average cue over the full song (Fig. S2A). This revealed that the middle of the song was most strongly correlated with the cues from the full song, because cue direction often changed halfway through the song (Fig. S1). The difference of the correlation obtained from the behavior and from the full stimulus indicates stimulus thirds that are more or less influential on the behavior than expected from the stimulus statistics (Fig. S2C, 1C).

Modelling

The stimulus is defined as a sequence s(t) with t=[1, … , T], T being the number of syllables in the stimulus, and a sign indicating the side of the cue relative to the reference speaker ("-1" away from the reference, "+1" towards the reference, "0" neutral).

Simple averaging model

As a baseline, the simple averaging model simply averages the directional information over the full song and the predicted response ρ is then given by the sign of that average: ρ=f(x(T+1)), with x(t+1) = x(t) + Z(s(t)), x(0)=0. The sign function Z returns -1 if s(t)<0 and +1 otherwise. To account for saturation effects, we set f to be a sigmoidal, which was fitted to minimize the mean squared error over all stimuli between the prediction ρ of the simple averaging model and the males' turning response r. However, this only marginally increased the performance of the simple averaging models (r^2 linear: 0.72, r^2 sigmoidal: 0.75).

Drift-diffusion model

For the drift-diffusion model, the integrated information x after syllable t is given by

$$x(t+1) = \begin{cases} -\theta, \text{if } x(t) < -\theta \\ \theta, \text{if } x(t) > \theta \\ x(t)/\tau + w(t)s(t) + \eta(t)\sigma, \text{otherwise} \end{cases}$$

with x(0)=0, an integration time constant τ, and a decision threshold θ. Noise η(t) was drawn at each time step from a normal distribution with zero mean and unit variance. For timing cues, w was fixed to 1.0 for all models. For level cues, w was fixed to 1.0 for models that did not differentially weight timing and level cues ("single cue" in Table 1) and optimized during fitting for models that did ("two cues"). For models with urgency gain, the sensory weight changed over time and was defined as w(t)=w*(1+(t-1)γ), with γ being the urgency gain (Cisek et al., 2009). Experiments with alternative implementations in which the urgency gain reduced the thresholds time (Tajima et al., 2016) yielded similar results of negligible gain. The decision threshold θ was sticky – once it was crossed, integration ceased and x(t) was fixed to ±θ. The predicted response, ρ, was determined by the sign of the integrated information after the last syllable, Z(x(T+1)) averaged over 1000 different instantiations of the noise η. The simple

averaging model can be considered a special case of a drift-diffusion model with $w=1$ for level and timing cues, $\sigma=0$, $\tau=\infty$, and $\theta=\infty$.

Model fitting and evaluation

The parameters of the drift-diffusion models were optimized by minimizing the mean-squared error between the predicted and the males' responses using a Genetic Algorithm ((Mitchell, 1998) see (Clemens and Ronacher, 2013) for details). To speed up convergence, upper and lower bounds were defined for all parameters: $0<w_L<10$, $1<\tau<40$, $0<\sigma<5$, $0<\theta<10$, $0<\gamma<10$. We ensured that these bounds did not affect the final parameter estimates. Fits were evaluated using leave-one-out cross-validation. That is, the model parameters were fitted on all but one stimulus (and its mirror version) and a prediction was then generated for the left-out stimuli. Doing this for all stimuli resulted in 81 parameter estimates and 157 predictions. The squared Pearson's coefficient of correlation, r^2, between the predictions and the males' responses was used to quantify model performance. Different models were compared using Akaike's Information Criterion (AIC), which penalizes models with many parameters. The AIC score is given by $AIC= 2k + n \ln(e)$, where k is the number of parameters of the model, n is the number of samples used for fitting the model, and e is the sum of squared residuals between the predicted and the male's responses: $e=\sum(r-\rho)^2$. Smaller AIC scores are better.

Supplemental Figures

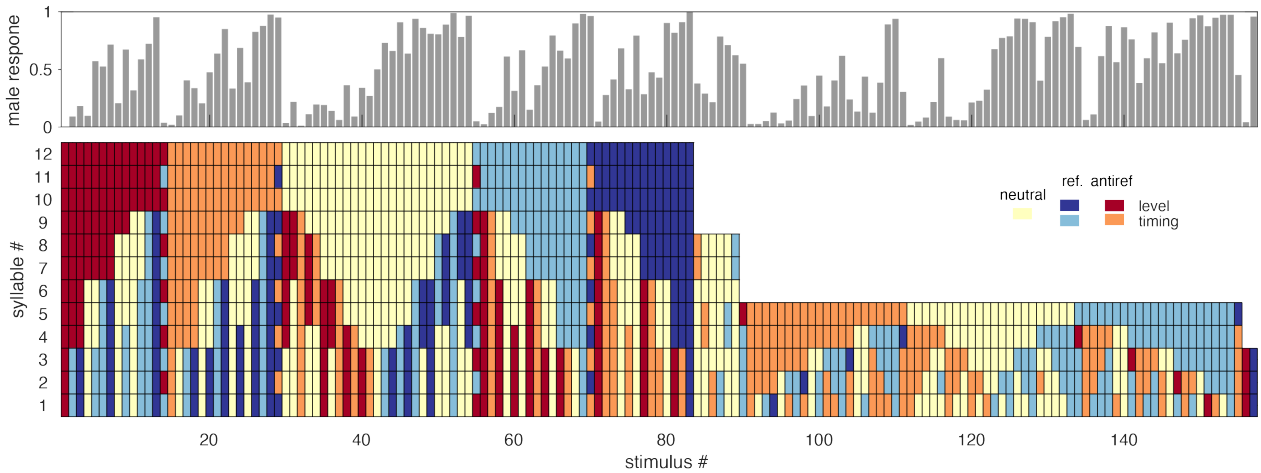


Figure S1: Stimulus patterns (bottom, color coded, see legend) and male responses (top).

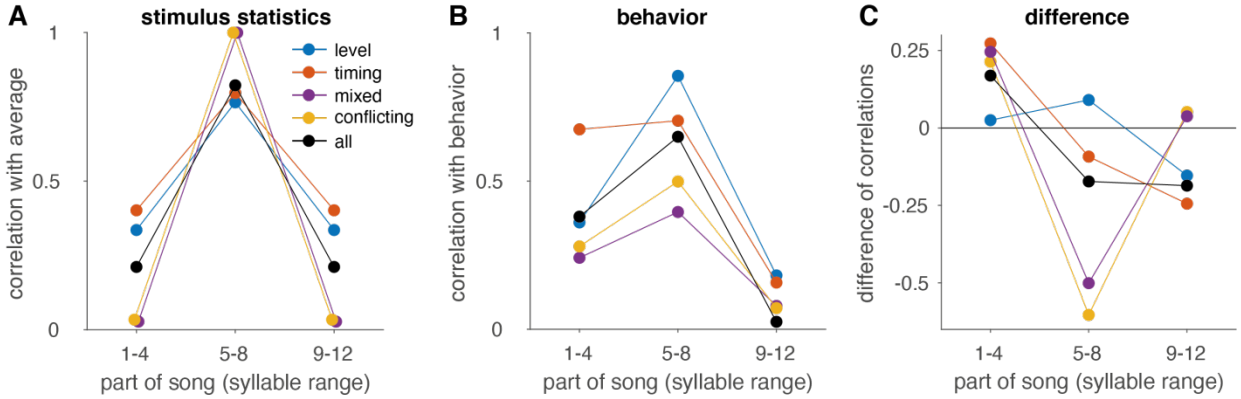


Figure S2: Correlations in the stimulus and in the behavior.
A Correlation (r^2) of the average directional cue from different thirds of the songs with 12 syllables with the average directional cue over the full song. The song's middle is most strongly correlated with the average directional cue from the stimuli than the beginning and the end because cue direction often changed half-way through the stimulus (see Fig. S1).
B Same as in A but showing the correlation between the males' turning responses and the cues from different parts of the song.
C Difference between the curves in B and A (see Fig. S1). The beginning of the song has a greater influence on the behavior (difference is positive) than expected from the correlation structure of the stimuli, suggesting that males do not simply average the directional cues over the stimulus.
Calculated for different, non-exclusive sets of stimulus with 12 syllables: all (black, N=83), level only (blue, N=29), timing only (red, N=27), mixed (purple, stimuli containing timing and level cues, N=28), conflicting (orange, stimuli with cues from both sides, N=40).

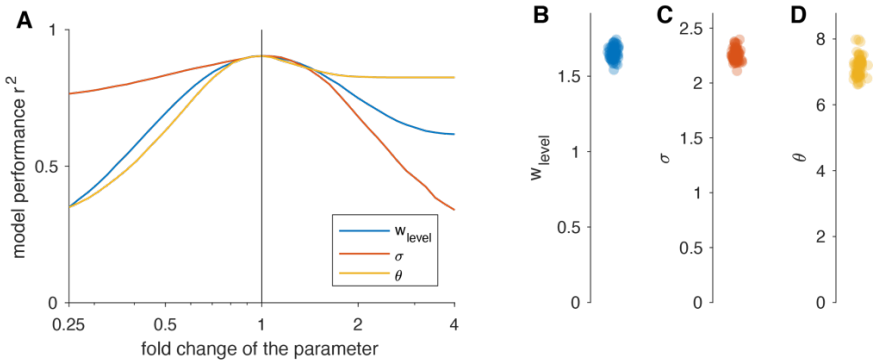


Figure S3 – Sensitivity analysis reveals that the parameters are well determined by the data.
A Performance r^2 of models with parameter perturbed around the optimal ones found in Table S1. Performance decreases for changes in the parameter, indicating that the parameter values are well determined by the data. Asymmetrical effects for σ and θ arise because increasing θ does not affect the outcome for most stimuli since the sign of the integrated evidence often determines the decision even without threshold crossing. Too much noise likely worsens predictions for stimuli while too little noise affects model predictions more weakly.
B, C, D Parameter values for weight of level cues w_{level}, the noise level σ, and the decision threshold θ are reproducible across cross-validation runs for the best fit model (compare Table 1).

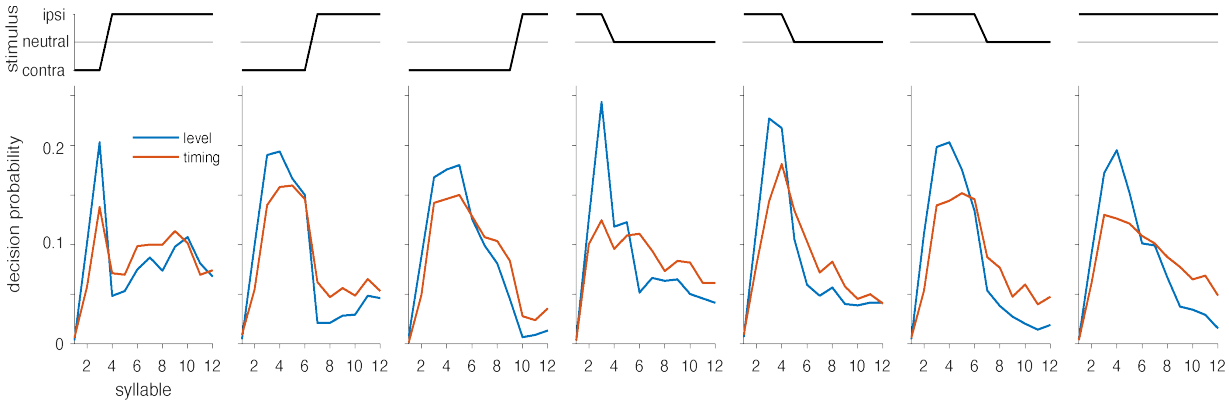


Figure S4 – Cue sequence (top) and decision time distributions (bottom) for the stimuli in Figure 2D.
Stimuli were chosen to have matching temporal cue sequences but different cue types (level – blue, timing – red).

Property	Female	Male
Cost of errors	High	Low
Cost of Slowness	Low	High
Integration property	pattern	direction & pattern
Number syllables in typical input	~30 syllables per male calling song	~12-15 syllables per female response song
Time constant τ (syllables)	Very long (213 syllables)	Long (25 syllables)
Pos. sensory evidence w	Very weak rel. to threshold (1)	Strong rel. to threshold (1 and 1.65)
Neg. sensory evidence	Very strong rel. to threshold (-47)	Not tested
Noise σ	Very high (104)	Medium (2.25)
Decision threshold θ	Very high (187)	Low (7.14)

Table S1 – Comparison of female and male costs and integration dynamics.
Female values from (Clemens et al., 2014).

References

Bargmann, C.I. (2012). Beyond the connectome: How neuromodulators shape neural circuits. Bioessays *34*, 458–465.

Bonachea, L.A., and Ryan, M.J. (2011). Localization Error and Search Costs during Mate Choice in Túngara Frogs, Physalaemus pustulosus. Ethology *117*, 56–62.

Brody, C.D., and Hanks, T.D. (2016). Neural underpinnings of the evidence accumulator. Current Opinion in Neurobiology *37*, 149–157.

Chittka, L., Dyer, A.G., Bock, F., and Dornhaus, A. (2003). Bees trade off foraging speed for accuracy. Nature *424*, 388–388.

Chittka, L., Skorupski, P., and Raine, N.E. (2009). Speed–accuracy tradeoffs in animal decision making. Trends in Ecology & Evolution *24*, 400–407.

Cisek, P., Puskas, G.A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. Journal of Neuroscience *29*, 11560–11571.

Clemens, J., and Ronacher, B. (2013). Feature extraction and integration underlying perceptual decision making during courtship behavior. Journal of Neuroscience *33*, 12136–12145.

Clemens, J., Aufderheide, J., and Ronacher, B. (2017). Relative weighting of acoustic information during mating decisions in grasshoppers indicates signatures of sexual selection. J Comp Physiol A *203*, 891–901.

Clemens, J., Krämer, S., and Ronacher, B. (2014). Asymmetrical integration of sensory information during mating decisions in grasshoppers. Proc Natl Acad Sci U S A *111*, 16562–16567.

Constantinople, C.M., Piet, A.T., and Brody, C.D. (2019). An Analysis of Decision under Risk in Rats. Current Biology *29*, 2066–2074.e5.

Davidson, J.D., and El-Hady, A. (2019). Foraging as an evidence accumulation process. PLoS Comput Biol *15*, e1007060.

Deutsch, D., Pacheco, D.A., Encarnacion-Rivera, L.J., Pereira, T., Fathy, R., Calhoun, A., Ireland, E.C., Burke, A.T., Dorkenwald, S., McKellar, C., et al. (2020). The Neural Basis for a Persistent Internal State in Drosophila Females. bioRxiv *1*, 2020.02.13.947952.

Dickson, B.J. (2008). Wired for sex: the neurobiology of Drosophila mating decisions. Science (New York, N.Y.) *322*, 904–909.

Dunlap, A.S., McLinn, C.M., MacCormick, H.A., Scott, M.E., and Kerr, B. (2009). Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments. Behavioral Ecology *20*, 1096–1105.

El-Hady, A., Davidson, J.D., and Gordon, D.M. (2019). Editorial: An Ecological Perspective on Decision-Making: Empirical and Theoretical Studies in Natural and Natural-Like Environments. Front. Ecol. Evol. *7*, 712.

Franks, N.R., Dornhaus, A., Fitzsimmons, J.P., and Stevens, M. (2003). Speed versus accuracy in collective decision making. Proceedings of the Royal Society of London. Series B: Biological Sciences *270*, 2457–2463.

Gold, J.I.I., and Shadlen, M.N. (2007). The Neural Basis of Decision Making. Annu. Rev. Neurosci. *30*, 535–574.

Heitz, R.P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. Frontiers in Neuroscience *8*, 865.

Helversen, D.V. (1997). Acoustic communication and orientation in grasshoppers. Springer

Helversen, D.V., and Helversen, O.V. (1975). Verhaltensgenetische Untersuchungen am akustischen Kommunikationssystem der Feldheuschrecken (Orthoptera, Acrididae). J Comp Physiol A *104*, 301–323.

Helversen, O.V., and Helversen, D.V. (1994). Forces driving coevolution of song and song recognition in grasshoppers. *39*, 253–284.

Helversen, von, D., and Helversen, von, O. (1995). Acoustic pattern recognition and orientation in orthopteran insects: parallel or serial processing? Journal of Comparative Physiology a: Neuroethology, Sensory, Neural, and Behavioral Physiology *177*, 767–774.

Helversen, von, D. (1972). Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke Chorthippus biguttulus (Orthoptera, Acrididae). Journal of Comparative Physiology a: Neuroethology, Sensory, Neural, and Behavioral Physiology *81*, 381–422.

Helversen, von, D., and Rheinlaender, J. (1988). Interaural intensity and time discrimination in an unrestraint grasshopper: a tentative behavioural approach. J Comp Physiol A *162*, 333–340.

Ings, T.C., and Chittka, L. (2008). Speed-Accuracy Tradeoffs and False Alarms in Bee Responses to Cryptic Predators. Current Biology *18*, 1520–1524.

Kriegbaum, H. (1989). Female choice in the grasshopper Chorthippus biguttulus. Naturwissenschaften *76*, 81–82.

Kriegbaum, H., and Helversen, von, O. (1992). Influence of Male Songs on Female Mating Behavior in the Grasshopper Chorthippus biguttulus (Orthoptera: Acrididae). Ethology *91*, 248–254.

Kuijper, B., Pen, I., and Weissing, F.J. (2012). A Guide to Sexual Selection Theory. Annu. Rev. Ecol. Evol. Syst. *43*, 287–311.

Marder, E., and Thirumalai, V. (2002). Cellular, synaptic and network effects of neuromodulation. Neural Networks *15*, 479–493.

Michelsen, A., and Rohrseitz, K. (1997). Sound localisation in a habitat: an analytical approach to quantifying the degradation of directional cues. Bioacoustics *7*, 291–313.

Mitchell, M. (1998). An Introduction to Genetic Algorithms (Complex Adaptive Systems) (A Bradford Book).

O'Connell, R.G., Shadlen, M.N., Wong-Lin, K., and Kelly, S.P. (2018). Bridging Neural and Computational Viewpoints on Perceptual Decision-Making. Trends in Neurosciences *41*, 838–852.

Palmer, J., Huk, A.C., and Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. Journal of Vision *5*, 1–1.

Piet, A.T., El-Hady, A., and Brody, C.D. (2018). Rats adopt the optimal timescale for evidence integration in a dynamic environment. Nat Commun *9*, 1–12.

Ratcliff, R., and Mckoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural Computation *922*, 873–922.

Reddi, B.A., and Carpenter, R.H. (2000). The influence of urgency on decision time. Nature Neuroscience *3*, 827–830.

Reichert, M.S. (2015). Effects of noise on sound localization in male grasshoppers, Chorthippus biguttulus. Animal Behaviour *103*, 125–135.

Reichert, M.S., and Ronacher, B. (2019). Temporal integration of conflicting directional cues in sound localization. J. Exp. Biol. jeb.208751.

Romer, H. (2013). Masking by Noise in Acoustic Insects: Problems and Solutions. Animal Communication and Noise 33–63.

Sponberg, S., Dyhr, J.P., Hall, R.W., and Daniel, T.L. (2015). Luminance-dependent visual processing enables moth flight in low light. Science *348*, 1245–1248.

Stine, G.M., Zylberberg, A., Ditterich, J., and Shadlen, M.N. (2020). Differentiating between integration and non-integration strategies in perceptual decision making. eLife *9*, 8213.

Stöckl, A.L., Kihlström, K., Chandler, S., and Sponberg, S. (2017). Comparative system identification of flower tracking performance in three hawkmoth species reveals adaptations for dim light vision. Philosophical Transactions of the Royal Society B: Biological Sciences *372*, 20160078.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal policy for value-based decision-making. Nat Commun *7*, 12400.

Wang, M.-Y., Brennan, C.H., Lachlan, R.F., and Chittka, L. (2015). Speed–accuracy trade-offs and individually consistent decision making by individuals and dyads of zebrafish in a colour discrimination task. Animal Behaviour *103*, 277–283.

Wolf, H. (1986). Response patterns of two auditory interneurons in a freely moving grasshopper (Chorthippus biguttulus L.). Journal of Comparative Physiology a: Neuroethology, Sensory, Neural, and Behavioral Physiology *158*, 689–696.

Yang, C.F., and Shah, N.M. (2014). Representing Sex in the Brain, One Module at a Time. Neuron *82*, 261–278.

Zhang, S.X., Miner, L.E., Boutros, C.L., Rogulja, D., and Crickmore, M.A. (2018). Motivation, Perception, and Chance Converge to Make a Binary Decision. Neuron 1–20.

Zhang, S.X., Rogulja, D., and Crickmore, M.A. (2019). Recurrent Circuitry Sustains Drosophila Courtship Drive While Priming Itself for Satiety. Current Biology *0*.

Zylberberg, A., Fetsch, C.R., Shadlen, M.N., and Frank, M.J. (2016). The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. eLife *5*, e17688.