Hierarchical Bayesian Modeling: Does it Improve Parameter Stability?

Benjamin Scheibehenne (benjamin.scheibehenne@unibas.ch)
Economic Psychology, Department of Psychology, Missionsstrasse 62a
4055 Basel, Switzerland

Thorsten Pachur (pachur@mpib-berlin.mpg.de)
Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94
14195 Berlin, Germany

Abstract

Fitting multi-parameter models to the behavior of individual participants is a popular approach in cognitive science to measuring individual differences. This approach assumes that the model parameters capture psychologically meaningful and stable characteristics of a person. If so, the estimated parameters should show, to some extent, stability across time. Recently, it has been proposed that hierarchical procedures might provide more reliable parameter estimates than non-hierarchical procedures. Here, we examine the benefits of hierarchical parameter estimation for assessing parameter stability using Bayesian techniques. Using the transfer-of-attention-exchange model (TAX; Birnbaum & Chavez, 1997), a highly successful account of risky decision making, we compare parameter stability based on hierarchically versus non-hierarchically estimated parameters. Surprisingly, we find that parameter stability for TAX is not improved by using a hierarchical Bayesian as compared to a non-hierarchical Bayesian approach. Further analyses suggest that this is because the shrinkage induced by hierarchical estimation overcorrects for extreme yet reliable parameter values. We suggest that the benefits of hierarchical techniques may be limited to particular conditions, such as sparse data on the individual level or very homogenous samples.

Keywords: cognitive modeling; parameter consistency; risky choice; hierarchical Bayesian modeling; transfer-of-attention-exchange model

Introduction

In cognitive science, a highly popular approach to describing and understanding behavior is to develop models with adjustable parameters that can be fitted to data. As parameters of cognitive models are usually supposed to represent meaningful aspects of cognitive processing, they are often used to study, measure, and describe individual differences between people. For illustration, consider cumulative prospect theory (CPT; Tversky & Kahneman, 1992), one of the most prominent models of decision making under risk. According to CPT, responses to a risky alternative (which lead to different outcomes with particular probabilities) are a function of several factors including a person's sensitivity to probability information and her relative overweighting of losses as compared to gains ("loss aversion"). In the model, both probability sensitivity and loss aversion can be quantified by adjustable parameters, and several studies have employed CPT to investigate how differences in age (Harbaugh, Krause, & Vesterlund, 2002), gender (e.g., Fehr-Duda, Gennaro, & Schubert, 2006), or personality (Pachur, Hanoch, & Gummerum, 2010) affect risky decision making. Cognitive modeling thus allows individual differences in behavior to be decomposed into underlying cognitive components.

Using individually fitted model parameters to measure individual differences relies on the assumption of *parameter stability*—that is, that the parameter values estimated for a person remain relatively invariant across time (Yechiam & Busemeyer, 2008). This applies in particular when modeling risky decision making, where it is often assumed that people's choices and their cognitive underpinnings reflect stable preferences (Yechiam & Ert, 2011). In principle, however, it is possible that differences in parameter estimates between people simply reflect unsystematic variability (i.e., noise) rather than stable characteristics. In that case, fitting parameters of cognitive models would not be very useful because the results obtained would not generalize beyond a given task or situation.

Glöckner and Pachur (2012) found some evidence for temporal stability of the parameters of CPT: parameters fitted to individual participants' choices at each of two separate experimental sessions were (moderately) correlated. But does this finding also extend to other models of risky decision making? And—more importantly—do conclusions regarding a model's parameter stability depend on how the parameters are estimated? Whereas parameters are traditionally estimated independently for each single participant, it has recently been proposed that more reliable estimates might be achieved by using hierarchical Bayesian procedures, which exploit group-level distributions to inform individual-level estimation (e.g., Gelman & Hill, 2007; Lee & Webb, 2005).

Our goal is to examine whether conclusions regarding the parameter stability of a cognitive model are affected by the statistical method used to obtain the estimates. In particular, we compare hierarchical Bayesian techniques against non-hierarchical Bayesian procedures in a decision-making context. We investigate this issue for the transfer-of-attention-exchange model (TAX; Birnbaum & Chavez, 1997), which has been claimed to provide a better account of decision making under risk than CPT (Birnbaum, 2008). For example, Birnbaum (2008) showed that TAX can correctly account for several violations of CPT, such as violations of gain–loss separability, coalescing, and stochastic dominance, while being able to also accommodate apparent loss aversion and risk aversion.

Hierarchical Bayesian Parameter Estimation

The application of hierarchical Bayesian techniques is becoming an increasingly popular tool to estimate cognitive models, including models of judgment and decision making (Lee & Wagenmakers, 2013; Nilsson, Rieskamp, & Wagenmakers, 2011; Scheibehenne, Rieskamp, & Wagenmakers, 2013). Hierarchical Bayesian techniques are attractive because the approach naturally lends itself to the hierarchical data structure inherent in many psychological experiments, where a single individual provides many observations and researchers aim to draw conclusions on the aggregate group level. The alternative, "independence" approach, by contrast, is to first estimate the parameters of each individual participant separately and then aggregate these measures in a second step (Gelman & Hill, 2007). While feasible, this approach ignores possible similarities between individuals and does not take into account that some participants might allow more precise and reliable estimates than others. Bayesian hierarchical techniques account for these differences and thus promise to yield more consistent and accurate estimates (Rouder & Lu, 2005).

The Bayesian approach achieves this because the imposed hierarchical structure simultaneously informs the individual level, such that potentially unreliable individual estimates can borrow strength from the other estimates (Gelman, Carlin, Stern, & Rubin, 2004). Furthermore, parameter estimates that are deemed unlikely given the distribution of the remaining parameter values (i.e., because they are located at the extremes of the distribution) are pulled closer towards the group mean and implicitly receive less weight. This property is referred to as "shrinkage." For these reasons, it has been argued that hierarchical methods often provide a more thorough evaluation of models in cognitive science (Shiffrin, Lee, Kim, & Wagenmakers, 2008).

Though increasingly popular, Bayesian hierarchical implementations have been developed for only relatively few cognitive models of decision making under risk (but see Nilsson et al., 2011; Wetzels, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2010). Below we develop, to our knowledge for the first time, a hierarchical model for estimating individual participants' TAX parameters.

Transfer-of-Attention-Exchange Model

TAX is a model of how people evaluate risky alternatives that can lead to certain outcomes *x*, each of which occurring with probability *p*. For instance, consider whether you would prefer to play a lottery with a 90% chance of winning $100 (otherwise nothing) or a lottery with a 10% chance of winning $1000 (otherwise nothing). According to TAX, the valuation of a lottery is a weighted average of the utilities of the outcomes; the weight that each outcome receives depends on its rank among all possible outcomes (the *n* outcomes being ordered such that $x_1 < x_2 < x_3 \ldots x_n$) and its probability. To account for the typically found risk aversion (risk seeking) in gains (losses), is the model assumes that attention (i.e., weight) is "transferred" from better (worse) to worse (better) outcomes. Specifically, the valuation, *V*, of a lottery *A* is calculated as

$$V(A) = \frac{\sum_{i=1}^{n}\left[t(p_i) + \frac{\delta}{n+1}\sum_{j=1}^{i-1}t(p_j) - \frac{\delta}{n+1}\sum_{j=i}^{n}t(p_i)\right]u(x_i)}{\sum_{i=1}^{n}t(p_i)}, \quad (1)$$

where δ is a free parameter governing the attention shift from higher to lower outcomes (or vice versa); with $0 < \delta < 1$ attention is shifted from higher (lower) to lower (higher) outcomes in gains (losses), with $0 > \delta > -1$ the opposite would occur. The function $u(x)$ is the utility function, $u(x) = x^\beta$, transforming objective outcomes into subjective utilities. The free parameter β indicates the curvature of the value function and reflects the decision maker's sensitivity to outcome information (with lower values of β indicating lower sensitivity). $t(p)$ is the probability weighting function, transforming objective into subjective probabilities, and equals $t(p) = p^\gamma$. γ is a free parameter reflecting the decision maker's sensitivity to probability information (with lower values of γ indicating lower sensitivity). To derive the predicted probability that lottery *A* is preferred over lottery B, we used an exponential version of Luce's choice axiom:

$$p(A,B) = \frac{e^{\theta \cdot V(A)}}{e^{\theta \cdot V(A)} + e^{\theta \cdot V(B)}}, \quad (2)$$

where θ is a choice sensitivity parameter, indicating how sensitively a decision maker reacts to differences in the valuation of lotteries *A* and *B*. To summarize, TAX as implemented here has four free parameters: attention shift (δ), outcome sensitivity (β), probability sensitivity (γ), and choice sensitivity (θ).

Data We applied TAX to model the data reported in Glöckner and Pachur (2012). In this study, 63 participants (25 male, mean age 24.7 years) indicated their preference between two-outcome monetary lotteries at two experimental sessions that were one week apart. At each session, the participants were presented (on a computer) with 138 lottery problems that contained pure gain, pure loss, and mixed lotteries, all drawn from sets of lottery problems used in previously published studies; 38 of the problems were shown at both sessions (see Glöckner & Pachur for details). The outcomes of the lotteries ranged from –€1000 to €1200. At the completion of each session, one of the chosen lotteries was picked randomly, played out, and the participant received an additional payment proportional to the outcome.

Parameter Estimation

To estimate the free parameters of TAX, we implemented two Bayesian versions of the model—a hierarchical version and an independent (i.e., non-hierarchical) version. Bayesian modeling requires a detailed specification of the likelihood function and the prior probability distributions of all model parameters. For the independent version, we specified the likelihood function based on Equations (1) and (2). The priors for the free parameters were set to uniform probability distributions that span a "reasonable" range that excluded theoretically implausible values and allowed for ample space to include parameter values found in previous research (Michael Birnbaum, personal communication). In particular, the priors ranged from –2 to 2 for the δ parameter and from 0 to 5 for the β, γ, and θ parameters.

In the hierarchical version, we utilized the same functions as in the independent version but partially pooled the individual parameters using normally distributed group-level distributions. Uninformative priors were assigned to the respective means and standard deviation of these group-level distributions. The group-level means were assumed to be normally distributed with mean 0 and variance 1. The prior on the group-level standard deviation was uniformly distributed ranging from 0 to 10. To ensure proper parameter scaling, the group-level parameters were linked onto the individual level through a probit transformation (Rouder & Lu, 2005). As this transformation yields a parameter range from 0 to 1 on the individual level, an additional, linear linkage function was interposed that stretched the parameter range to match the scale used in the independent model implementation outlined above (i.e., a range from –2 to 2 for the δ parameter, and a range from 0 to 5 for the β, γ, and θ parameters).

For both the individual and the hierarchical model we estimated the joint posterior parameter distributions using Monte Carlo Markov Chain methods implemented in JAGS, a sampler that utilizes a version of the BUGS programming language (Lunn, Spiegelhalter, Thomas, & Best, 2009; Plummer, 2011) that was called from the R statistics software (version 2.14.0; R Core Team, 2012). A total of 10,000 representative samples were drawn from the posterior distributions after a "burn-in" period of 1,000 samples. The sampling procedure was efficient as indicated by a low autocorrelation of the samples, the Gelman–Rubin statistic, and visual inspection of the chains.

Quantifying Parameter Stability

To the extent that the parameters of a cognitive model capture stable characteristics of an individual, the parameters should be invariant across time—at least for relatively short time intervals and under comparable measurement conditions (Bland & Altman, 1986). One way to quantify parameter stability (or reliability) is to correlate individual parameter estimates between two points in time (i.e., test and re-test). Higher correlations indicate higher parameter stability.

As outlined above, one rationale for using hierarchical Bayesian techniques for parameter estimation is to obtain more reliable estimates. Thus, one might expect a higher test–retest correlation when parameters are estimated hierarchically than when they are estimated for each participant independently. To test this prediction, we calculated correlations between the parameter values estimated for each participant at the two measurement points (t1 and t2), separately for the individual model and the hierarchical model.

Correlations were calculated based on the mean posterior parameter estimates for each measurement point, using Bayesian techniques implemented in BUGS. A Bayesian approach to calculating correlations allows correlation coefficients to be compared based on their posterior distributions. This avoids many problems inherent in traditional frequentist statistical procedures that rely on null-hypothesis significance testing (Kruschke, 2011).

Results

Table 1 reports the best-fitting TAX parameter values on the group level, obtained from the hierarchical model. As indicated by the δ parameter being larger than 0, participants displayed risk aversion in gains and risk seeking in losses, and some reduced sensitivity to outcomes (β being smaller than 1) and probabilities (γ being smaller than 1). Overall, the parameter values obtained are within the range of values obtained or used in previous applications of TAX (e.g., Birnbaum, 2008).

Table 1: Best-fitting group-level TAX parameters and their 95% highest density intervals (HDI_{95}).

	TAX parameters			
	δ	β	γ	θ
t1				
M	.33	.65	.64	.14
HDI_{95}	[.25,.40]	[.62,.68]	[.57,.71]	[.11,.16]
t2				
M	.35	.63	.61	.16
HDI_{95}	[.27,.43]	[.60,.65]	[.51,.72]	[.13,.20]

Figure 1 shows Pearson's product–moment correlations (across participants) between t1 and t2 for each of the four TAX parameters. As can be seen, the mean correlation coefficient for the δ and the γ parameters is slightly higher when they are estimated hierarchically than when they are estimated independently. However, this difference is not credible, as the 95% highest posterior density interval (HDI_{95}) includes zero. For the β parameter, the correlation is slightly higher when parameters are estimated independently, and for the θ parameter the test–retest correlation is clearly lower for the hierarchical than for the independent estimates. A similar picture emerges based on Spearman's rank correlations (not shown).

Why Does Hierarchical Estimation Fail to Improve Parameter Stability?

The results indicate that applying a hierarchical TAX model does not yield higher parameter stability on the individual level. At first sight, this seems surprising given the supposed advantages of hierarchical techniques that "borrow strength" from distributional information on the group level to improve estimations on the individual level.

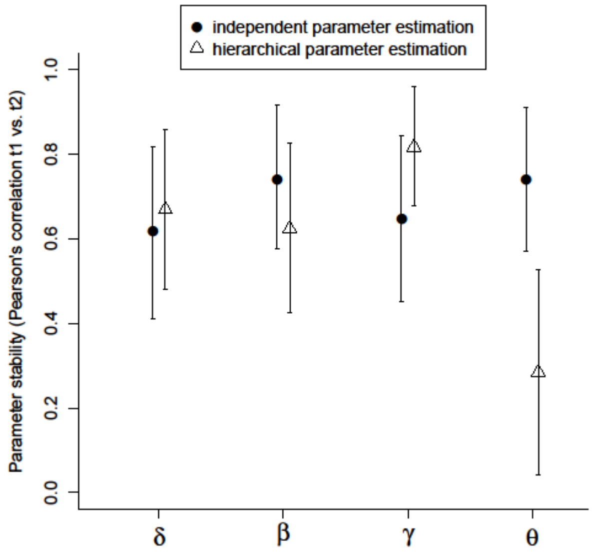

Figure 1: Stability of each TAX parameter as indicated by the mean product–moment correlation across participants between t1 and t2. Circles indicate independent estimates, triangles indicate hierarchical estimates. Error bars = HDI_{95}.

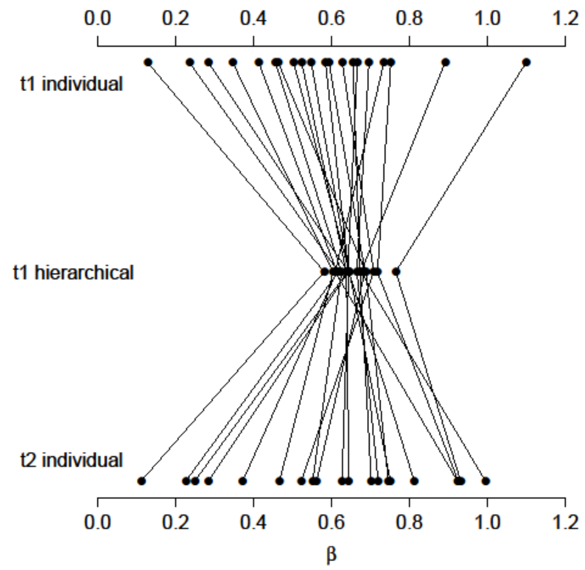

Figure 2: Mean posterior estimates of the β parameter of TAX separately for each individual at t1 and t2 (upper and lower row) and the hierarchically estimated parameters at t1 (middle row) for a representative subset of 20 participants.

To explore the reasons for this result, it is instructive to take a closer look at the distribution of the parameter estimates obtained. For illustration, Figure 2 displays the posterior means for the independently estimated β parameter values at t1 and t2 (upper and lower row, respectively) as well as the hierarchically informed estimates at t1 (middle row) for a subset of 20 representative participants; for each person the estimates are connected by a line. As could be expected, given the partial pooling enforced through the introduction of the higher level group distribution in the hierarchical model, the hierarchical estimates show a lower dispersion than the individually estimated parameters (the same holds for the hierarchical estimates at t2, which are not shown). This shrinkage is particularly pronounced for extreme parameter estimates, that is, those that are far away from the group-level mean. The reason is that these estimates appear rather unlikely with respect to the group-level distribution and are thus implicitly treated as outliers in the hierarchical model.

Unwarranted Shrinkage Importantly, however, Figure 2 further shows that the shrinkage of the hierarchical method is not necessarily warranted: for the independently estimated parameter values there is rather good correspondence between t1 and t2 even for participants with rather extreme parameter values. That is, individuals who have a high β value at t1 also tend to have a high β value at t2; the same applies for small β values. Thus, our analysis shows that in the context on hand extreme estimates often reflect meaningful and reliable characteristics of individuals. The partial pooling enforced by the hierarchical modeling somewhat distorts the individual parameter estimates by pulling them too much towards the group-level mean.

Diminished test-retest correlation The unwarranted shrinkage imposed by hierarchical modeling does not inevitably lead to lower test–retest correlations. After all, it could be that the compressed hierarchical estimates are nevertheless more reliable and thus stable over time than the (more dispersed) parameter values estimated on the individual level. As we will outline next, however, that does not seem to be the case.

Figure 3 displays a scatterplot for the θ parameter separately for the independent and the hierarchical estimates. The θ parameter provides an instructive example because here the difference between the correlations for the individual and the hierarchical estimates is particularly large (Figure 1). Figure 3 shows that the high correlation for the independent estimates is partly due to some individuals having high values on the θ parameter at both measurement points. As mentioned above, although these values are much higher than for most individuals in the group, they nevertheless seem to be reliable in the sense that they are equally high at both measurement points. In contrast, the range of the hierarchically estimated parameters is much narrower (note that the axis scales in the figure were adjusted to best display the data). Furthermore, the hierarchical model seems to affect the individual parameter estimates to different degrees. This occurs because the influence of the group-level depends, among other aspects, on the variance and the mean of the individual estimates. As indicated by the shape of the scatterplot in the lower panel of Figure 3, this effect pulls the parameter estimates towards the mean and thus leads to a lower (linear) relationship between the two measurement points. In that sense, the hierarchical method also induces shrinkage on the correlation coefficients. In situations where the correlation of the individually estimated parameters is reduced due to unreliable outliers, however, applying hierarchical techniques will shrink these outliers and may then yield higher parameter stability.

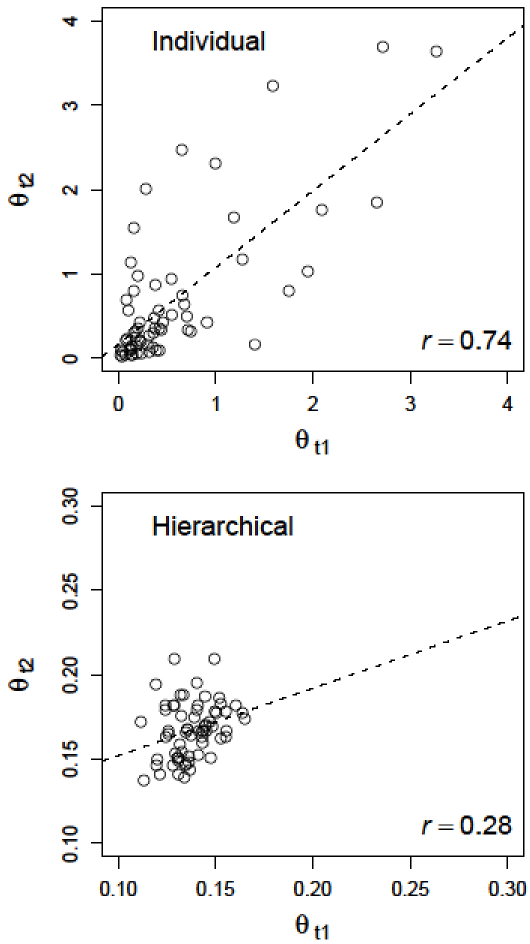

Figure 3: Scatter plot of the mean posterior estimates for the θ parameter at t1 and t2. Each point represents one participant. The upper panel shows the parameter values obtained by individual estimation; the lower panel shows the parameter values obtained by hierarchical estimation. Note that the value ranges on the axes are much smaller in the lower panel.

Discussion

The psychological content and generalizability of a cognitive model hinges on the extent to which its parameters reflect stable characteristics of an individual. Conclusions regarding a model's parameter stability may be affected by the statistical procedures used to estimate these parameters. Specifically, researchers must decide whether to employ hierarchical techniques or to estimate each person individually.

Our analyses show that the free parameters of the TAX model are rather consistent across time, indicating that the model captures stable aspects of decision makers' risk attitude and their outcome and probability sensitivity. This finding parallels previous results obtained for CPT based on the same data using maximum likelihood estimates (Glöckner & Pachur, 2012). Most importantly—and rather unexpectedly—our analysis provided no evidence that hierarchical Bayesian techniques yield more stable parameter estimates than the alternative approach of estimating each participant independently from the others.

Why did the shrinkage of the hierarchical procedure yielding distorted estimates? In principle, one possibility is that the distribution of the individual parameter values is bi-modal, which would render group-level means futile. As indicated by visual inspection, however, the parameter distributions for our data were mostly unimodal in shape, so this cannot explain why the hierarchical procedure distorted the estimates.

Another possibility could be the prior distribution used for shrinkage. To achieve an optimal balance between complete pooling and complete independence, the degree of shrinkage in the hierarchical model is represented by a free parameter (representing the variance of the group-level distributions) estimated from the data. In principle, the choice of prior on the variance could impose an unwarranted amount of shrinkage (i.e., a low variance), for instance, if much weight is put on low variances, or if the prior does not allow for large variances in the first place. For the current data, however, the posterior estimates for the group-level standard deviations were far away from the upper boundaries of the uniform prior distributions on the group-level. The choice of prior on the variance of the group-level distributions is thus an unlikely reason for the undue amount of shrinkage.

Generalizability Although our demonstration focused on one particular cognitive model, we suspect that the conclusions will hold for other models as well—particularly in the domain of judgment and decision making; here, people often rely on different strategies (e.g., Pachur & Olsson, 2012; Scheibehenne et al., 2013) and parameter heterogeneity thus reflects genuine differences between people. In such a case, the parameter estimates will not regress towards the mean if more data or more precise measures are collected.

Advantages of Hierarchical Approaches The case on hand may be different from situations in which hierarchical Bayesian techniques have been shown to outperform independent parameter estimates. In a classic example, Efron and Morris, (1975) predicted the success rate of professional baseball players for an entire season based on their success rate early in the season. This prediction was greatly improved through the application of hierarchical techniques. Presumably, this improvement occurred because the differences in the success rates of professional baseball players are rather small (they are all pretty good players) and random noise will equal out throughout the season.

Under this condition, there will be regression towards the mean, which benefits hierarchical Bayesian techniques.

Another situation in which hierarchical Bayesian estimates presumably provide more accurate results than independent estimates is when only very little data is available for each individual, yielding high uncertainty on the individual level. Here, the unreliability on the individual level might be reduced through partial pooling.

Finally, hierarchical modeling techniques might be beneficial for comparisons on the group level (Gelman & Hill, 2007), where the goal is not to improve the reliability on the individual level but to derive robust estimates for the group as a whole. As a consequence, the implicit weighting imposed through hierarchical estimation methods might also be advantageous for making out-of-sample predictions for new group members.

Summary

Our results indicate that hierarchical Bayesian techniques do not necessarily improve the reliability of individual parameter estimates. Therefore, researchers aiming to predict individual behavior may be better advised to rely on individual estimates instead. As discussed above, hierarchical models might have specific strengths in situations in which very little information is available on the individual level, when the group is very homogenous, or when the goal is to describe and compare groups as a whole.

References

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review*, *115*, 463–501.

Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes*, *71*, 161–194.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *i*, 307–310.

Efron, B., & Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association 70,* 311–319.

Fehr-Duda, H., de Gennaro, M., & Schubert, R. (2006). Gender, financial risk, and probability weights. *Theory and Decision*, *60*, 283-313.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: Parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*, 21–32.

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, *5*, 53-84.

Kruschke, J. K. (2011). *Doing Bayesian data analysis: A tutorial with R and BUGS.* Amsterdam: Elsevier.

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

Lee, M. D., & Webb, M. R. (2005). Modeling individual differences in cognition. *Psychonomic Bulletin and Review*, *12*, 605–621.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in Medicine, 28*, 3049–3067.

Nilsson, H., Rieskamp, J., & Wagenmakers, E.-J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, *55*, 84–93.

Pachur, T., Hanoch, Y., & Gummerum, M. (2010). Prospects behind bars: Analyzing decisions under risk in a prison population. *Psychonomic Bulletin and Review*, *17*, 630–636.

Pachur, T., & Olsson, H. (2012). Type of learning task impacts performance and strategy selection in decision making. *Cognitive Psychology*, *65*, 207-240.

Plummer, M. (2011). *JAGS Version 3.1.0.* Retrieved from http://mcmc-jags.sourceforge.net/

R Core Team (2012). R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* Vienna, Austria.

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin and Review, 12,* 573–604.

Scheibehenne, B., Rieskamp, J., & Wagenmakers, E.-J. (2013). Testing adaptive toolbox models: A Bayesian hierarchical approach. *Psychological Review*, *120*, 39–64.

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.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

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.

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

Yechiam, E., & Ert, E. (2011). Risk attitude in decision making: Search for trait-like constructs. *Topics in Cognitive Science*, *3*, 166–186.