Pre–registration for: A Confirmatory Test of the Diffusion Model Explanation for the Worst Performance Rule

Gilles Dutilh[1], Joachim Vandekerckhove[2], Alexander Ly[3], Dora Matzke[3], Andreas Pedroni[1], Renato Frey[1], Jörg Rieskamp[1], & Eric-Jan Wagenmakers[3]

[1] University of Basel
[2] University of California Irvine
[3] University of Amsterdam

Correspondence concerning this article should be addressed to:
Gilles Dutilh
University of Basel, Department of Psychology
Missionsstrasse 60/62, 4055 Basel, Switzerland
Ph: (+41) 61–267–3905
E–mail may be sent to gilles.dutilh@gmail.com.

Abstract

People with higher IQ-scores also tend to perform better on elementary cognitive-perceptual tasks, such as deciding quickly whether an arrow points to the left or the right (Jensen, 2006). The worst performance rule (WPR) finesses this relation by stating that the association between IQ and elementary-task performance is most pronounced when this performance is summarized by people's slowest responses. Previous research has shown that the WPR can be accounted for in the Ratcliff diffusion model by assuming that the same ability parameter –drift rate– mediates performance in both elementary tasks and higher-level cognitive tasks. In this preregistration proposal, we aim to test four qualitative predictions concerning the WPR and its diffusion model explanation in terms of drift rate. In the first stage, the diffusion model will be fit to data from 1000 participants completing a perceptual two-choice task; crucially, the fitting will happen after randomly shuffling the key variable, i.e., each participant's score on a working memory capacity test. In the second stage, after all modeling decisions have been made, the key variable is unshuffled and the adequacy of the predictions evaluated by means of confirmatory Bayesian hypothesis tests. By temporarily holding back the mapping of the key predictor we retain flexibility for proper modeling of the data (e.g., outlier exclusion) while preventing biases from unduly influencing the results.

Keywords: Response time model; preregistration; predictions; intelligence.

Over the past decades, the field of mental chronometry has revealed several robust associations between high-level cognitive ability (e.g., IQ, working memory) and response times (RT) in elementary cognitive-perceptual tasks (Jensen, 2006; van Ravenzwaaij, Brown, & Wagenmakers, 2011). The main finding is that people with relatively high IQ-scores tend to respond relatively quickly in simple RT tasks that do not appear to involve deep cognitive processing; one example of such a task is the random dot kinematogram which requires participants to detect the direction of apparent motion in a cloud of dot stimuli.

Another important finding is known as the worst-performance rule (WPR): the fact that the worst performance in these simple tasks —that is, the slowest responses— is most indicative of high–level cognitive ability (Baumeister & Kellas, 1968; Larson & Alderton, 1990a). In this preregistered study, we aim to assess the presence and intensity of the WPR in a large data set. In addition, we test a prediction from the Ratcliff diffusion model (Ratcliff, 1978; Ratcliff, Schmiedek, & McKoon, 2008a), namely that speed of information processing is the factor that underlies the WPR.

The Worst Performance Rule

Since the seminal work by Baumeister and Kellas (1968), the WPR has been shown to exert itself in various forms. In its most general form, the WPR holds that the worst performance on multi–trial elementary cognitive-perceptual tasks is more predictive for *g–loaded* measures than is the best performance on these tasks (Coyle, 2003). This prediction is usually confirmed by demonstrating that higher RT bands correlate more strongly than lower RT bands with both IQ measures (e.g., Larson & Alderton, 1990b; Jensen, 1982; Larson & Alderton, 1990b) and working memory capacity (WMC; e.g., Unsworth, Redick, Lakey, & Young, 2010). For example, Figure 1 presents the results from Larson and Alderton (1990b), showing that the negative correlation between RT and IQ gets stronger as RT lengthens.

The WPR expresses itself in several related ways as well. Coyle (2001), for example, found that the worst performance on a word-recall task (i.e., the lowest number of words from a list recalled by each participant) correlates higher with IQ than the best performance on this task (i.e., the highest number of words from a list recalled by each participant). Furthermore, Kranzler (1992) and Ratcliff, Thapar, and McKoon (2010) showed that the WPR is strongest for multi-trial tasks that are relatively complex.

Several explanations for the WPR have been proposed. The most dominant explanation holds that performance on cognitive tasks of any level in any domain (e.g., IQ, WMC, speeded perceptual choice) is facilitated by the general neural processing speed of an individual's brain (Jensen, 2006). Inspired by this idea, Ratcliff, Schmiedek, and McKoon (2008b) suggested that the drift rate parameter of the diffusion model reflects precisely this speed of processing.

The Ratcliff Diffusion Model

The diffusion model (Ratcliff, 1978) describes the observed RT distributions of correct and error responses on two-choice tasks as the finishing times of a diffusion process with absorbing bounds. When presented with a stimulus, a decision maker is assumed to

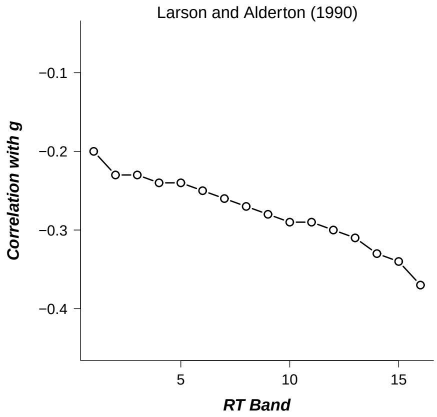

Figure 1. An example of the worst performance rule. The negative correlation of RT with IQ gets stronger as RTs lengthen. Data from Larson & Alderton (1990b).

accumulate noisy evidence from that stimulus (i.e., the meandering lines in Figure 2) until either of two pre-set evidence boundaries is reached and the associated response is initiated. On average, the accumulation of evidence approaches the correct boundary at a speed that is quantified by the drift rate parameter. Due to noise in the accumulated evidence, the diffusion process sometimes reaches the incorrect boundary, leading to error responses. This within-trial noise is also responsible for the right-skewed distribution of RT. In the model's most extended form the diffusion process is governed by seven parameters, including drift rate. Thus, drift rate is a key parameter of the diffusion model, as it corresponds to the signal-to-noise ratio in the evidence accumulation process; hence, drift rate quantifies the speed of information processing.

Ratcliff et al. (2008b) pointed out an important property of the diffusion model for the explanation of the WPR: increasing drift rate acts to reduce RT. Crucially, this reduction is most pronounced for higher percentiles of RT (cf. Van Ravenzwaaij, Brown, & Wagenmakers, 2011), as is illustrated in the upper part of Figure 2. The figure shows RT distributions that originate from two different drift rates. The solid vertical lines indicate the .1 quantiles of the distributions resulting from a high drift rate (dark line) and a low drift rate (grey line). The dashed vertical lines indicate the .9 quantiles of these distributions. Clearly, the change in drift rate leads to a larger shift of the slow .9 quantile than of the fast .1 quantile. Thus, differences in drift rate and differences in IQ have the same qualitative effect on RT, in the sense that both are most strongly expressed in the slowest RTs. This observation adds credibility to the idea that the diffusion model's drift rate parameter quantifies the speed of processing that is thought to underlie the WPR as well as other associations between higher-level and lower-level cognitive tasks. In order to test this idea, several empirical studies related drift rate to IQ and WMC. Ratcliff, Thapar, and McKoon (2011) and Ratcliff et al. (2010) showed that IQ correlated positively with drift rate in recognition memory tasks. Ratcliff et al. (2010) further showed that IQ correlated positively with drift rate in a lexical decision task and a numerosity judgement task. A study by Leite (2009), however, found no evidence of a correlation between IQ and drift

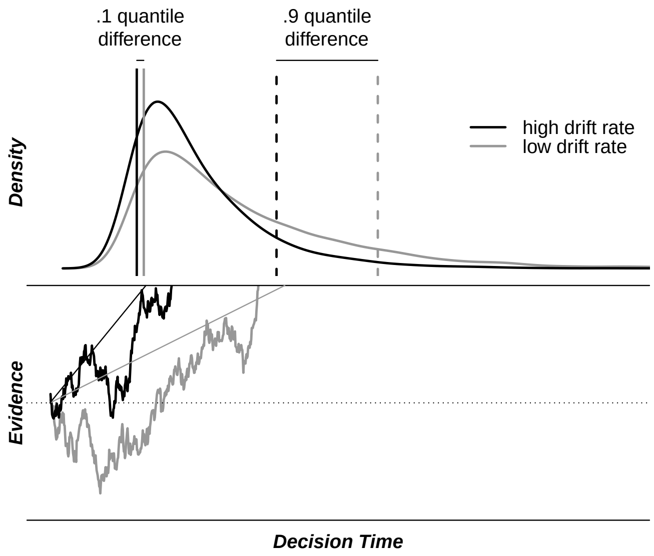

Figure 2. The Ratcliff diffusion model. Noisy evidence is accumulated until one of two pre-set boundaries is reached. The lower half of the figure shows two exemplary accumulation paths (meandering lines) and two different drift rates (the average rate of information accumulation, straight lines). The upper part shows the correct RT distributions that result from a low and a high drift rate. Vertical lines indicate the shift in $.1^{st}$ (solid lines) and $.9^{th}$ (dashed lines) percentiles caused by a change in drift rate.

rate in either a brightness discrimination task or a letter discrimination task. Schmiedek, Oberauer, Wilhelm, Süß, and Wittmann (2007) showed that WMC could be predicted from drift rate on a range of RT tasks.[1]

Van Ravenzwaaij et al. (2011) made another important observation about the relation of drift rate and RT. The diffusion model holds that both stimulus difficulty and subject ability are expressed in drift rate. In fact, drift rate can be viewed as a pair of scales weighting two intrinsically related constructs: difficulty and ability. The drift rate is the deflection of the pointer of this scale and is most pronounced in the slowest RTs, that is, in the worst performance. From this observation, Van Ravenzwaaij et al. (2011) suggested that difficulty, just as ability (e.g., IQ), should be reflected most strongly in the higher ranges of RT, a prediction that was empirically confirmed by Van Ravenzwaaij et al. (2011). From this same interconnection of IQ and difficulty we hypothesize that the WPR is more pronounced for difficult than for easy items of an elementary RT task. Figure 3 illustrates this hypothesis with a concrete example. The figure shows four hypothetical correct RT distributions generated by four drift rates that differ across IQ group and stimulus difficulty. The effect of IQ on slow (.9 quantile) responses is larger than the effect on fast (.1 quantile) responses.

[1] In fact, Schmiedek et al. (2007) constructed a measurement model to distill for each participant a latent factor for drift rate, boundary separation, and non-decision time.

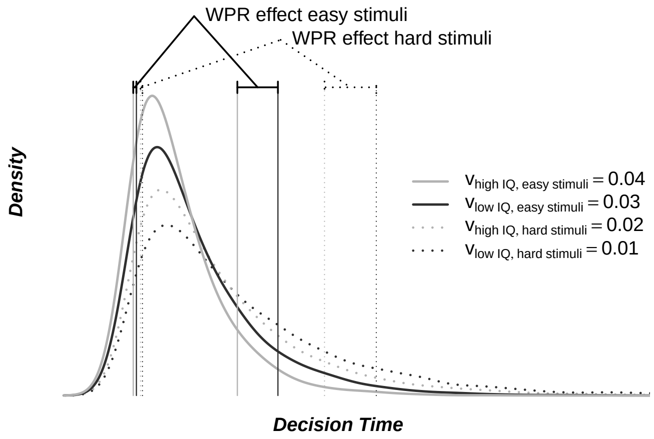

Figure 3. Four hypothetical drift rates v for easy stimuli (solid lines) and difficult stimuli (dotted lines), for participants with a relatively high IQ (light lines) and participants with a relatively low IQ (dark lines). The density lines show the predictions of the diffusion model, given these drift rates. The vertically drawn quantile lines show that the IQ effect on the higher ranges of RT (i.e., the .9 quantile) relative to the lower range of RT (i.e., the .1 quantile) is stronger for the difficult than for the easy stimuli.

This difference is more pronounced for difficult stimuli (dotted lines) than for easy stimuli (solid lines). This prediction is closely in line with the observations of Kranzler (1992) and Ratcliff et al. (2010), who showed that more complex tasks show a more pronounced WPR.

Overview of Hypotheses

The current study proposes a rigorous, preregistered test of four hypotheses related to the WPR and the account provided by the Ratcliff diffusion model. First, we test the existence of the WPR. Second, we test the prediction that the WPR is larger for difficult than for easy trials in a simple RT task. Third, we test the prediction that the diffusion model drift rate parameter correlates with WMC. Fourth, we test the prediction that the correlation between drift rate and WMC is higher for difficult trials than for easy trials from the perceptual RT task. We test these hypotheses by analyzing an existing data set with 1000 participants for which we measured both perceptual choice RT and WMC. A detailed account of the design, hypothesis, and proposed analyses is provided below.

Data Collection and Method

The data at hand have been collected in a large-scale study on the genetic underpinnings of risk preferences, funded by the Swiss National Science Foundation. For this study, 1000 participants (500 participants in Berlin, Germany; 500 in Basel, Switzerland) were tested on a range of psychological tasks. Among the participants, 65% were students, and 62% were female. The age range spans 18-36 years with a mode at 24 years. For the current study, we will analyze the data of two relevant tasks: a WMC test and a perceptual two-choice RT task.

Working Memory Capacity Battery

To measure working memory we used the WMC battery developed by Lewandowsky, Oberauer, Yang, and Ecker (2010). This battery was constructed as a tool to measure working memory capacity with a heterogeneous set of tasks that involves both verbal and spatial working memory. A pre-defined measurement model described in Lewandowsky et al. (2010) allows the calculation of a single WMC score for each participant. Lewandowsky et al. (2010) show that this score has a strong internal consistency and correlates highly with Raven's test of fluid intelligence ($r = .67$).

Speeded Perceptual Two-Choice Task

In the elementary RT task, participants were presented with 10×10 matrices of black and white dots (Figure 4). Participants were instructed to indicate whether the matrix contained more black or more white dots by pressing either of two mouse buttons. In this simple perceptual task, difficulty can be manipulated by adjusting the number of black and white dots. Participants saw 90 easy trials (proportion of black and white dots: 60/40, 40/60) and 90 difficult trials (proportions 55/45, 45/55). In addition, there were trials with an equal proportion of black and white dots. These stimuli are "undoable", and are of no special interest in this perceptual task but were included for comparison with another task conducted in the large scale study. In the current analyses, we nonetheless include these trials in order to facilitate the estimation of the diffusion model parameters. Participants received no feedback, but were instructed to respond as fast and accurately as possible. A "too slow" message was displayed after responses slower than 3.5 seconds. Our task originates from Dutilh and Rieskamp (in press) and resembles tasks that have been modeled successfully with the diffusion model, such as the brightness discrimination task (Ratcliff & Rouder, 1998) and the numerosity task (Ratcliff et al., 2010).

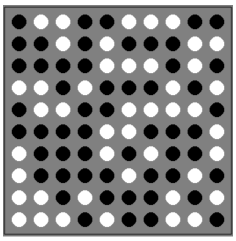

Figure 4. Example of a stimulus in the perceptual RT task. Participants pressed the left or right mouse button to indicate quickly whether the stimulus contained more black or more white dots.

Registered Analysis Plan

In this study, we aim to test four key hypotheses in a manner that is described in detail below. For all hypotheses, we use the Bayes factor to quantify the degree of confirmation provided by the data (Jeffreys, 1961); we will also provide the posterior distribution for the parameters of interest. The registered analysis plan will be carried out on the complete data set (subject to the outcome-blind decisions by the modeler; see the next section on the two-stage analysis process). In a second, exploratory analysis, we will test the hypotheses separately for the relatively homogeneous student group and the relatively heterogeneous non-student group.

Note that, with 1000 participants, we collected data that are sufficiently informative to pass Berkson's "interocular traumatic test" (Edwards, Lindman, & Savage, 1963) such that the confirmatory hypothesis tests serve merely to confirm what is immediate apparent from a cursory visual inspection of the data.

Analysis of Hypothesis 1: Worst Performance Rule

For each participant we obtain a single WMC score from the WMC battery. Furthermore, for each participant we obtain the 1/6, 2/6, 3/6, 4/6, and 5/6 quantiles of correct RTs; it is possible to use more quantiles, but only at the cost of reducing the precision with which the mean RT within each bin is estimated. Hypothesis 1 states that the correlation between WMC and mean RT within each quantile is negative (i.e., higher WMC is associated with faster responding). More specifically, Hypothesis 1 states that the absolute magnitude of this correlation increases monotonically from the fastest to the slowest quantile (i.e., the WPR). Hypothesis 1a refers to the WPR for easy stimuli, and Hypothesis 1b refers to the WPR for difficult stimuli.

Both Hypothesis 1a and 1b will be tested separately, in the following manner. Denote by ρ_i the estimated Pearson correlation coefficient for quantile i. Then, the simplest linear version of the WPR predicts that $\rho_i = \beta_0 + \beta_1 I_i$, where I_i indicates the quantile, β_0 is the intercept of the regression equation, and β_1 is the slope. We then use the Bayes factor (Jeffreys, 1961; Kass & Raftery, 1995) to quantify the support that the data provide for two competing hypotheses: the null hypothesis $\mathcal{H}_0 : \beta_1 = 0$ versus the WPR alternative hypothesis $\mathcal{H}_1 : \beta_1 < 0$. Under $\mathcal{H}_1$, we assign each ρ_i an independent uniform prior from -1 to 0, in order to respect the fact that all correlations are predicted to be negative. Furthermore, we assign a uniform prior to β_0 that ranges from -1 to 0, in order to respect the fact that even for the fastest RTs, the correlation is not expected to be positive. Finally, we assign a uniform prior to β_1 that ranges from its steepest possible value to 0. Specifically, since the quantiles are on the scale from zero to one, and the highest possible value of the intercept β_0 equals 0, the assumption of linearity across the scale implies that the steepest slope is -1. Hence, we assign β_1 a uniform prior from -1 to 0.

With the model specification in place, the Bayes factor between $\mathcal{H}_0 : \beta_1 = 0$ versus $\mathcal{H}_1 : \beta_1 \sim U[-1, 0]$ can be obtained using an identity known as the Savage-Dickey density ratio (e.g., Dickey & Lientz, 1970; Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010). Specifically, this involves focusing on parameter β_1 in $\mathcal{H}_1$ and comparing the prior ordinate at $\beta_1 = 0$ to the posterior ordinate at $\beta_1 = 0$, that is, by computing $\mathrm{BF}_{10} = p(\beta_1 = 0 \mid \mathcal{H}_1)/p(\beta_1 = 0 \mid y, \mathcal{H}_1)$, where y denotes the observed data. Bayes factors higher than 1 favor $\mathcal{H}_1$ and provide support for the WPR. All parameters will be estimated simultaneously using a hierarchical Bayesian framework and Markov chain Monte Carlo (MCMC, e.g., Lee & Wagenmakers, 2013).

Analysis of Hypothesis 2: Stronger Worst Performance Rule for More Difficult Stimuli

The WPR tested under Hypothesis 1 is predicted to be more pronounced for difficult stimuli than for easy stimuli. In the previous WPR model, $\rho_i = \beta_0 + \beta_1 I_i$; now denote β_1 for the difficult stimuli by β_{1d} and β_1 for the easy stimuli by β_{1e}. Hypothesis 2 holds that $\beta_{1e} > \beta_{1d}$. We multiply both parameters by -1 so that we obtain variables on the probability scale, and hence $\beta^*_{1d} > \beta^*_{1e}$. We use a dependent prior structure (Howard, 1998), apply a probit transformation, and orthogonalize the parameter space (Kass & Vaidyanathan, 1992). Specifically, denoting the probit transformation by Φ^{-1}, we write $\Phi^{-1}(\beta^*_{1d}) = \mu + \delta/2$ and $\Phi^{-1}(\beta^*_{1e}) = \mu - \delta/2$. We assign the probitized grand mean parameter μ an uninformative distribution, that is, $\mu \sim N(0, 1)$, and then use the Bayes factor to contrast two models: the null hypothesis $\mathcal{H}_0 : \delta = 0$ versus the alternative hypothesis $\mathcal{H}_2 : \delta > 0$. We complete the model specification for $\mathcal{H}_2$ by assigning the difference parameter δ a default folded normal prior defined only for positive values, that is, $\delta \sim N(0, 1)^+$. As before, parameter estimates are obtained from MCMC sampling in a hierarchical Bayesian model and Bayes factors will be computed using the Savage-Dickey density ratio test on parameter δ under $\mathcal{H}_2$.

Analysis of Hypothesis 3: Working Memory Capacity Correlates Positively with Drift Rate

We fit the diffusion model to the data using hierarchical Bayesian estimation (e.g., Wabersich & Vandekerckhove, 2014; Wiecki, Sofer, & Frank, 2013). This hierarchical method allows us to exploit the vast number of participants and estimate parameters even for participants whose data contain little information (for example due to a small number of errors, which are crucial for diffusion model parameter estimation). Hypothesis 3 holds that WMC correlates positively with drift rate. Hypothesis 3a refers to the positive correlation between WMC and drift rate for the easy stimuli, and Hypothesis 3b refers to the positive correlation between WMC and drift rate for the difficult stimuli. Both Hypothesis 3a and 3b will be tested separately, in the following manner.

First WMC is included within the hierarchical structure. WMC will then be correlated with drift rate estimates (Hypothesis 3a: for the easy stimuli; Hypothesis 3b: for the difficult stimuli) in a hierarchical structure. The null hypothesis holds that there is no correlation, $\mathcal{H}_0 : \rho = 0$, whereas the alternative hypothesis holds that the correlation is positive, $\mathcal{H}_3 : \rho > 0$. Specifically, we assign ρ a uniform prior from 0 to 1. Bayes factors can be obtained by a Savage-Dickey density ratio test on parameter ρ under $\mathcal{H}_3$.

Analysis of Hypothesis 4: Stronger Correlation Between Working Memory and Drift Rate for More Difficult Stimuli

Hypothesis 4 holds that WMC correlates more strongly with drift rates for difficult stimuli than with drift rates for easy stimuli. Denote by ρ_d the WMC-drift rate correlation for the difficult stimuli, and by ρ_e the WMC-drift rate correlation for the easy stimuli. Hypothesis 4 states that $\rho_d > \rho_e$. Moreover, both ρ_d and ρ_e are assumed to be positive, so that both are on the probability scale. Consequently, the proposed analysis mimics that of Hypothesis 2: We use a dependent prior structure, apply a probit transformation, and orthogonalize the parameter space. We write $\Phi^{-1}(\rho_d) = \mu + \delta/2$ and $\Phi^{-1}(\rho_e) = \mu - \delta/2$. We assign the probitized grand mean parameter μ an uninformative distribution, that is, $\mu \sim N(0, 1)$, and then use the Bayes factor to contrast two models: the null hypothesis $\mathcal{H}_0 : \delta = 0$ versus the alternative hypothesis $\mathcal{H}_4 : \delta > 0$. We complete the model specification for $\mathcal{H}_4$ by assigning the difference parameter δ a default folded normal prior defined only for positive values, that is, $\delta \sim N(0, 1)^+$. As before, parameter estimates are obtained from MCMC sampling in a hierarchical Bayesian model and Bayes factors will be computed using the Savage-Dickey density ratio test on parameter δ under $\mathcal{H}_4$.

Two-Stage Analysis

We pursue an unbiased method to test the diffusion model account of the WPR. Therefore, we propose a two-stage analysis with a special status for coauthor JV who fits the diffusion model to data (e.g., Vandekerckhove & Tuerlinckx, 2007; Vandekerckhove, Tuerlinckx, & Lee, 2011; Vandekerckhove & Tuerlinckx, 2008; Wabersich & Vandekerckhove, 2014). In the first stage we provide JV with the perceptual RT data and a *randomly permuted* version of the WMC variable. With these data in hand, JV produces code to fit the model while respecting the analysis choices outlined above (i.e., Hypothesis 1-4). This first stage allows JV to model the data at will, by excluding outliers, introducing contaminant processes, adding transformations, and generally make any other reasonable modeling choice. Since the crucial WMC score variable is randomly permuted, the correlation between drift rate and WMC estimated in this stage-one model is meaningless. The first stage is terminated when JV indicates the model code is ready. At this point the code is fixed and made available on the Open Science Framework (`https://www.osf.io`). In the second stage the true sequence of WMC scores is revealed, and the code created by JV is applied to the data in a deterministic manner to address each of the hypotheses outlined above.

This two-stage analysis is both flexible and fair. It is flexible because the modeler retains the freedom to exclude data and make adjustments to the model to account for eventual peculiarities of the data. And it is fair because the modeling choices are not outcome-driven, that is, guided by expectations about the main hypotheses.

References

Baumeister, A. A., & Kellas, G. (1968). Reaction time and mental retardation. *International review of research in mental retardation*, *3*, 163–193.

Coyle, T. R. (2001). IQ is related to the worst performance rule in a memory task involving children. *Intelligence*, *29*, 117–129.

Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence*, *31*, 567–587.

Dickey, J. M., & Lientz, B. P. (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *The Annals of Mathematical Statistics*, *41*, 214–226.

Dutilh, G., & Rieskamp, J. (in press). Comparing perceptual and preferential decision making. *Psychonomic Bulletin and Review*.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193–242.

Howard, J. V. (1998). The 2×2 table: A discussion from a Bayesian viewpoint. *Statistical Science*, *13*, 351–367.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Jensen, A. R. (1982). Reaction time and psychometric g. In *A model for intelligence* (pp. 93–132). Springer.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences.* Elsevier.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Kass, R. E., & Vaidyanathan, S. K. (1992). Approximate Bayes factors and orthogonal parameters, with application to testing equality of two binomial proportions. *Journal of the Royal Statistical Society, Series B*, *54*, 129–144.

Kranzler, J. H. (1992). A test of Larson and Alderton's (1990) worst performance rule of reaction time variability. *Personality and Individual Differences*, *13*, 255–261.

Larson, G. E., & Alderton, D. L. (1990a). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, *14*, 309–325.

Larson, G. E., & Alderton, D. L. (1990b). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, *14*, 309–325.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* Cambridge University Press.

Leite, F. P. (2009). Should IQ, perceptual speed, or both be used to explain response time? *The American journal of psychology*, *122*, 517–526.

Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. H. (2010). A working memory test battery for matlab. *Behavior Research Methods*, *42*, 571–585.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.

Ratcliff, R., Schmiedek, F., & McKoon, G. (2008a). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence*, *36*, 10-17.

Ratcliff, R., Schmiedek, F., & McKoon, G. (2008b). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence*, *36*, 10–17.

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, *140*, 464.

Ratcliff, R., Thapar, J., & McKoon, G. (2010). Individual differences, aging, and IQ in two–choice tasks. *Cognitive Psychology*, *60*, 127–157.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414.

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence*, *38*, 111–122.

van Ravenzwaaij, D., Brown, S., & Wagenmakers, E.-J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition*, *119*, 381–393.

Van Ravenzwaaij, D., Brown, S., & Wagenmakers, E.-J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition*, *119*, 381–393.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*, 1011-1026.

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*, 61-72.

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two–choice response times. *Psychological Methods*, *16*, 44–62.

Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior Research Methods*, *46*, 15–28.

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, *60*, 158–189.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift–diffusion model in Python. *Frontiers in Neuroinformatics*, *7*, doi:10.3389/fninf.2013.00014.