**In press for *Decision***

# $\hat{w} = .2, \hat{a} = .8, \hat{c} = .6$: So What? On the Meaning of Parameter Estimates from Reinforcement-Learning Models

Helen Steingroever[a], Ruud Wetzels[b], and Eric-Jan Wagenmakers[a]

[a] University of Amsterdam, The Netherlands

[b] PricewaterhouseCoopers, Amsterdam, The Netherlands

Abstract

In their comment on Steingroever, Wetzels, and Wagenmakers (SWW; 2014), Konstantinidis, Speekenbrink, Stout, Ahn, and Shanks (KSSAS; 2014) convincingly argue why a wide range of sophisticated model comparison methods is required to select a good model for the Iowa gambling task (IGT). While we agree with KSSAS on this count, the focus of SWW was not on model comparison. Here we clarify our initial goal, that is, to illustrate why assessment of absolute model performance is necessary to avoid premature conclusions about the psychological processes that drive performance on the IGT. In addition, we elaborate on the advantages and drawbacks of both the post hoc absolute fit method and the simulation method. Finally, we highlight the distinction between statistical aspects of model adequacy and psychological relevance of parameter estimates.

*Keywords:* Posterior predictives; experience-based decision making; Iowa gambling task; mathematical modeling

In their comment on Steingroever et al. (2014), Konstantinidis et al. (KSSAS; 2014) convincingly argue that the examination of reinforcement-learning (RL) models for the Iowa gambling task (IGT) is only at its starting point. In particular, KSSAS argue that, in model comparison, a uni-dimensional approach may be ill-advised. According to KSSAS, a more balanced and comprehensive assessment of model adequacy requires one to integrate results obtained from different model comparison methods. KSSAS also state that model comparison methods should take into account model complexity (see for example the generalization criterion; Busemeyer & Wang, 2000). In addition, KSSAS argue that more model comparison efforts are required to test the psychological meaningfulness of parameter estimates obtained from fitting RL models to IGT data. Altogether, KSSAS provide a focused summary of the current state of research on RL models for the IGT, while making crucial suggestions for future research.

However, KSSAS also express skepticism about the simulation method. According to KSSAS, it is premature to claim that the simulation method offers a more useful method of model discrimination than the post hoc fit method. In particular, KSSAS argue that "the use of parameters estimated with the post hoc fit method could result in a considerable underestimation of a model's simulation performance" (p. 187). While we agree with this claim in general terms,[1] it is important to note that the parameters obtained from the post hoc method are the ones that applied researchers base their conclusions on, and the validity of these conclusions was the focus of Steingroever et al. (2014).

In the remainder of this comment, we clarify the goal of Steingroever et al. (2014); we elaborate on the advantages and drawbacks of both the post hoc absolute fit method

[1]The parameter space partitioning study from Steingroever, Wetzels, & Wagenmakers, 2013a, suggests that, for example, in the case of the PVL model and the stylized data set showing a pronounced preference for the good decks, other parameter combinations may produce better simulation performance.

Correspondence concerning this article should be addressed to: Helen Steingroever, Ph: +31 20525 8869, E-mail: helen.steingroever@gmail.com.

and the simulation method; and we highlight the distinction between statistical aspects of model adequacy and psychological relevance of parameter estimates.

## Clarification of Our Initial Goal

RL models are applied to IGT data whenever researchers are interested in making inferences on the psychological processes that drive IGT performance (see Steingroever et al., 2013a, for references). Unfortunately, systematic and detailed evaluations of model performance are virtually absent from the applied literature. Such evaluations, however, are necessary to decide whether or not the conclusions from the estimated model parameters are trustworthy. Consequently, estimated model parameters reported in applied studies might not be indicative of the psychological processes they seek to represent. In this context, Steingroever et al. (2014) pointed out that the conclusions from the model parameters are valid only when the applied model provides an adequate account of the observed data. This means that before drawing conclusions from the model parameters, researchers have to assess absolute model performance. It is therefore insufficient to report only relative assessment measures (i.e., the fit of a RL model compared to a baseline model or another RL model). In order to assess absolute model performance, Steingroever et al. (2014) focused on two methods: the post hoc absolute fit method and the simulation method. These methods allow one to assess whether the applied model provides an adequate account for the observed data in absolute terms. Only when this requirement has been fulfilled can the estimated parameters be mapped to the psychological processes of interest to any degree of confidence.[2]

Steingroever et al. (2014) applied the post hoc absolute fit method and the simulation method to seven data sets and three different models. This comprehensive analysis revealed that in many data sets the neglect of absolute model performance can result in mislead-

[2]Good performance on other measures, such as parameter consistency, test of generalizability, and test of specific influence (e.g., Ahn, Busemeyer, Wagenmakers, & Stout, 2008; Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2010; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008; see also section "Statistical Adequacy versus Psychological Relevance") are also important to establish the psychological relevance of parameter estimates.

ing conclusions about psychological processes. Thus, Steingroever et al. (2014)'s goal was to emphasize the importance of assessing absolute model performance before interpreting model parameters, and their goal was neither to identify a good comparison method for RL models nor to identify a good RL model for the IGT.

A compelling illustration of why it is important to assess absolute model performance prior to the interpretation of model parameters is given by Anscombe's quartet (Anscombe, 1973; Andraszewicz et al., 2015), presented in Figure 1. Each panel shows a different data set consisting of eleven (x, y) pairs of observations. The four data sets are constructed in such a way that the two variables $x$ and $y$ have the same means and variances, and that the linear regression coefficient and Bayes factor (Jeffreys, 1961) are identical. The Bayes factor of 23 indicates a strong support for the presence of a linear association in all four data sets. However, the visual presentation of the data sets suggests that the statistical model assuming that there is a linear relationship between the two variables is only valid in the upper left panel. Hence, only in the upper left panel can the model parameter $\rho$ be interpreted in a meaningful way.

### Current Way of Assessing Model Performance

As pointed out in the previous section, systematic and detailed evaluations of IGT model performance are virtually absent from the applied literature. The studies that did assess model performance used various model selection criteria, such as Bayesian information criterion (BIC) and $G^2$ (Steingroever et al., 2014).[3] However, the disadvantage of these criteria is that they are relative measures, and thus provide no information on whether a given model is able to provide an adequate account of the data.

To illustrate the risks and drawbacks of the current way of assessing model performance, consider the study of Yechiam, Busemeyer, Stout, and Bechara (2005). These

[3]This index compares the performance of two models (i.e., the accuracy of one-step-ahead predictions when provided with intermediate feedback on the observed choices and payoffs); the first model is an RL model that aims to explain trial-to-trial dependencies and learning effects; the second model is a baseline model that assumes constant choice probabilities across all trials (equal to the individual's overall choice proportions from each deck; Busemeyer & Stout, 2002).

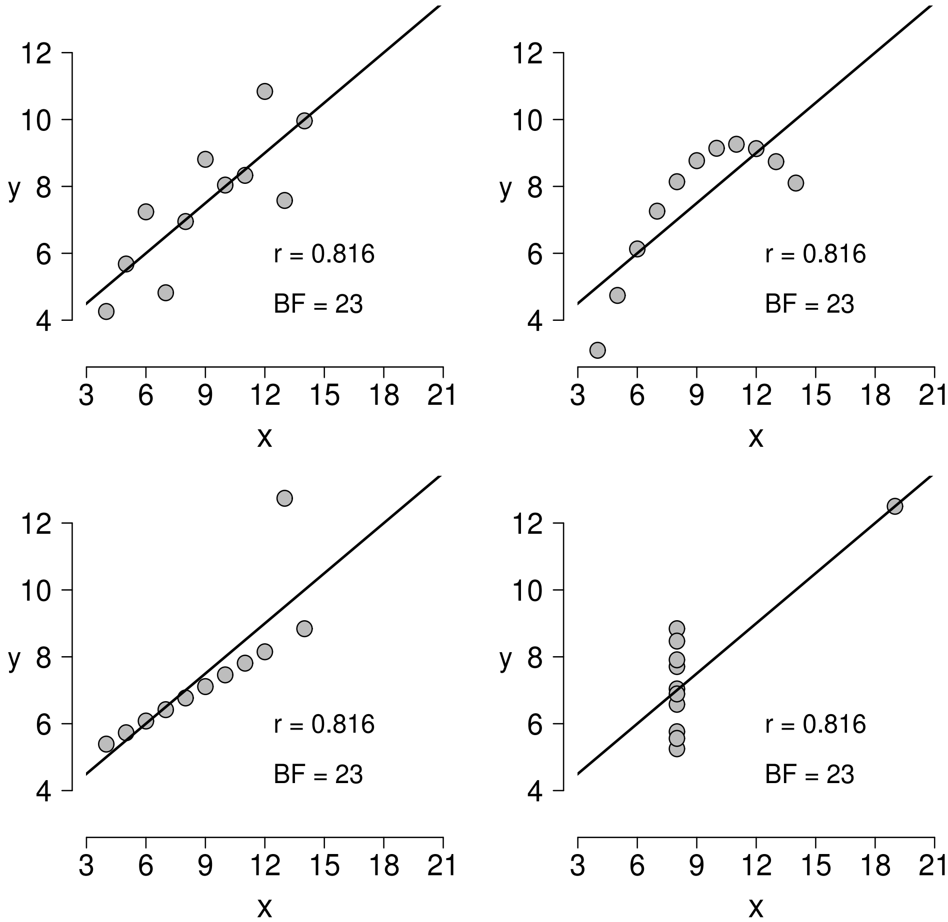


*Figure 1*. Anscombe's Quartet. Four different data sets that are identical on summary statistics (i.e., means and variances of the two variables $x$ and $y$, the linear regression coefficient, and the Bayes factor, Jeffreys, 1961). However, the visual presentation of the data sets suggests that they differ strongly, and that the statistical model assuming that there is a linear relationship between the two variables is only valid for the data from the upper left panel. Hence, only in the upper left panel can the model parameter $\rho$ be interpreted in a meaningful way.

authors fit the EV model to data of 10 groups of people suffering from various neuropsychological disorders, and mapped these groups according to the differences between their model parameters and those of their control group. The purpose of this analysis was to characterize the decision-making deficits of each clinical group in terms of underlying psychological processes, and to examine whether differences in neuropsychological disorders can be explained by differences in specific psychological processes that underlie performance on the IGT. Specifically, Yechiam et al. (2005) argued that their results suggest the presence of three different clusters. First, there is a cluster of young polydrug abusers and young alcohol abusers with parameter values similar to their respective control groups. Second, there is a cluster of chronic cocaine abusers, chronic cannabis abusers, Huntington's patients, seniors, and patients with bilateral damage to the ventromedial prefrontal cortices, with parameter values indicating higher attention to gains and greater recency effects than their respective control groups. Finally, there is a cluster of Parkinson's patients, Asperger's patients, and patients with lesions in the right somatosensory and insular cortex, with parameter values indicating higher attention to losses than their respective control groups. Yechiam et al. (2005) justified the validity of the estimated model parameters by arguing that the EV model, from which the estimates were obtained, outperforms the baseline model for the majority of the participants (see Table 1 in Yechiam et al., 2005). We agree that it is important to establish the superiority of the EV model compared to the baseline model. However, the fact that the EV model outperforms the baseline model does not guarantee that the fit is adequate; it is possible that the EV model has a better fit than the baseline model, but nonetheless provides a poor fit to the data in absolute terms. And if a model provides a poor fit to the data, this suggests that the parameter estimates are non-interpretable and that they do not link to the psychological processes they seek to represent (cf. Anscombe's Quartet shown in Figure 1). In sum the current procedure to validate the link between model parameters and psychological processes is based solely on assessment of relative model adequacy. Although we agree that relative model adequacy is of great importance, so is the assessment of absolute model adequacy.

**Additional Ways to Assess Model Performance**

In Steingroever et al. (2014) we illustrated why conclusions from model parameters may be misleading when these parameters are interpreted without having confirmed the adequacy of the model through sensible checks. We proposed two such checks –the post hoc fit and simulation method– as straightforward, general, and easy-to-implement methods to assess absolute model performance. These two methods both assess absolute model performance by considering the data generated by the models relative to the data that were observed. We showed the relevance and feasibility of these methods by applying them to three different RL models and seven data sets. We argued how these methods shed light on which steps are necessary to avoid premature conclusions about the model parameters.

Two important advantages of these methods are that they assess absolute model performance (i.e., not relative to a baseline model or another RL model), and that they do not require data from a second experiment (in contrast to the generalization criterion method proposed by KSSAS). We consider good post hoc fit and simulation performance as minimum requirements that need to be fulfilled before one can claim that conclusions about the model parameters are meaningful. This does not mean that we believe relative model adequacy to be unimportant; both relative and absolute measures of model adequacy are needed to obtain a complete assessment of the extent to which parameter estimates are valid indicators of the associated psychological processes.

To illustrate the importance of assessing absolute model performance using the post hoc fit method and the simulation method, consider the data from Premkumar et al. (2008; left panel of Figure 2), and imagine that the authors had conducted a BIC comparison of the EV, PVL and PVL-Delta models, and that the PVL model was the winning model. Current practice dictates that the estimated parameters of the PVL model are then used to draw conclusions about the psychological processes underlying IGT performance. However, the following illustration shows that these conclusions might be invalid, and that two more steps need to be taken following the BIC assessment. In a second step, post hoc absolute fit performance needs to be assessed. This performance is shown in the middle panel of

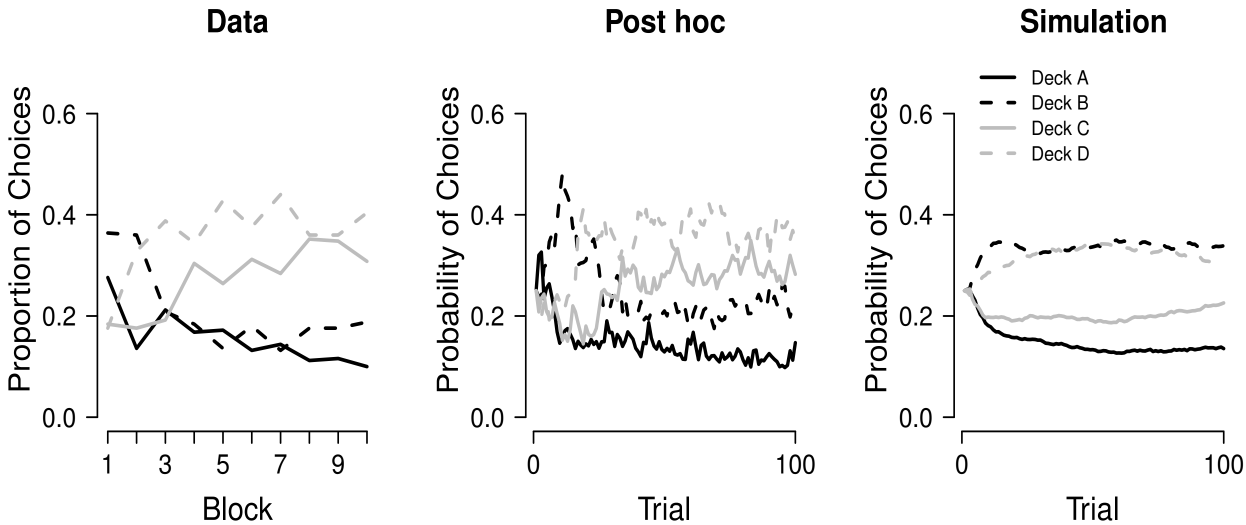


*Figure 2*. Good post hoc absolute fit performance, but poor simulation performance of the PVL model for the data set of Premkumar et al. (2008). The first panel presents the observed mean proportions of choices from each deck across 10 blocks. Each block contains 10 trials. The second and third panel present the mean probabilities of choosing each deck on each trial as obtained with the post hoc absolute fit performance and the simulation performance, respectively.

Figure 2. A comparison of the left and middle panels suggests that the PVL model makes accurate one step ahead predictions for that data set. Thus, the post hoc fit method can be used to determine whether, in addition to the better BIC score of the PVL model relative to the EV and PVL-Delta model (i.e., first step), the PVL model has sufficient post hoc fit performance also in absolute terms.

However, these two steps are still not sufficient to conclude that the parameters are meaningful. In a third step, simulation performance needs to be assessed. This performance is shown in the right panel of Figure 2. A comparison of the left and right panels suggests that, when we use the estimated parameters to generate data, the model predicts a qualitatively different choice pattern, that is, a preference for the decks with infrequent losses (i.e., bad deck B and good deck D), even though the participants in the observed data set show an overall preference for the good decks. Thus, even though the PVL model has sufficient relative and absolute post hoc fit performance, the estimated parameters may not be indicative of the psychological processes that drove the participants' performance on the IGT—a conclusion based on the poor simulation performance of the PVL model.

## Post Hoc Absolute Fit Method versus Simulation Method

In their comment, KSSAS endorse the post hoc fit method. However, instead of providing arguments supporting the post hoc fit method, they mainly provide arguments against the simulation method. First, KSSAS state that the simulation method is a crude generalization test. We agree with KSSAS on this count; in Steingroever et al. (2014), we introduced the simulation method as the least demanding test of generalizability (compared to different implementations of the generalization criterion; see Ahn et al., 2008; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008).

Second, KSSAS state that the simulation method ignores model complexity. We also agree with KSSAS on this count. However, we never claimed otherwise – our focus was not on model selection, and we proposed neither the simulation method nor the post hoc fit method as a tool for model comparison (see above).

Third, KSSAS state that the simulation method ignores the fact that different models are required to fit different individuals. We also agree with KSSAS on this count. This is why we used a Bayesian hierarchical implementation of the RL models that naturally accounts for similarities and differences between individuals (Shiffrin, Lee, Kim, & Wagenmakers, 2008). While we agree with KSSAS that a mixture model might be even more sensible, it should be noted that this is a valid criticism of all current model comparison approaches in the field of RL models because none of them use a mixture model implementation.

Fourth, KSSAS state that the simulation method, in contrast to the post hoc fit method, can favor models that are unlikely to generalize to new datasets. KSSAS illustrate this claim with a deterministic model that contains only one parameter, say $\theta$. This parameter identifies a sequence of choices within the set of all possible choice sequences. However, we believe this example is flawed; the following illustration shows that, in contrast to KSSAS's claim, the post hoc fit method would also favor this model. Here we consider only a two-trial IGT, but our conclusions generalize to any trial length. In a two-trial IGT, there are $4^2$ possible choice sequences that are indexed by the parameter $\theta$ (Table 1).

Table 1
*Illustration of the deterministic model proposed by Konstantinidis et al. All possible choice sequences are presented as a function of the model parameter $\theta$. Here we only consider a two-trial IGT, but our conclusion generalize to any trial length.*

| $\theta$ | Choice sequence | |
|---|---|---|
| 1 | A | A |
| 2 | A | B |
| 3 | A | C |
| 4 | A | D |
| 5 | B | A |
| . | . | . |
| . | . | . |
| 16 | D | D |

The resulting payoffs are neglected because they do not influence the predictions of this model. Suppose that the observed choice sequence was (A, D). Consequently, the estimate for $\theta$ is $\hat{\theta} = 4$. If we now apply the post hoc fit method using $\hat{\theta}$, we obtain $P[S_A(t = 1)] = .25$ for the predicted choice probability on the first trial because all decks are equally likely on the first trial. For the predicted choice probability on the second trial, we obtain $P[S_D(t = 2)] = 1$. Note that, in the case of this deterministic model, this prediction is independent of the information on the choice and payoff on the first trial. If we use a longer IGT, the predicted choice probabilities are always one for the chosen deck, and zero for the unchosen decks (except for the first trial; see appendix in Steingroever et al., 2014, for a recipe of the post hoc fit method). Thus, the post hoc fit method would clearly favor this overly complex model.

KSSAS provided only one argument supporting the post hoc fit method, that is, that RL models should include choice inertia. However, choice inertia (also called choice perseveration) and choice mimicry refer to two different concepts. Choice inertia is a psychological process describing participants' tendency to repeat previous choices (see for example Worthy, Pang, & Byrne, 2013). Choice mimicry, on the other hand, refers to a model's ability to fine-tune its parameters to obtain an accurate fit to the exact sequences of observed payoffs and choices. This fine-tuning is facilitated by the fact that the post-hoc fit method uses the payoffs and choices twice; once to obtain the best-fitting parameters, and once to make one-step-ahead "predictions" given the payoffs and choices from all previous trials. Thus, it is possible to observe choice mimicry in the post hoc absolute fit performance of a model, even though the model does not incorporate choice inertia as one of the psychological processes underlying IGT performance.

In addition, one crucial drawback of model selection criteria, such as BIC and $G^2$ –criteria that are based on the post hoc fit method– is that these criteria do not fully account for model complexity. A more complete account for model complexity (i.e., number of parameters, functional form of the model, and extension of the parameter space; Myung & Pitt, 1997) is offered by the Bayes factor (Jeffreys, 1961; see Steingroever, Wetzels, & Wagenmakers, submitted, for a comparison of RL models using Bayes factors).

Finally, it is important to note that, in a strict sense, the post hoc fit method does not predict, but it post-dicts.[4] This is because in a first step, we use the data (i.e., choices, wins, losses of a given participant) to estimate the model parameters. In a second step, we use the estimated parameters together with the choices, wins, losses observed up-to and including the current trial to "predict" the choice on the next trial. But this is a postdiction because we already used all of the data to estimate the model parameters, and thus use all information for the one-step-ahead "predictions".

## Statistical Adequacy versus Psychological Relevance

In cognitive modeling, it is important to differentiate between statistical aspects of model adequacy (e.g., good fit to the observed data and good predictions for new data), and psychological relevance of parameter estimates (e.g., good parameter recovery, and good performance on tests of selective influence). In other words, in a statistical sense a model can be adequate because it provides a good fit to the observed data or because it makes

[4]See Wagenmakers, Grünwald, and Steyvers (2006) for a description of the accumulative prediction error—a method that, in contrast to the post hoc fit method, indeed assesses whether a model can predict the next choice given only the information from the previous trials.

good predictions for the behavior in new environments. However, this statistical dimension of model adequacy is not necessarily related to the psychological dimension that considers the link between model parameters and psychological processes. For example, a model is psychologically inadequate if its parameters cannot be recovered, or if it performs poorly on a test of selective influence—a test that investigates whether experimental manipulations of certain psychological processes are reflected by the parameters that are sought to measure the corresponding psychological processes (e.g., Steingroever et al., 2013b; Wetzels et al., 2010). It is therefore possible that a model has adequate statistical properties, but that the parameter recovery is relatively poor, compromising the model's utility in applied settings (e.g., the Value-Plus-Perseveration model; Ahn et al., 2014; Worthy et al., 2013). On the other hand, a model with parameters that clearly link to distinct psychological processes does not necessarily need to be able to account for the finer details of performance in order to be practically useful. For example, the EZ diffusion model (Wagenmakers, van der Maas, & Grasman, 2007) cannot account for the fact that error responses are differently distributed than correct responses; nevertheless, the EZ diffusion model is useful for estimating model parameters, especially when sample size is low (van Ravenzwaaij & Oberauer, 2009).

## Conclusion

To conclude, KSSAS provide a focused summary of the current state of model comparison efforts in the field of RL models for the IGT, and they make crucial suggestions for future efforts on finding a good model for the IGT. Even though we agree with KSSAS on these counts, the focus of Steingroever et al. (2014) was not on model comparison. But we focused on the assessment of absolute model performance. Our goal was to illustrate why applied researchers should carefully assess absolute model performance before they draw conclusions from the estimated parameters. In particular, premature conclusions can be avoided by carefully assessing both post hoc fit performance and simulation performance.

## References

Ahn, W.-Y., Busemeyer, J. R., Wagenmakers, E.-J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science*, *32*, 1376 - 1402.

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience Psychology and Economics*, *4*, 95 - 110.

Ahn, W.-Y., Vasilev, G., Lee, S. H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., & Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Frontiers in Psychology*, *5*.

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R. P. P. P., Verhagen, J., & Wagenmakers, E.-J. (2015). An introduction to Bayesian hypothesis testing for management research. *Journal of Management*, *41*, 521 - 543.

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*, 17 - 21.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, *14*, 253 - 262.

Busemeyer, J. R., & Wang, Y.-M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, *44*, 171 - 189.

Jeffreys, H. (1961). *Theory of probability* (Third ed.). Oxford University Press, Oxford, England.

Konstantinidis, E., Speekenbrink, M., Stout, J. C., Ahn, W.-Y., & Shanks, D. R. (2014). To simulate or not? Comment on Steingroever, Wetzels, and Wagenmakers (2014). *Decision*, *1*, 184 - 191.

Myung, I. J., & Pitt, M. A. (1997). Applying occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, *4*, 79 - 95.

Premkumar, P., Fannon, D., Kuipers, E., Simmons, A., Frangou, S., & Kumari, V. (2008). Emotional decision-making and its dissociable components in schizophrenia and schizoaffective disorder: A behavioural and MRI investigation. *Neuropsychologia*, *46*, 2002 - 2012.

Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, *32*, 1248 - 1284.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2013a). A comparison of reinforcement-learning models for the Iowa gambling task using parameter space partitioning. *The Journal of Problem Solving*, *5*, Article 2.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2013b). Validating the PVL-Delta model for the Iowa gambling task. *Frontiers in Psychology*, *4*.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2014). Absolute performance of reinforcement-learning models for the Iowa gambling task. *Decision*, *1*, 161 - 183.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (submitted). Bayes factors for reinforcement-learning models of the Iowa gambling task.

van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: EZ, fast-dm, and dmat. *Journal of Mathematical Psychology*, *53*, 463 - 473.

Wagenmakers, E.-J., Grünwald, P., & Steyvers, M. (2006). Accumulative prediction error and the selection of time series models. *Journal of Mathematical Psychology*, *50*, 149 - 166.

Wagenmakers, E.-J., van der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3 - 22.

Wetzels, R., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2010). Bayesian parameter estimation in the Expectancy Valence model of the Iowa gambling task. *Journal of Mathematical Psychology*, *54*, 14 - 27.

Worthy, D. A., Pang, B., & Byrne, K. A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. *Frontiers in Psychology*, *4*.

Yechiam, E., & Busemeyer, J. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, *12*, 387 - 402.

Yechiam, E., & Busemeyer, J. R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior*, *63*, 370 - 394.

Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, *16*, 973 - 978.

Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. *Journal of Mathematical Psychology*, *51*, 75 - 84.