

Speed-accuracy trade-off behavior: Response caution adjustment or mixing task strategies?

Leendert van Maanen (lvmaanen@gmail.com)
Department of Psychology, University of Amsterdam
Weesperplein 4, 1018 XA, Amsterdam, The Netherlands

Abstract

The speed-accuracy trade-off (SAT) effect refers to the behavioral trade-off between fast yet error-prone responses and accurate but slow responses. Multiple theories on the cognitive mechanisms behind SAT exist. One theory assumes that SAT is a consequence of strategically adjusting the amount of evidence required for overt behaviors, such as perceptual choices. Another theory hypothesizes that SAT is the consequence of mixing different task strategies. In this paper these theories are disambiguated by assessing whether the fixed-point property of mixture distributions holds, in both simulations and data. I conclude that, at least for perceptual decision making, there is no evidence for mixing different task strategies to trade off accuracy of responding for speed.

Keywords: speed-accuracy trade-off; SAT; fixed-point property; fp; mixture distributions; evidence accumulator models; diffusion model.

Introduction

In sports, acting fast is often as important as acting precise. For example, a basketball player trying to make the winning shot in the dying seconds of the game may be satisfied with less precision in his attempt given the severe time pressure of the clock. On the other hand, if he has just been awarded a free throw without any time pressure, accuracy in his attempt is vital. In experimental psychology, the ability to trade speed of responding for accuracy of responding is referred to as the speed-accuracy trade-off (SAT, Schouten & Bekker, 1967; Wickelgren, 1977). SAT-related effects have been shown in many different experimental paradigms (e.g., Dutilh et al., 2011; Meyer et al., 1988; Wagenmakers et al., 2008).

Response Caution Adjustment

The most prominent theory about the neural and cognitive mechanisms of SAT is *Response Caution Adjustment* (RCA, Bogacz et al., 2010). This view entails that SAT is a consequence of strategically adjusting the amount of evidence required for overt behaviors, such as perceptual choices. According to this view, perceptual choice behavior can be best described as the accumulation of evidence for each choice alternative. That is, given a particular stimulus, the decision maker accumulates over time which alternative is most likely to be the correct response. A response is then provided once a certain minimal level of evidence is exceeded (Figure 1A). Computational models that quantify this process have accounted for many different aspects of decision-making behavior (for reviews see Mulder et al., 2014; Ratcliff & McKoon, 2008), including SAT.

SAT occurs in the accumulator framework through response caution adjustment (Figure 1B). If a decision maker

is pressed for time (or has any other reason why speed-of-responding is important), the minimal level of evidence required for a response may be set to a lower value. If a decision maker is more cautious, then the minimal level of evidence may be set to a higher value. A high value automatically results in longer decision times – and hence longer response times (RT) – since the amount of evidence required to make a decision is larger, and thus takes longer to accrue. However, because of the stochastic nature of the evidence accumulation process, the increased decision time is accompanied by a larger probability of being correct. This is because the probability of accumulating enough evidence for the *incorrect* response alternative is lower as the threshold is set higher.

Mixing Task Strategies

The RCA theory of SAT has been tested in many different studies (e.g., Rae et al., 2014; Mulder et al., 2010, 2013), and in addition is also consistent with many neuroscientific findings (Boehm et al., 2014; Forstmann et al., 2008, 2010; Ho et al., 2012; Van Maanen et al., 2011; Winkel et al., 2012).¹ Nevertheless, alternative theories have been proposed about the nature of SAT. However, no model comparison between different theoretical proposals for SAT has so far been attempted.

One alternative theory of SAT that warrants a formal comparison with RCA is what I refer to here as the *Mixing Task Strategies* (MTS) theory. This theory entails that participants switch between two modes of responding during a task, depending on the speed and accuracy requirements (Ollman, 1966; Meyer et al., 1988). Under accuracy stress, participants respond through a stimulus-controlled process, which is thought to yield optimal – yet relatively slow – performance. Under speed stress, participants are thought to recruit an additional guess process on a large proportion of trials. Because this is hypothesized to be a fast process, the average response times decreases. However, because the guess process leads to chance performance on a certain proportion of trials, accuracy drops as well. This mixture idea lies at the heart of more modern models of SAT, such as the phase-transition model by Dutilh et al. (2011) and a recent ACT-R model of SAT (Schneider & Anderson, 2012).

The essential property of the Mixing Task Strategies theory is that participants use two modes of responding, but in different proportions. In fact, a strong prediction is that any exper-

¹For completeness, it should be mentioned that many of these formal modeling approaches also required the “non-decision time” parameter to vary between speed-stressed and accuracy-stressed conditions.

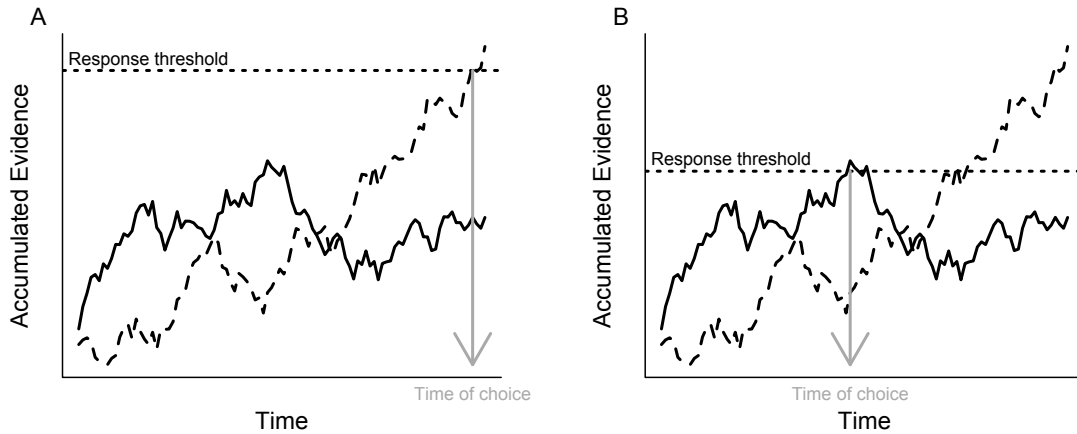


Figure 1: A. An illustration of two evidence accumulation processes, one depicted by a solid line, one by a dashed line. The process that reaches the response threshold the earliest is selected. B. A decreased threshold (panel B vs A) may yield a faster, possibly incorrect, choice.

imental condition that has intermediate speed and accurate stress, should have an in intermediate mixing proportion of the two modes as well. In this paper will test this strong prediction for a simple perceptual choice task (Forstmann et al., 2008) using the fixed-point property of mixture distributions (Falmagne, 1968).

Fixed-Point Property

The fixed-point property (Falmagne, 1968) is a general property of mixture distributions with two base distributions, that can be easily applied to response time data (Van Maanen et al., 2014). Because the probability density of a binary mixture distribution is always the weighted sum of the densities of the two base distributions, it follows that there is (at least) one value that has the same density, independent of the mixture proportions (for a proof, see Falmagne 1968; reiterated in Van Maanen et al. 2014). In terms of mixture distributions of response times, this implies that there will be one RT for which the probability of providing a response at that particular time is equal for all mixtures.

The fixed-point property is illustrated in Figure 2. The figure shows the probability densities of four binary mixture distributions. Each is a mixture of two shifted Wald distribution functions with common scale ($\lambda = 5000$) and shift ($\theta = 100$), but different means ($\mu_1 = 300$ and $\mu_2 = 500$).² The legend in Figure 2 refers to the mixture proportion, here represented as the proportion of the data that comes from the second base distribution (with $\mu_2 = 500$). As is clear from the figure, all densities cross each other at a common RT value, referred to as the *crossing point*. In the Results section below, we will test for the presence of the fixed-point property in empirical

data by assessing whether across participants, the crossing points of pairs of distributions with different mixture proportions are indeed the same.

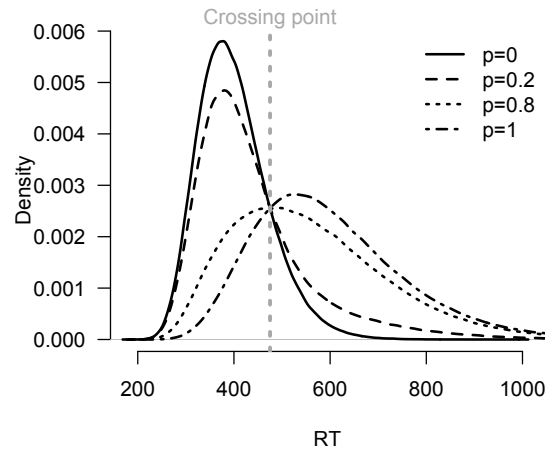


Figure 2: Binary mixture distributions with different mixture proportions always cross at a common RT.

The fixed-point property is predicted by the MTS theory of SAT. That is, if observed RT distributions in SAT are a mixture of the guess process and the stimulus-controlled process, and the mixture proportions differ as a result of the amount of speed stress, then the fixed-point property should be present in the data. On the other hand, the fixed-point property is not predicted by the RCA theory. These predictions will be fleshed out in the next section.

Simulations

To understand which of the theories of SAT predicts the fixed-point property in RT distributions, I generated data under the two theories, for three levels of speed stress. In the RCA

²I chose the shifted Wald distribution function as an example because of its wide applicability in RT data (e.g., Anders et al., 2015; Heathcote, 2004), but the fixed-point property does not depend on the choice of distribution function.

simulation, all trials are drawn from a simple random-walk process with positive drift (cf. Bogacz et al., 2006), and the speed-stress levels are simulated by three different settings of an absorbing boundary. In the MTS simulation, only a proportion of the trials is drawn from that random-walk process, with the remaining trials drawn from a guess process. The three levels of speed-stress are simulated by different mixture proportions.

Response Caution Adjustment Simulation

For the RCA simulation I used a pure drift diffusion model (Bogacz et al., 2006):

$$dx = \mu dt + N(0, \sigma^2 dt) \text{ with } x(0) = a/2. \quad (1)$$

The speed of evidence accumulation is represented by the constant drift μdt , with standard deviation σ . On each trial, a decision is made once the evidence x exceeds one of two boundaries at $x = 0$ and $x = a$. The response time is then determined by the time when one of the boundaries is crossed, plus a fixed non-decision time intercept t_0 . Similar models have been applied to many decision making paradigms to study the cognitive (e.g., Donkin & Van Maanen, 2014; Mulder et al., 2013; Palmer et al., 2005; Ratcliff, 1978; Van Maanen et al., 2012b,a) and neural (e.g., Forstmann et al., 2008, 2010; Ratcliff et al., 2009) mechanisms underlying choice behavior. In particular, this model has been used extensively to study SAT. Overall, SAT has been linked to changes in the boundary parameter a (e.g., Forstmann et al., 2008, 2010; Mulder et al., 2013; Van Maanen et al., 2011; Winkel et al., 2012).

To generate RT distributions for this model, I simulated 10,000 trials in each condition, with the following parameters: $\mu = 0.2$; $\sigma = 0.3$; $t_0 = 200$; $a_1 = 0.3$; $a_2 = 0.6$; $a_3 = 0.72$. Table 1 presents mean RTs for correct responses and accuracy of these simulations, to illustrate that indeed a SAT is simulated.

Table 1: Summary of simulated data.

Model	Mean RT (ms)	Accuracy
RCA		
- $a_1 = 0.3$	458	.67
- $a_2 = 0.6$	1103	.80
- $a_3 = 0.72$	1416	.84
MTS		
- $p_1 = 0.6$	885	.68
- $p_2 = 0.75$	987	.72
- $p_3 = 1.0$	1104	.80

Figure 3A displays kernel density estimates of the RT distributions for correct responses under the RCA theory. The standard deviation of the smoothing kernel is set at 1,000 ms, above the minimal value of 1 standard deviation in the data, as suggested by Van Maanen et al. (2014). It is clear that these

density functions do not all cross at the same RT. Figure 3B shows this even clearer. Here, the differences between each pair of speed-stress levels (i.e., boundary settings) are shown. The RTs where these differences are zero are the crossing points. The absence of the fixed-point property in this simulation is apparent from the multiple crossing points.

Mixing Task Strategies Simulation

The MTS simulation generates data from a stimulus-controlled and a guess process. The stimulus-controlled process is identical to the RCA simulation, except that the boundary setting of the pure drift diffusion is always set at $a = 0.6$. The guess process is simulated by a random draw from a Bernoulli process representing the choice, and an independent draw from a normal distribution with mean $\mu_{guess} = 400$ and $\sigma_{guess} = 100$ representing the response time. Of note is that the mean RT of the guess process is below the mean RT of the stimulus-controlled process, as it represents the faster speed-stressed trials (see the mean RT for the RCA simulation with $a_2 = 0.6$ in Table 1).

Table 1 again presents mean RT for correct responses as well as accuracy for the simulations under the MTS theory. This shows that MTS is indeed consistent with a general SAT effect. Figure 3C shows the kernel density estimates of the RT distributions (with the same smoothing kernel as for the RCA simulations); Figure 3D the density differences. These figures confirm that the MTS theory predicts a fixed-point in the data, as all crossing points in Figure 3D align.

Analysis of Behavioral Data

Simulation of an RCA and an MTS model suggest that a fixed-point in the data is consistent with the MTS theory, but not with the RCA theory. To disentangle these alternative accounts in the domain of perceptual decision making, I re-analyzed data from Forstmann et al. (2008). In this study, participants were asked to perform a random-dot motion task while being stressed for either speed, accuracy, or both on a trial-by-trial basis. This task has been used extensively in the context of SAT (e.g., Palmer et al., 2005; Forstmann et al., 2008; Van Maanen et al., 2011; Mulder et al., 2013) and SAT effects have been explained by the RCA theory. However, a formal comparison with the MTS theory has never been performed. In this particular experiment, the presence of three levels of speed-stress enables a test of the MTS hypothesis that the fixed-point property holds in the data. If the MTS theory is correct, then the proportion of guess responses should be lower for accuracy-stressed trials than for speed-stressed trials. Stressing both speed and accuracy (or rather not stressing anything) should yield a proportion of guess responses that is in between these two extremes.

The Task

In the random-dot motion task, participants had to indicate from a cloud of semi-randomly moving dots what the overall direction of motion is. Prior to each stimulus, participants were presented with one of three cues for 1,000 ms. The cues

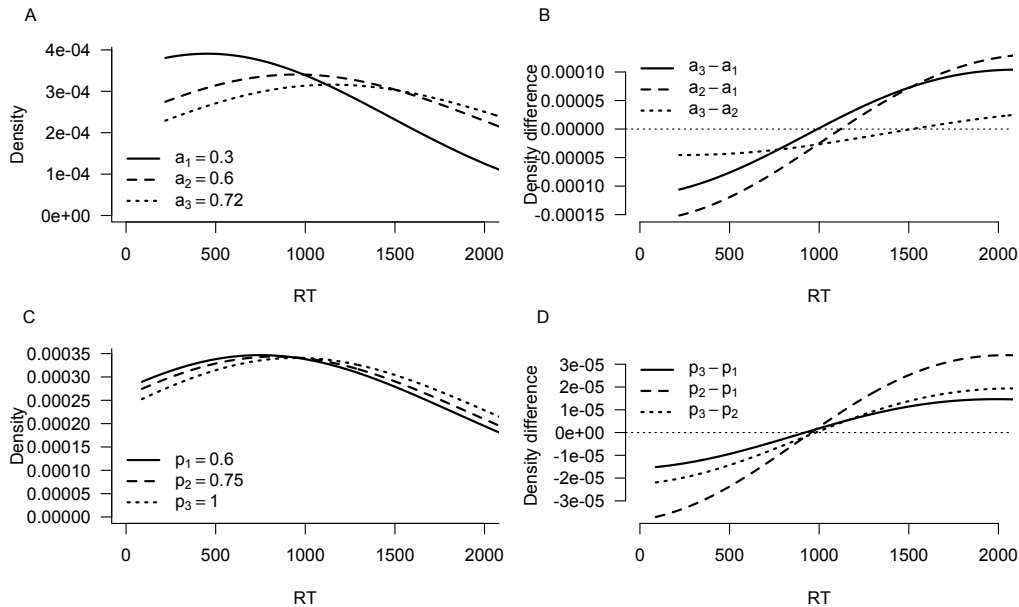


Figure 3: A. Kernel density estimates of three simulated conditions under the Response Cautious Adjustment theory; lines represent different threshold settings (a). B. Density differences of each pair of conditions from A. C. Densities of three simulated conditions under the Mixing Task Strategies theory; lines represent the different proportions p of trials that are form the stimulus-controlled process. D. Density differences of each pair of conditions from C.

could be either “SN” (referring to the German “Schnell”), “NE” (“Neutral”, stressing neither speed nor accuracy), or “AK” (“Akurat”). After a variable interval of 500 ms, the stimulus appeared for another 1,000 ms, followed by 350 ms feedback. Feedback reflected the previously presented cue. Thus, when the cue was either “SN” or “NE”, feedback was given on response speed; when the cue was either “AK” or “NE”, feedback was given on response accuracy. The experiment consisted of 840 trials, equally distributed across the conditions. A total of 20 participants took part in the experiment (see Forstmann et al. 2008 for more details on the experimental procedure).

Results

To assess the presence of the fixed-point property, I only analyzed correct responses (additional simulations showed that the influence of incorrect responses on the crossing points was marginal). The kernel density estimates were computed using a kernel with a standard deviation of 300 ms. Figure 4A and B illustrate that there is no fixed-point in the data. For these figures I aggregated all data points to compute one density function per condition. However, to formally assess the presence of the fixed-point property would be to test within-subjects whether the crossing points are the same (Van Maanen et al., 2014). Because standard frequentist analyses can only test for the presence of a *difference* between conditions, we prefer to apply Bayesian statistics (Rouder et al., 2012). A Bayesian ANOVA (Rouder et al., 2012) quantifies the probability that the observed crossing points are sampled from one underlying population (i.e., when the fixed-point prop-

erty holds) or are sampled from multiple populations (when the fixed-point property does not hold).

Crossing points of the density differences per condition and participant were computed and are presented in Figure 4C. A Bayesian within-subjects ANOVA yields a Bayes factor of 53 in favor of multiple populations of crossing points. This means that the data are 53 times more likely to be generated by such a model than by a model assuming one true population. This result is clearly not in agreement with the fixed-point property, and by extension not in agreement with the MTS theory.

Discussion & Conclusion

The data from Forstmann et al. (2008) is not consistent with an important signature of binary mixture distributions. The absence of the fixed-point property therefore speaks against a MTS theory of SAT. A Bayesian analysis shows that it is in fact 53 times more likely that the data are not from binary mixture distributions. This result is consistent with an RCA theory of SAT. To some extent, this is not surprising, given the excellent fits of cognitive models that implement the RCA theory, both on this data set as well as on related data (e.g., Forstmann et al., 2010; Van Maanen et al., 2011; Mulder et al., 2010, 2013). However, no formal model comparison had so far been attempted. Theoretically, the MTS theory could have generated data that would be excellently fit by RCA models (cf. model mimicry, Ratcliff, 1988; Ratcliff & Smith, 2004). The phase-transition model of Dutilh et al. (2011, an instance of MTS), has been compared to other mod-

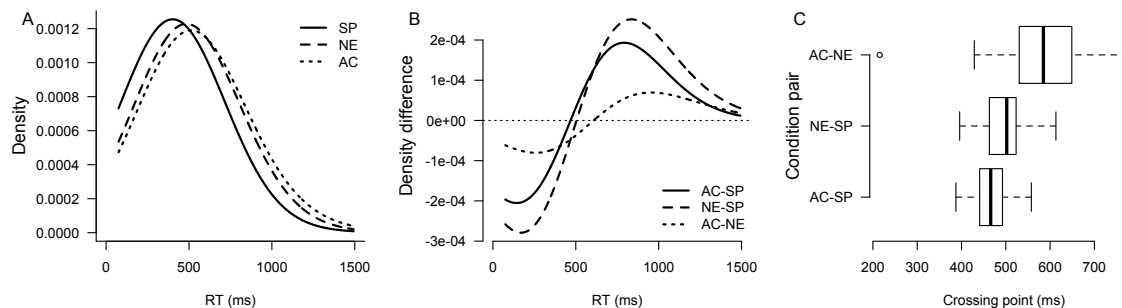


Figure 4: A: Densities of the correct RT distributions in the data. B. Density differences of each conditions pair from A. C. Boxplots indicating the distribution of crossing points per condition pair.

els, but the authors did not include an RCA model in their model comparison. Therefore, although they argue against RCA, it cannot be excluded based on their study.

It is entirely possible that the effects that are collectively referred to as SAT effect depend on different cognitive mechanisms. For example, if presenting a speed-stress cue results to increased preparation (e.g., motor preparation, Rhodes et al. 2004) independently of which mode is actually used on that specific trial, then a fixed-point would also not be observed. This is because the observed response time distributions are not pure mixtures of two base distributions, but rather constitute multiple processes.

Additionally, an experimental paradigm that promotes true guessing behavior may indeed still best be explained by MTS, while an experiment where guessing never leads to satisfactory behavior may be best explained by RCA. Under this view, the best explanation of SAT may be a mixture of RCA and MTS. Nevertheless, the current model and analyses strongly suggests an important role for adjusting control when people are confronted with situations in which the importance of response speed varies.

To disentangle the MTS and RCA theories, I took advantage of the different predictions that these two models make with respect to mixtures of behaviors. The fixed-point property provides an excellent tool to test these predictions.³ Similar predictions may be found in other domains where multiple strategies for a task may (or may not) be expected. Examples include multiple reasoning strategies that may be involved in reasoning tasks (Meijering et al., 2010) or varying proportions of fast-and-automatic processing and slow and deliberate processing, such as can be found in motor sequence learning (Rhodes et al., 2004) or developmental transitions (Van Rijn et al., 2003). For these kinds of response time data, the presence or absence of the fixed-point property seems to be an easy test of multiple competing task processes.

³Van Maanen et al. (2014) includes R code for testing the fixed-point property.

Acknowledgments

Thank you to Birte Forstmann for providing the data and to Hedderik van Rijn voor providing comments on an earlier version of this paper.

References

- Anders, R., Alario, F. X., & Van Maanen, L. (2015). An instrumental cognitive model for speeded and/or simple response tasks. In N. A. Taatgen, M. Van Vugt, J. P. Borst, & K. Mehlhorn (Eds.), *Proceedings of the 13th International Conference on Cognitive Modeling*.
- Boehm, U., Van Maanen, L., Forstmann, B. U., & Van Rijn, H. (2014). Trial-by-trial fluctuations in CNV amplitude reflect anticipatory adjustment of response caution. *Neuroimage*, *96*, 95–105.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700–765.
- 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.
- Donkin, C., & Van Maanen, L. (2014). Piéron's law is not just an artifact of the response mechanism. *Journal of Mathematical Psychology*, *62-63*, 22–32.
- Dutilh, G., Wagenmakers, E.-J., Visser, I., & van der Maas, H. L. J. (2011). A phase transition model for the speed-accuracy trade-off in response time experiments. *Cognitive Science*, *35*, 211–250.
- Falmagne, J. (1968). Note on a simple fixed-point property of binary mixtures. *British Journal of Mathematical and Statistical Psychology*, *21*, 131–132.
- Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S. D., Wagenmakers, E.-J., Bogacz, R., & Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 15916–15920.

- Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 17538–17542.
- Heathcote, A. (2004). Fitting wald and ex-wald distributions to response time data: an example using functions for the s-plus package. *Behavior Research Methods Instruments & Computers*, *36*, 678–694.
- Ho, T. C., Brown, S., van Maanen, L., Forstmann, B. U., Wagenmakers, E.-J., & Serences, J. T. (2012). The optimality of sensory processing during the speed-accuracy tradeoff. *Journal of Neuroscience*, *32*, 7992–8003.
- Meijering, B., Van Maanen, L., Van Rijn, H., & Verbrugge, R. (2010). The facilitative effect of context on second-order social reasoning. In C. R., & S. Ohlsson (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*.
- Meyer, D. E., Irwin, D. E., Osman, A. M., & Kounios, J. (1988). The dynamics of cognition and action: mental processes inferred from speed-accuracy decomposition. *Psychological Review*, *95*, 183–237.
- Mulder, M., Bos, D., Weusten, J., van Belle, J., van Dijk, S., Simen, P., van Engeland, H., & Durston, S. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of ADHD. *Biological Psychiatry*, *68*, 1114–1119.
- Mulder, M., Keuken, M., Van Maanen, L., Boekel, W., Forstmann, B. U., & Wagenmakers, E.-J. (2013). The speed and accuracy of perceptual decisions in a random-tone pitch task. *Attention, Perception & Psychophysics*, *75*, 1048–1058.
- Mulder, M., Van Maanen, L., & Forstmann, B. U. (2014). Perceptual decision neurosciences – a model-based review. *Neuroscience*, (pp. 872–884).
- Ollman, R. (1966). Fast guesses in choice reaction time. *Psychonomic Science*, *6*, 155–156.
- 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*, 376–404.
- Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1226–1243.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R. (1988). A note on mimicking additive reaction time models. *Journal of Mathematical Psychology*, *32*, 192–204.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the eeg. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 6539–6544.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.
- Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. A. (2004). Learning and production of movement sequences: behavioral, neurophysiological, and modeling perspectives. *Human Movement Science*, *23*, 699–746.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374.
- Schneider, D. W., & Anderson, J. R. (2012). Modeling fan effects on the time course of associative recognition. *Cognitive Psychology*, *64*, 127–160.
- Schouten, J. F., & Bekker, J. A. (1967). Reaction time and accuracy. *Acta Psychologica*, *27*, 143–153.
- Van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E. J., Ho, T. C., Serences, J. T., & Forstmann, B. U. (2011). Neural correlates of trial-to-trial fluctuations in response caution. *Journal of Neuroscience*, *31*, 17488–17495.
- Van Maanen, L., De Jong, R., & Van Rijn, H. (2014). How to assess the existence of competing strategies in cognitive tasks: a primer on the fixed-point property. *PLoS One*, *9*, e106113.
- Van Maanen, L., Grasman, R. P. P. P., Forstmann, B. U., Keuken, M., Brown, S. D., & Wagenmakers, E. J. (2012a). Similarity and number of alternatives in the random-dot motion paradigm. *Attention, Perception & Psychophysics*, *74*, 739–753.
- Van Maanen, L., Grasman, R. P. P. P., Forstmann, B. U., & Wagenmakers, E. J. (2012b). Piéron's law and optimal behavior in perceptual decision-making. *Frontiers in Decision Neuroscience*, *5*, article 143.
- Van Rijn, H., Van Someren, M., & Van der Maas, H. L. J. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science*, *27*, 227–257.
- 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.
- Wickelgren, W. (1977). Speed-accuracy tradeoff and information-processing dynamics. *Acta Psychologica*, *41*, 67–85.
- Winkel, J., Van Maanen, L., Ratcliff, R., Van der Schaaf, M., Van Schouwenburg, M., Cools, R., & Forstmann, B. U. (2012). Bromocriptine does not alter speed-accuracy tradeoff. *Frontiers in Decision Neuroscience*, *6*, Article 126.