The speed and accuracy of perceptual decisions in a random-tone pitch task

**Martijn J. Mulder · Max C. Keuken ·
Leendert van Maanen · Wouter Boekel ·
Birte U. Forstmann · Eric-Jan Wagenmakers**

Abstract Research in perceptual decision making is dominated by paradigms that tap the visual system, such as the random-dot motion (RDM) paradigm. In this study, we investigated whether the behavioral signature of perceptual decisions in the auditory domain is similar to those observed in the visual domain. We developed an auditory version of the RDM task, in which tones correspond to dots and pitch corresponds to motion (the random-tone pitch task, RTP). In this task, participants have to decide quickly whether the pitch of a "sound cloud" of tones is moving up or down. Stimulus strength and speed–accuracy trade-off were manipulated. To describe the relationship between stimulus strength and performance, we fitted the proportional-rate diffusion model to the data. The results showed a close coupling between stimulus strength and the speed and accuracy of perceptual decisions in both tasks. Additionally, we fitted the full drift diffusion model (DDM) to the data and showed that three of the four participants had similar speed–accuracy trade-offs in both tasks. However, for the RTP task, drift rates were larger and nondecision times slower, suggesting that some DDM parameters might be dependent on stimulus modality (drift rate and nondecision time), whereas others might not be (decision bound). The results illustrate that the RTP task is suitable for investigating the dynamics of auditory perceptual choices. Future studies using the task might help to investigate modality-specific effects on decision making at both the behavioral and neuronal levels.

Keywords Perceptual decision making · Random-dot motion task · Random-tone pitch task · Response time models · Auditory

M. J. Mulder (✉) · M. C. Keuken · L. van Maanen · W. Boekel · B. U. Forstmann · E.-J. Wagenmakers
Cognitive Science Center Amsterdam, University of Amsterdam, Nieuwe Achtergracht 129,
1018 TV Amsterdam, The Netherlands
e-mail: m.j.mulder@uva.nl

The field of perceptual decision making investigates how decisions are made on the basis of noisy sensory information (see Summerfield & Tsetsos, 2012, for a review). When people make perceptual decisions, it is generally assumed that sensory evidence for each of the alternatives is collected until a boundary is reached and a course of action is chosen. Typically, experiments designed to study perceptual decision making involve a two-alternative forced choice paradigm. A popular example of such a paradigm is the random-dot motion (RDM) task (e.g., Britten, Shadlen, Newsome, & Movshon, 1992; Forstmann et al., 2008; Gold, 2003; Gold & Shadlen, 2000; Hanks, Ditterich, & Shadlen, 2006; Heekeren, Marrett, Ruff, Bandettini, & Ungerleider, 2006; Morgan & Ward, 1980; Mulder et al., 2010; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Newsome & Paré, 1988; Palmer, Huk, & Shadlen, 2005; Roitman & Shadlen, 2002; van Ravenzwaaij, Mulder, Tuerlinckx, & Wagenmakers, 2012; Watamaniuk & Sekuler, 1992). In this task, participants have to decide as quickly and accurately as possible whether a "cloud" of dots appear to move to the left or the right. The paradigm has proven to be extremely useful for the study of perceptual decisions, since manipulations of the decision process can be implemented easily. For example, difficulty can be manipulated on a continuous scale by simply changing the number of coherently moving dots (e.g., Palmer et al., 2005). Another example involves a manipulation of the time period in which the participant is allowed to decide. Such a

manipulation is associated with the typical signatures of changes in the response time (RT) and accuracy of the perceptual choice: By stressing the importance of speed, choices will be faster, but more prone to errors (a speed–accuracy trade-off; e.g., Forstmann et al., 2008; Palmer et al., 2005).

Sequential-sampling models can describe the underlying mechanism of the decision process. These models, including the drift diffusion model (DDM; Ratcliff, 1978), are based on the assumption that noisy sensory evidence accumulates until a boundary is reached (Fig. 1A; for reviews, see Bogacz, 2007; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Wagenmakers, 2009). The model permits data to be decomposed into parameters that map onto latent psychological processes. For example, the rate of evidence accumulation (drift rate) depends on the quality of the stimulus and reflects the difficulty of a choice, whereas the boundary separation reflects the trade-off between the speed and accuracy of the perceptual decision. Furthermore, the DDM has proven to be neurobiologically plausible, as studies with human and nonhuman primates have shown neural correlates of the different components of the model (see Gold & Shadlen, 2007; Heekeren, Marrett, & Ungerleider, 2008). For instance, the firing rates of neurons in the monkey lateral intraparietal
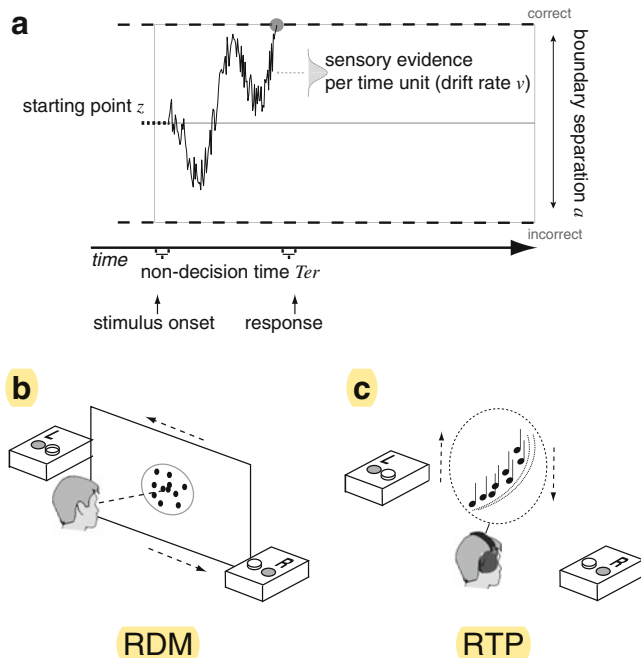
Fig. 1 Model and tasks. (A) The drift diffusion model assumes that noisy information is accumulated until a boundary is reached (a). Drift rate (v) reflects the quality of the sensory evidence, and nondecision time (T_{er}) reflects the time other than the decision time (e.g., processing of sensory information and/or execution of a motor response). (B) The classic version of the random-dot motion (RDM) task, in which participants have to decide whether a cloud of dots appears to move to the left or to the right. (C) The auditory random-tone pitch (RTP) task, in which participants have to decide whether a sound cloud of random tones moves up or down a pitch scale

area reflect the accumulation process prior to the actual choice, which is initiated when the neurons reach a critical firing rate (Gold & Shadlen, 2007).

However, research applying such a model-based approach seems to be dominated by tasks that tap the visual system. Although some studies have investigated perceptual decision processes in other domains as well (e.g., for olfaction, see Uchida & Mainen, 2003; for haptics, see Luna, Hernandez, Brody, & Romo, 2005; Mountcastle, Steinmetz, & Romo, 1990; Pleger et al., 2006; Preuschhof, Schubert, Villringer, & Heekeren, 2010; van Ede, de Lange, & Maris, 2012), research on perceptual choices in the auditory domain is relatively scarce (but see Binder, Liebenthal, Possing, Medler, & Ward, 2004; Cudahy, 1975; De Lucia, Tzovara, Bernasconi, Spierer, & Murray, 2012; Flanagan & Saslow, 1958; Harris, 1952; Klatt, 1973; Russ, Orr, & Cohen, 2008).

In this study, we developed an auditory random-tone pitch (RTP) task to investigate the behavioral signature of auditory perceptual choices. In this task, participants were asked to decide whether a mixture of randomly changing tones moved up or down a pitch scale.

Our main goal was to develop an auditory version of the RDM paradigm that would allow for flexible adaptation of stimulus strength on a continuous scale. As such, the psychophysical features of the auditory stimulus were closely related to those used in the RDM task. In addition, we applied a speed manipulation to illustrate how the task could be used to study changes in the decision process that might or might not be modality specific. The results of the auditory RTP task will be described with respect to performance on the visual RDM task. To this end, we fitted two versions of the DDM to the behavioral data. First, we fitted the proportional-rate diffusion model (Palmer et al., 2005) to the data, to show that the descriptive results (mean RT and accuracy) could be described by a model that has strong theoretical restrictions. Second, we fitted the full DDM (Ratcliff, 1978) to the data and measured the effect of the experimental conditions on the model parameters.

For both tasks, we expected participants to show a smaller boundary separation in the speed than in the accuracy condition (i.e., a speed–accuracy trade-off). Furthermore, we expected higher drift rates for easier stimuli. In addition, we explored the extent to which task modality moderates the speed–accuracy trade-off and the effects of stimulus difficulty.

Method

Procedure

Four of the authors [MK, WB, LM and MM.; mean (SD) age = 29.5 (6.8) years] participated in the experiment. All four participants had already undergone substantial training in both

paradigms, minimizing contamination from practice effects. In the RDM paradigm, the participants were asked to indicate the direction of motion of a cloud of randomly moving white dots on a black background (Fig. 1B). In the RTP paradigm, they were asked to indicate the direction of a "sound cloud" of random tones on a pitch scale (Fig. 1C). All four participants had an additional session of ~1 h of practice on both tasks before participating in the experiment.

In both the RDM and RTP tasks, both difficulty and speed–accuracy trade-off (SAT) were manipulated. Difficulty was manipulated by changing the quality of the perceptual stimulus. For the RDM task, this was done by manipulating the number of coherently moving dots. For the RTP task, this was done by manipulating the number of coherently changing tones on a pitch scale. For both tasks, we used six levels of difficulty, represented by six levels of coherence (0 %, 5 %, 10 %, 20 %, 40 %, and 80 %). For the SAT manipulation, each participant performed three blocks of each task in which they were instructed to be as accurate as possible. To determine a participant-specific time limit, we fitted the proportional-rate diffusion model to the data of the accuracy session and estimated the halfway time threshold from the chronometric curve (predicted by the proportional-rate diffusion model; Palmer et al., 2005). This threshold was then used as a time limit in the following speed session (three blocks). Across all sessions, the participants received one point for each correct choice. No points were given for incorrect choices. In the speed session, the feedback "too slow" was given when a response exceeded the time limit. In the auditory version, the participants additionally received a buzz sound. In each block, the participants performed 600 trials for each modality (100 trials per difficulty level), resulting in 1,800 trials for the accuracy condition and 1,800 trials for the speed condition (the total number of trials was 7,200, which lasted about 5 h). The order of the modalities was interleaved within sessions (i.e., a block of one modality was followed by a block of the other modality) and counterbalanced across participants (i.e., for two participants, the session started with an auditory block, whereas for the other two participants, the session started with a visual block).

Apparatus

Visual and auditory stimuli were generated on a personal computer (Intel Core 2 Quad 2.66-GHz processor, 3 GB RAM, running MS Windows XP SP3) using the Psychophysics Toolbox, version 3.0.8 (Brainard, 1997; Pelli, 1997) for MATLAB (version 2007b; MathWorks, Natick, MA).

Visual stimuli

In the RDM paradigm, participants were asked to decide whether a cloud of white dots on a black background appeared to move to the left or the right. Responses were given by pressing a left (the letter "Z") or right (the letter "M") key on a keyboard. The motion stimuli were similar to those described elsewhere (e.g., Gold & Shadlen, 2000; Mulder et al., 2010; Mulder et al., 2012; Palmer et al., 2005) and were created by changing the locations of dots at each successive video frame (monitor refresh rate = 60 Hz, resulting in 16.7 ms per frame). On the first three frames of the motion stimulus, dots were located in random positions. For each of these frames, the dots were repositioned after two subsequent frames (i.e., the dots in Frame 1 were repositioned in Frame 4, the dots in Frame 2 were repositioned in Frame 5, etc.). For each dot, the new location was either random or at a fixed distance from its former location, in the current direction of the motion stimulus. The probability that a dot would move in line with the motion direction was defined as the *coherence* (see also Britten et al., 1992; Gold & Shadlen, 2003; Palmer et al., 2005).
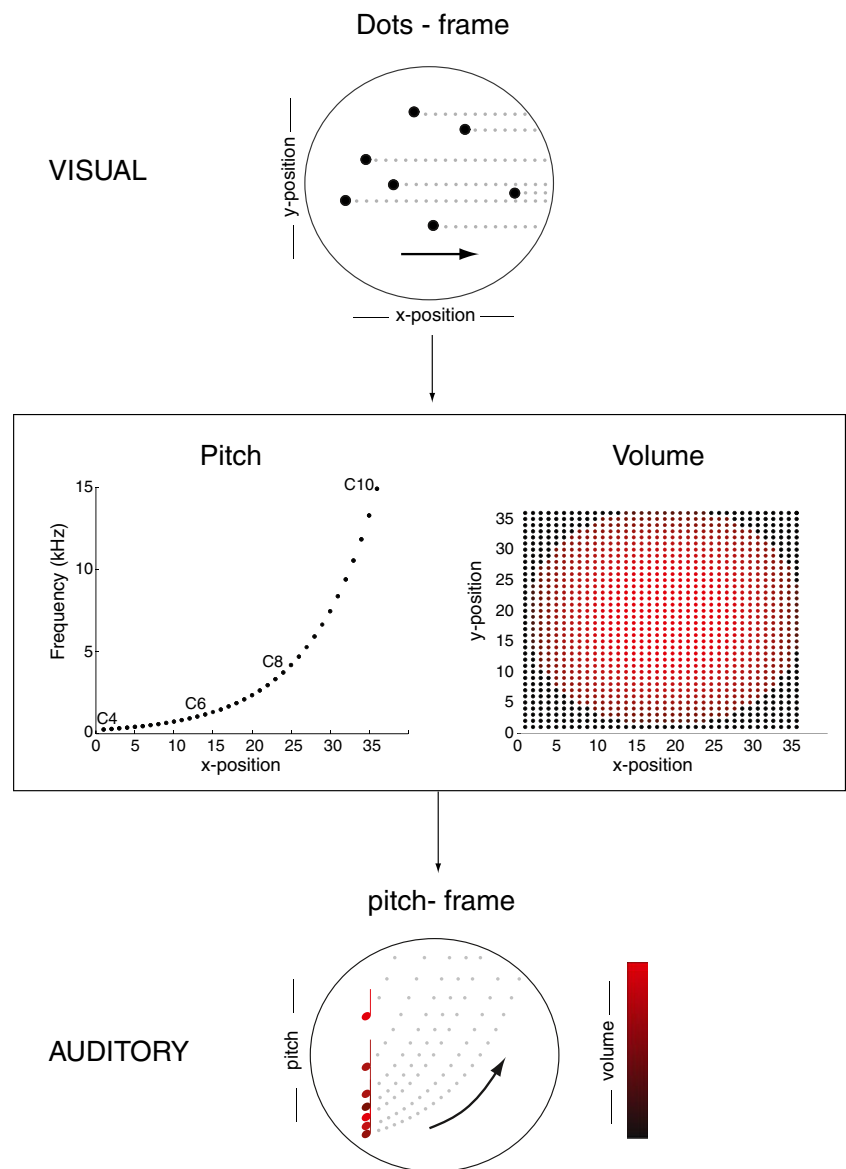
Auditory stimuli

In the RTP paradigm, participants were asked to decide the direction of a "sound cloud" of random tones moving up or down a pitch scale. Responses were given by pressing a left key (the letter "Z") for "up" or a right key (the letter "M") for "down" on a keyboard. To keep the psychophysical features of the stimuli similar to their visual counterparts, we used the locations of the moving dots to calculate the locations on the pitch scale: The x position of each dot was translated to a location on a pitch scale, and both the x and y positions were used to determine the volume on a ramping scale, with 0 being outside the aperture of the motion stimulus and 1 in the middle (see Fig. 2). The rationale behind using both the x and y positions to set the volume was to mimic the circular visual border of the RDM stimulus: The closer a dot was to the aperture, the lower the additive effect it had on the stimulus. At the first, and at each subsequent third frame, the different tones were added together and played for a duration of three frames (~50 ms). The resulting "sound cloud" was a set of bleeps that either increased or decreased in pitch. The stimulus strength again depended on coherence—the probability that a tone moved coherently with the pitch direction (see Fig. 2; for examples, go to http://martijnmulder.wordpress.com/stuff/). For each tone, the pitch frequency could vary between 261.63 and 16744 Hz (which is equal to the audio frequency range between C4 and C10). Each successive step on the x-axis resulted in an auditory change of 200 cents, which is equal to a whole tone on a musical scale (C4, D4, E4, F#4 . . . C10).

Behavioral analyses

First, we fitted the proportional-rate diffusion model (Palmer et al., 2005) to the mean RT and accuracy data for

each task separately. This model has strong theoretically motivated constraints and is able to adequately describe the relationship between coherence and behavioral performance on a continuous scale. The proportional-rate model predicts (1) a close coupling between mean RT, accuracy, and stimulus strength; (2) a linear scaling between stimulus strength and mean drift rate; and (3) equal RTs for correct and incorrect choices (Palmer et al., 2005). These restrictions result in a model that is less complex than the full DDM, while showing an intuitive relationship between stimulus strength and performance. Psychometric and chronometric functions of stimulus strength were fitted to the mean RT and accuracy data using a maximum likelihood procedure (see Palmer et al., 2005, for details).

Additionally, we used the DMAT toolbox to fit the full DDM to the data, in order to obtain parameter values for each condition separately, without the theoretical restrictions

of the proportional-rate diffusion model. The full DDM consists of seven parameters: three for the decision process (i.e., boundary separation a, mean starting point z, and mean drift rate v), a parameter for nondecision processes (i.e., nondecision time T_{er}), and three parameters for across-trial variability (i.e., variability in starting point s_z, variability in nondecision time s_t, and variability in stimulus quality η; Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Tuerlinckx, 2002). The DDM assumes that sensory evidence in favor of one or the other alternative starts to accumulate at a drift rate v from a starting point z, until a boundary is reached. We assume that the effects of difficulty (i.e., the stimulus quality or coherence) are reflected by changes in the drift rate v, whereas the speed–accuracy trade-off is controlled by the height of the boundary separation a. However, the nondecision time T_{er} might be affected by the speed manipulation as well (Mulder et al., 2010;

Osman et al., 2000; Rinkenauer, Osman, Ulrich, Muller-Gethmann, & Mattes, 2004; Voss, Rothermund, & Voss, 2004). As such, when fitting the DDM to the data, we allowed the boundary separation and nondecision time to fluctuate across the speed and accuracy conditions, whereas drift rate was allowed to vary across difficulty conditions. All other parameters were held fixed across conditions.

We used the Diffusion Model Analysis Toolbox (DMAT) to fit the DDM to the individual data (Vandekerckhove & Tuerlinckx, 2007, 2008). For this analysis, the easiest coherence level (80 % coherence) was discarded from the data, because the number of incorrect trials required by DMAT was not reached, due to a ceiling effect (see Fig. 3).
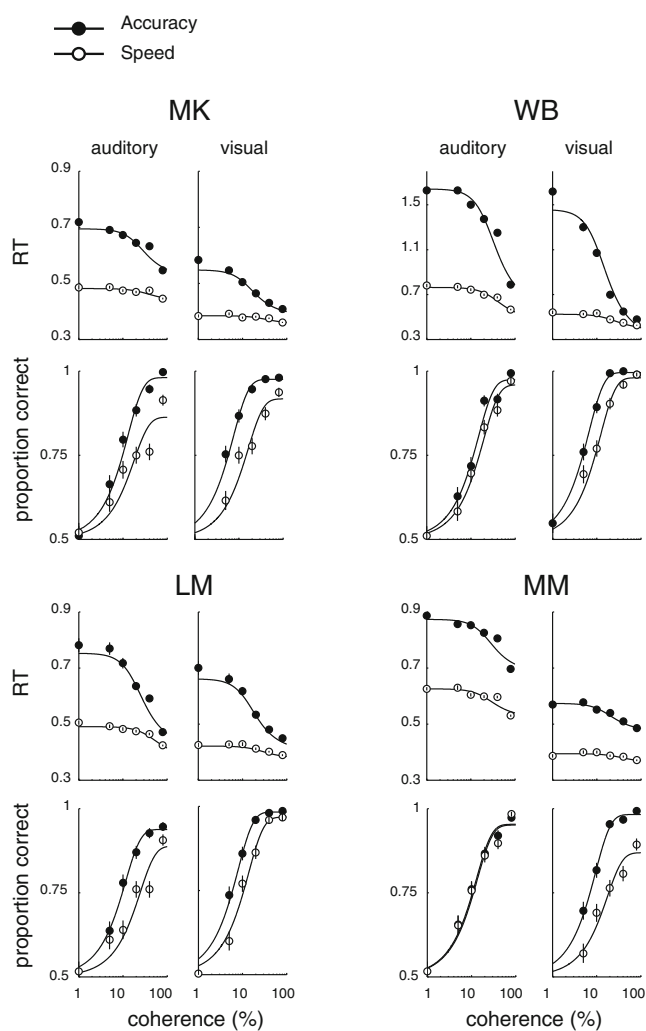

Fig. 3 Descriptive results for each participant, with accuracy (upper graphs) and response time (RT, lower graphs) as functions of stimulus strength (% coherence). Data points are shown for the speed (open) and accuracy (closed) conditions in the auditory (left panels) and visual (right panels) tasks. The continuous lines represent predictions of the proportional-rate diffusion model showing psychometric curves (upper panels) and chronometric curves (lower panels). Error bars represent *SEM*s. For the parameter values, see Tables S1, S2, and S3 in the supplementary materials

Furthermore, for each condition we excluded outlier trials, as defined by an RT that exceeded the conditional average RT plus three times the standard deviation. We used the DMAT Nelder–Mead SIMPLEX optimization algorithm (Nelder & Mead, 1965) to maximize the likelihood of observing a proportion of responses within a given number of RT bins (10th, 30th, 50th, 70th, and 90th percentiles). Quantile probability plots were generated in order to display the quality of the model's fits to the data (see Fig. 4).

Parameter permutations

To obtain confidence intervals for the DDM parameters, we used the nonparametric bootstrap (Efron & Tibshirani, 1993). The DDM was fitted to 1,000 resampled data sets, and the resulting histogram of best-fitting parameter values was used to obtain confidence intervals to test specific differences across conditions.

Results

Below, we will first describe the effects of stimulus difficulty and SAT for both the RTP and RDM tasks. Then we will show the results of the full-DDM analyses, comparing parameter changes across different modalities.

Descriptive results

Fitting the proportional-rate diffusion model to the data clearly showed a coupling between stimulus difficulty and the speed and accuracy of perceptual choices in both the RTP and RDM tasks (see the supplementary materials for the parameter values). More specifically, for each participant, RTs decreased and accuracy increased as a function of coherence (see Fig. 3). Furthermore, for both modalities we observed decreases in accuracy and RT for all difficulty levels when speed was stressed. This effect on accuracy was less apparent for participant MM in the RTP task, in whose data the accuracy levels of the speed-and-accuracy manipulation seemed to overlap (see Fig. 3). Interestingly, when comparing RTs between the visual and auditory tasks, most of the participants showed slower RTs for the auditory stimuli, reflected in the upward vertical shifts of the chronometric curves in Fig. 3. In addition, for participants LM and MM, the upper bounds of the psychometric curves of the auditory task did not reach an accuracy level of 100 %, suggesting perhaps a larger lapse rate for the auditory as opposed to the visual stimuli for these participants. Taken together, both the RTP and RDM tasks showed a close relationship between stimulus strength and performance on a continuous scale. However, the observed differences in the chronometric and psychometric curves reflected
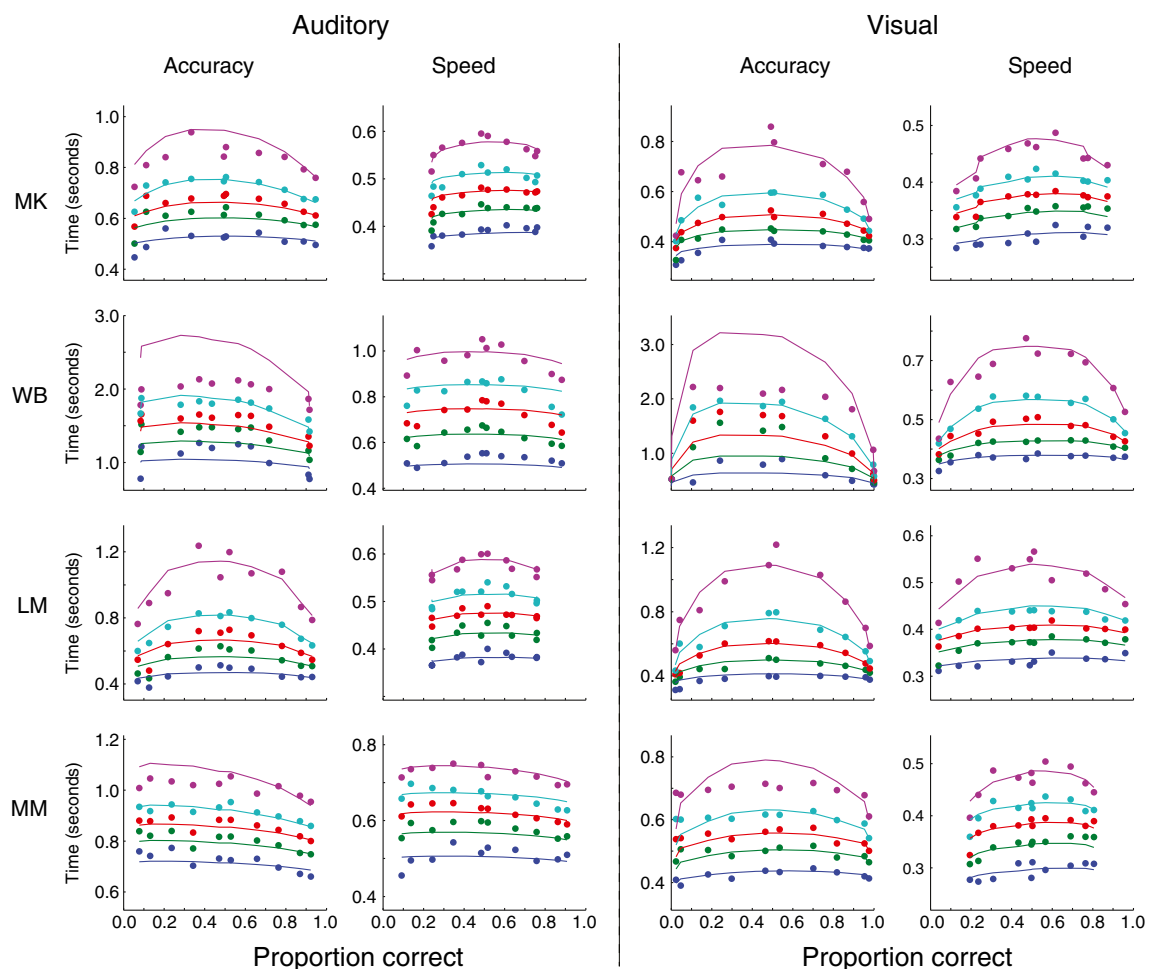
Fig. 4 Individual quantile probability plots for the accuracy and speed sessions within each task modality. Each graph represents the proportion of correct choices and response time (RT) distributions for each difficulty level (data points), as well as the drift diffusion model quantile probability functions describing them (lines). RT distributions are represented by five quantiles (gray shades in print, colors online), plotted along the *y*-axis for each condition. The coherence conditions are split into correct and incorrect responses and divided over the *x*-axis, representing response probabilities. Overall, the quantile probability functions sufficiently describe the data. Fits are worse for the higher quantiles in the accuracy sessions, possibly due to the lower number of incorrect trials for the accuracy sessions. This effect was most apparent for participant WB, for whom the model deviated from the data for most quantiles in the accuracy session

subtle differences in RTs and accuracy between the two tasks. To elaborate these findings, we fitted the full DDM to the data and investigated the differences in the underlying decision parameters between the auditory and visual task domains.

DDM fits

For each experimental modality (visual vs. auditory), the proportions of correct choices and the RT distributions are presented in quantile probability plots for the speed and accuracy conditions separately (Ratcliff & McKoon, 2008; see Fig. 4). These plots show the empirical data together with the quantile probability functions that indicate the fit from the diffusion model. Overall, the quantile probability functions adequately describe the data. However, the functions do deviate from the data at some points. Specifically, the fits are worse for the higher quantiles in the accuracy sessions (see participants WB and MM). This is possibly due to the relatively low number of incorrect trials for these sessions. The misfit was most apparent for participant WB, for whom the model deviated from the data for most quantiles in the accuracy session. A possible explanation for this deviation might be that the RTs were somewhat unusually long for the visual accuracy trials (see Fig. 3).

DDM parameters

Fitting the DDM to the data revealed that, for both tasks, manipulations of the speed–accuracy trade-off affected both boundary separation and nondecision time (see Fig. 5 and

Fig. 5 Results of fitting the full drift diffusion model to the data. For each participant and each modality, the effects of speed–accuracy trade-off on boundary separation and nondecision time are shown, together with the effects of stimulus strength (coherence) on drift rate. Note the close overlaps of the decision bounds across modalities. Error bars represent bootstrap 95 % confidence intervals
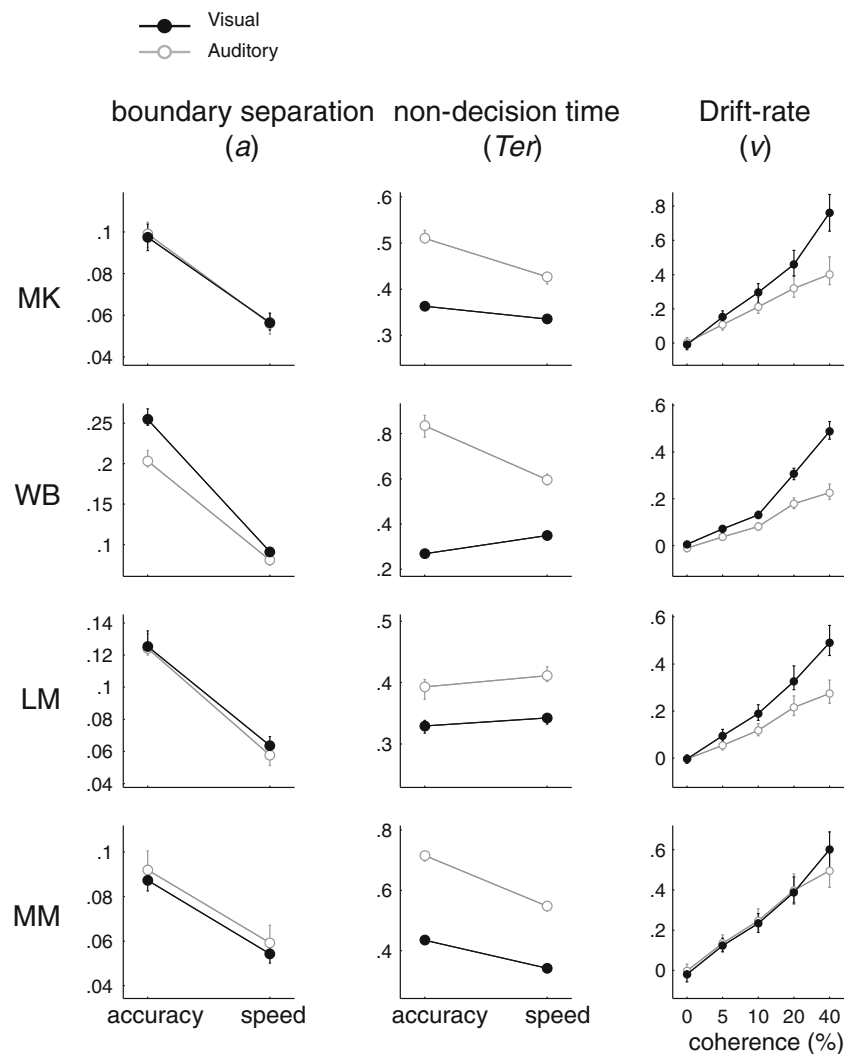
Tables 1 and 2). For three of the four participants (MK, LM, and MM), the change in the speed–accuracy trade-off (typically measured as the decrease in boundary separation from the accuracy to the speed sessions) does not differ between modalities (see Fig. 6 for results and bootstrap estimates of the 95 % confidence intervals). For nondecision time (T_{er}), we found an effect of speed stress as well, showing shorter nondecision times for the speed sessions

Table 1 Drift diffusion model parameters (with 95 % confidence intervals) for the auditory stimuli in each participant

	MK	WB	LM	MM
Boundary Separation (*a*)				
Accuracy	.099 (.093–.105)	.203 (.196–.216)	.124 (.120–.133)	.092 (.089–.101)
Speed	.056 (.051–.061)	.081 (.075–.087)	.058 (.051–.061)	.059 (.057–.067)
Nondecision Time (T_{er})				
Accuracy	.510 (.501–.528)	.835 (.785–.881)	.393 (.373–.405)	.715 (.697–.726)
Speed	.427 (.412–.437)	.596 (.574–.622)	.411 (.402–.426)	.548 (.532–.556)
Drift Rate (*v*)				
0 % coherence	.004 (−.026 to .033)	−.012 (−.027 to .002)	−.003 (−.021 to .016)	−.002 (−.034 to .032)
5 % coherence	.108 (.077–.145)	.037 (.024–.052)	.055 (.035–.078)	.135 (.096–.177)
10 % coherence	.213 (.174–.267)	.081 (.066–.098)	.119 (.095–.147)	.247 (.202–.306)
20 % coherence	.319 (.268–.395)	.179 (.158–.204)	.216 (.181–.264)	.398 (.328–.479)
40 % coherence	.401 (.342–.504)	.226 (.198–.263)	.275 (.234–.332)	.495 (.413–.596)

Table 2 Drift diffusion model parameters (with 95 % confidence intervals) for the visual stimuli in each participant

|	MK	WB	LM	MM
Boundary Separation (a)				
Accuracy	.097 (.091–.104)	.255 (.247–.268)	.125 (.121–.135)	.087 (.083–.094)
Speed	.057 (.053–.061)	.091 (.088–.096)	.064 (.062–.069)	.054 (.050–.060)
Nondecision Time (T_{er})				
Accuracy	.363 (.356–.371)	.268 (.253–.280)	.329 (.318–.339)	.435 (.425–.446)
Speed	.335 (.328–.340)	.349 (.338–.355)	.342 (.332–.346)	.342 (.334–.348)
Drift Rate (v)				
0 % coherence	−.009 (−.038 to .020)	.005 (−.005 to .015)	−.003 (−.022 to .014)	−.019 (−.057 to .012)
5 % coherence	.153 (.118–.188)	.071 (.058–.085)	.095 (.072–.123)	.123 (.092–.160)
10 % coherence	.297 (.235–.348)	.131 (.116–.146)	.190 (.161–.227)	.234 (.189–.282)
20 % coherence	.461 (.393–.542)	.307 (.281–.330)	.326 (.291–.392)	.388 (.336–.464)
40 % coherence	.762 (.655–.868)	.489 (.455–.530)	.491 (.437–.564)	.601 (.506–.688)

(see Fig. 5 and Tables 1 and 2). In addition, we found a difference between the visual and auditory modalities, with longer nondecision times for the auditory task (Fig. 6). Finally, for both tasks we found an effect of difficulty on drift rates, with larger drift rates for trials with higher coherence. However, for three of the four participants (MK, WB, and LM), the drift rates were higher for visual than for auditory stimuli, especially when stimulus strength increased (see Figs. 3 and 5). These results suggest that for these participants, the quality of the sensory evidence was lower for easy auditory than for easy visual stimuli, possibly reflecting an individual difference in modality-specific processing.

Discussion

Research in perceptual decision making is dominated by paradigms that tap the visual system. Here, we investigated whether the behavioral signature of perceptual decisions in the auditory domain is similar to those observed in the visual domain. We developed an auditory version of the RDM task, in which tones correspond to dots and pitch corresponds to motion (the random-tone pitch task, RTP) and manipulated stimulus difficulty and speed–accuracy trade-off. We showed that the relationship between stimulus strength and performance on the RTP task followed a pattern similar to the one observed for the RDM task. By fitting the full drift diffusion model to the data, we showed that the effects of modality mainly affected drift rate and nondecision time, with lower drift rates and longer nondecision times for the auditory task. In contrast, boundary separation seems to be less sensitive to modality-specific effects, as shown by the similar boundary separation values across both tasks for three of the four participants. Similar results were found for the proportional-rate model parameter values (see Table S3 in the supplementary materials). These

Fig. 6 Differences in speed–accuracy trade-offs (SAT), nondecision times, and drift rates between the visual and auditory tasks. Error bars represent bootstrap 95 % confidence intervals
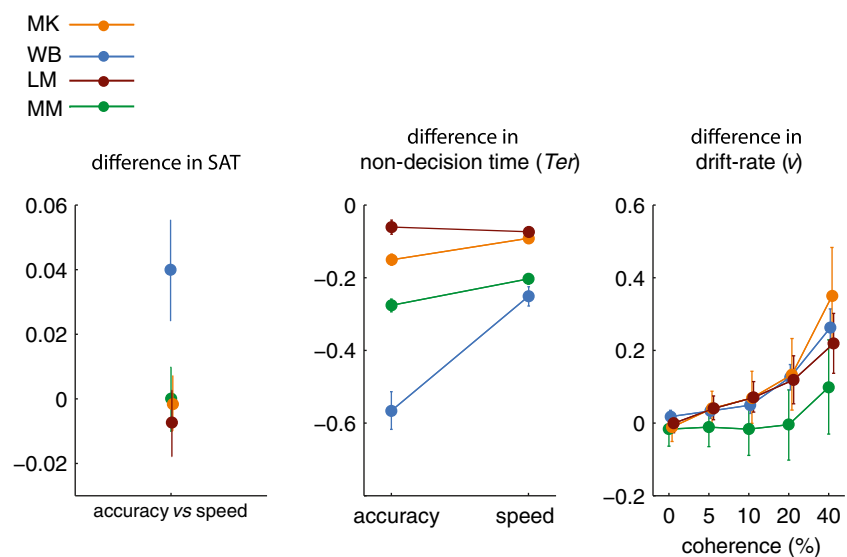
results might indicate that some components of the decision process might be inherent to the participant rather than the task.

Remarkably, the boundary separations for participant WB in the accuracy conditions were much higher than those for the other three participants. One explanation might be that participant W.B. was overly cautious (e.g., Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Forstmann et al., 2008; Ratcliff & McKoon, 2008; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). Indeed, this participant reported that he focused specifically on performing accurately, and therefore was deliberately very cautious. This explanation is in line with the RTs for the accuracy session, which were relatively slow, as compared also to those from other studies using the RDM task (Forstmann et al., 2008; Mulder et al., 2010; Mulder et al., 2012; Palmer et al., 2005; Ratcliff & McKoon, 2008; van Maanen, Grasman, Forstmann, Keuken, et al., 2012a; van Maanen, Grasman, Forstmann, & Wagenmakers, 2012b).

In contrast to boundary separation, drift rate and nondecision time showed considerable variability across participants and modalities. Overall, we found that easy auditory stimuli had lower drift rates than did easy visual stimuli (see Fig. 6). Typically, drift rate reflects the speed of the accumulation process, and therefore indicates the quality of the sensory evidence. Hence, the drift rates suggest that the RTP task is more difficult than the RDM task. This assertion is confirmed by an inspection of the behavioral results. Accuracy levels for the easy trials (20 % and 40 % coherence) were higher for the visual than for the auditory task (note that the 80 %-coherence trials were discarded from the DDM analyses, as participants made very few incorrect choices for this stimulus strength). Furthermore, RTs seem to be longer for the auditory than for the visual modality. Together, these results strongly suggest that the discrimination between auditory stimuli was harder than the discrimination between visual stimuli. One explanation of the differences in auditory performance across participants might be that they could have different experiences within the auditory domain, such as musical development (Foster & Zatorre, 2010; Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Schön, Magne, & Besson, 2004; Spiegel & Watson, 1984; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). Musical experience might enhance the processing of the auditory stimuli used in the RTP task, which in turn could increase drift rate.

In addition to drift rate, effects of nondecision time can account for modality-specific differences in RTs as well. As is shown in Fig. 6, three of the four participants had longer nondecision times for the auditory than for the visual decisions. Typically, nondecision time is associated with sensory-encoding and/or motor processes (Ratcliff & McKoon, 2008;

Vandekerckhove & Tuerlinckx, 2008; Voss et al., 2004; Zylberberg, Ouellette, Sigman, & Roelfsema, 2012). However, it has been shown that the encoding of auditory information is usually faster than the encoding of visual information (Brebner & Welford, 1980; Green & Vongierke, 1984). This suggests that the difference in nondecision times may be due primarily to effects at the motor level. One explanation is that the stimulus–response mapping might be different for the auditory than for the visual task. For the visual task, the direction of the motion stimulus is compatible with the response, resulting in preparatory effects in favor of that particular response (Buetti & Kerzel, 2008; Lien & Proctor, 2002). For the auditory task, however, the "vertical" direction of the tones might result in a longer RT, as the stimulus has to be associated with the particular response first. As such, stimulus-mapping effects might result in a delay (Zhang & Kornblum, 1998), resulting in prolonged nondecision times for auditory stimuli. In addition to the sensory modality effect, some participants showed effects of the speed manipulation on nondecision time, as well. Other studies have also shown effects of speed instructions on nondecision times (Mulder et al., 2010; Osman et al., 2000; Rinkenauer et al., 2004; Voss et al., 2004), suggesting effects of speed instructions on (pre) motor processes that occur after the decision process (Rinkenauer et al., 2004).

Taken together, the differences in drift rates and nondecision times between the two tasks suggest that sensory and motor processes are modality specific, whereas the boundary separation might be less sensitive to the type of information that is accumulated.

This study demonstrated how both the RTP and RDM tasks can be used to study perceptual decision making for different sensory modalities. Furthermore, the tasks illustrate the convenience of using stimuli on a continuous scale. For example, by fitting both the psychometric and chronometric functions (e.g., through the proportional-rate diffusion model), one is able to interpolate participant-specific speed and accuracy levels. This may be useful in experiments in which one wants to keep task difficulty similar across participants (e.g., Mulder et al., 2010; Mulder et al., 2012). Instead of using a fixed level of stimulus strength, one could interpolate the stimulus strength from the psychometric curve for a specific performance level (e.g., 80 % correct responses). Similarly, for a speed–accuracy trade-off paradigm, one may choose to interpolate a participant-specific time limit by using the chronometric curve, to make sure that the underlying effort to speed up is similar across participants. Such experimental controls might be particularly useful for neuroimaging studies in which one is interested in measuring the brain correlates of a specific decision mechanism. Furthermore, combining the RDM with the RTP task might be

especially useful for investigating the neural correlates of perceptual decision making within and across different sensory modalities.

Conclusion

We developed the RTP task, an auditory version of the random-dot motion task to investigate the dynamics of auditory perceptual choices on a continuous stimulus scale. Manipulations of difficulty and the speed–accuracy trade-off illustrated that some parameters might be independent of stimulus modality (i.e., boundary separation), whereas others are not (i.e., drift rate and nondecision time). Future studies on perceptual decision making using the RTP task will allow a more systematic investigation of these modality-specific effects on both the behavioral and neuronal levels.

References

Binder, J., Liebenthal, E., Possing, E., Medler, D., & Ward, B. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience, 7,* 295–301. doi:10.1038/nn1198

Bogacz, R. (2007). Optimal decision-making theories: Linking neurobiology with behaviour. *Trends in Cognitive Sciences, 11,* 118–125. doi:10.1016/j.tics.2006.12.006

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences, 33,* 10–16. doi:10.1016/j.tins.2009.09.002

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. doi:10.1163/156856897X00357

Brebner, J. M. T., & Welford, A. T. (1980). Introduction: An historical background sketch. In A. T. Welford (Ed.), *Reaction times* (pp. 1–23). London, UK: Academic Press.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12,* 4745–4765.

Buetti, S., & Kerzel, D. (2008). Time course of the Simon effect in pointing movements for horizontal, vertical, and acoustic stimuli: Evidence for a common mechanism. *Acta Psychologica, 129,* 420–428. doi:10.1016/j.actpsy.2008.09.007

Cudahy, E. (1975). Fundamental pitch discrimination for monotic and dichotic two-tone complexes. *Journal of the Acoustical Society of America, 57,* S27.

De Lucia, M., Tzovara, A., Bernasconi, F., Spierer, L., & Murray, M. M. (2012). Auditory perceptual decision-making based on semantic categorization of environmental sounds. *NeuroImage, 60,* 1704–1715. doi:10.1016/j.neuroimage.2012.01.131

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap.* New York, NY: Chapman & Hall.

Flanagan, J. L., & Saslow, M. G. (1958). Pitch discrimination for synthetic vowels. *Journal of the Acoustical Society of America, 30,* 435–442.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., et al. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences, 105,* 17538–17542. doi:10.1073/pnas.0805903105

Foster, N. E., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex, 20,* 1350–1359. doi:10.1093/cercor/bhp199

Gold, J. I. (2003). Linking reward expectation to behavior in the basal ganglia. *Trends in Neurosciences, 26,* 12–14.

Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature, 404,* 390–394. doi:10.1038/35006062

Gold, J. I., & Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *Journal of Neuroscience, 23,* 632–651.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Green, D. M., & Vongierke, S. M. (1984). Visual and auditory choice reaction-times. *Acta Psychologica, 55,* 231–247. doi:10.1016/0001-6918(84)90043-X

Hanks, T., Ditterich, J., & Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neuroscience, 9,* 682–689. doi:10.1038/nn1683

Harris, J. D. (1952). Pitch discrimination. *Journal of the Acoustical Society of America, 24,* 750–755.

Heekeren, H. R., Marrett, S., Ruff, D. A., Bandettini, P. A., & Ungerleider, L. G. (2006). Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proceedings of the National Academy of Sciences, 103,* 10023–10028. doi:10.1073/pnas.0603949103

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9,* 467–479. doi:10.1038/nrn2374

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology, 12,* 125–143.

Klatt, D. H. (1973). Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception. *Journal of the Acoustical Society of America, 53,* 8–16.

Lien, M.-C., & Proctor, R. W. (2002). Stimulus–response compatibility and psychological refractory period effects: Implications for response selection. *Psychonomic Bulletin & Review, 9,* 212–238. doi:10.3758/BF03196277

Luna, R., Hernandez, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary somatosensory cortex. *Nature Neuroscience, 8,* 1210–1219. doi:10.1038/nn1513

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research, 219,* 36–47. doi:10.1016/j.heares.2006.05.004

Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research, 20,* 431–435.

Mountcastle, V. B., Steinmetz, M. A., & Romo, R. (1990). Frequency discrimination in the sense of flutter: Psychophysical measurements correlated with postcentral events in behaving monkeys. *Journal of Neuroscience, 10,* 3032–3044.

Mulder, M. J., Bos, D., Weusten, J. M. H., van Belle, J., van Dijk, S. C., Simen, P., . . . Durston, S. (2010). Basic impairments in regulating the speed–accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 68,* 1114–1119. doi:10.1016/j.biopsych.2010.07.031

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience, 32,* 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. doi:10.1093/comjnl/7.4.308

Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8,* 2201–2211.

Osman, A., Lou, L., Muller-Gethmann, H., Rinkenauer, G., Mattes, S., & Ulrich, R. (2000). Mechanisms of speed–accuracy tradeoff: Evidence from covert motor processes. *Biological Psychology, 51,* 173–199.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5*(5), 1, 376–404.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. doi:10.1163/156856897X00366

Pleger, B., Ruff, C. C., Blankenburg, F., Bestmann, S., Wiech, K., Stephan, K. E., . . . Dolan, R. J. (2006). Neural coding of tactile decisions in the human prefrontal cortex. *Journal of Neuroscience, 26,* 12596–12601. doi:10.1523/JNEUROSCI.4275-06.2006

Preuschhof, C., Schubert, T., Villringer, A., & Heekeren, H. R. (2010). Prior information biases stimulus representations during vibrotactile decision making. *Journal of Cognitive Neuroscience, 22,* 875–887. doi:10.1162/jocn.2009.21260

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9,* 438–481. doi:10.3758/BF03196302

Rinkenauer, G., Osman, A., Ulrich, R., Muller-Gethmann, H., & Mattes, S. (2004). On the locus of speed–accuracy trade-off in reaction time: Inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General, 133,* 261–282. doi:10.1037/0096-3445.133.2.261

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22,* 9475–9489.

Russ, B. E., Orr, L. E., & Cohen, Y. E. (2008). Prefrontal neurons predict choices during an auditory same–different task. *Current Biology, 18,* 1483–1488. doi:10.1016/j.cub.2008.08.054

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41,* 341–349. doi:10.1111/1469-8986.00172.x

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *Journal of the Acoustical Society of America, 76,* 1690–1695. doi:10.1121/1.391605

Summerfield, C., & Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: Neural and computational mechanisms. *Frontiers in Neuroscience, 6,* 70. doi:10.3389/fnins.2012.00070

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs. non-musicians: An event-related potential and behavioral study. *Experimental Brain Research, 161,* 1–10. doi:10.1007/s00221-004-2044-5

Uchida, N., & Mainen, Z. F. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nature Neuroscience, 6,* 1224–1229. doi:10.1038/nn1142

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review, 14,* 1011–1026. doi:10.3758/BF03193087

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods, 40,* 61–72. doi:10.3758/BRM.40.1.61

van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *Journal of Neuroscience, 32,* 10408–10412. doi:10.1523/JNEUROSCI.1337-12.2012

van Maanen, L., Grasman, R. P., Forstmann, B. U., Keuken, M. C., Brown, S. D., & Wagenmakers, E.-J. (2012a). Similarity and number of alternatives in the random-dot motion paradigm. *Attention, Perception, & Psychophysics, 74,* 739–753. doi:10.3758/s13414-011-0267-7

van Maanen, L., Grasman, R. P., Forstmann, B. U., & Wagenmakers, E.-J. (2012b). Piéron's law and optimal behavior in perceptual decision-making. *Frontiers in Neuroscience, 5,* 143. doi:10.3389/fnins.2011.00143

van Ravenzwaaij, D., Mulder, M. J., Tuerlinckx, F., & Wagenmakers, E.-J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Frontiers in Psychology, 3,* 132. doi:10.3389/fpsyg.2012.00132

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220. doi:10.3758/BF03196893

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21,* 641–671. doi:10.1080/09541440802205067

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58,* 140–159. doi:10.1016/j.jml.2007.04.006

Watamaniuk, S. N. J., & Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research, 32,* 2341–2347. doi:10.1016/0042-6989(92)90097-3

Zhang, H., & Kornblum, S. (1998). The effects of stimulus–response mapping and irrelevant stimulus–response and stimulus–stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 3–19.

Zylberberg, A., Ouellette, B., Sigman, M., & Roelfsema, P. R. (2012). Decision making during the psychological refractory period. *Current Biology, 22,* 1795–1799. doi:10.1016/j.cub.2012.07.043