A tutorial on response-time extended multinomial processing tree models in social cognition

Liz Wilson[1], Raphael Hartmann[2], Karl Christoph Klauer[2]

Ruben Laukenmann[3], & Jimmy Calanchini[1]

[1] Department of Psychology, University of California, Riverside

[2] Department of Psychology, University of Freiburg

[3] School of Social Science, University of Mannheim

Address for correspondence: Jimmy Calanchini, Department of Psychology, 900 University Ave., University of California, Riverside, CA, 92521 Email: jimmy.calanchini@ucr.edu

Total Word Count for Research Article: 7,747 (excluding code, tables, figures, and references)

Abstract

Multinomial processing tree (MPT) models can provide novel insights into the cognitive processes underlying a wide variety of social cognitive judgments and behaviors. In previous research, MPTs have been used to disentangle the contributions of multiple latent processes to tasks configured to assess moral reasoning, processing fluency, decision-making, implicit biases, and social categorization, among many other topics. However, until recently, MPT models were limited in their application to categorical data. New methodological advances extend traditional MPT estimation methods by incorporating reaction time data, thereby expanding the breadth and depth of questions that can be investigated. This article provides a user-friendly step-by-step tutorial for response-time extended MPT methods with annotated code and data, using the Implicit Association Test as a working example.

A tutorial on response-time extended multinomial processing tree models in social cognition

Social and cognitive researchers have long been interested in the joint influence of multiple qualitatively distinct processes that drive thoughts and behaviors. Given that the phenomena psychologists study are often complex and multiply determined, researchers have developed tools to dissociate and quantify the unique role of different processes on task performance. One such tool is multinomial processing tree (MPT: Batchelder & Riefer, 1999) modeling, which belongs to a formal class of mathematical models that describe and quantify the multiple processes that account for responding in tasks. The general utility of MPT models has been recognized in a broad range of research areas including but not limited to recognition memory (Jacoby, 1991), source monitoring (Batchelder & Riefer, 1990; Klauer & Meiser, 2000), hindsight bias (Erdfelder & Buchner, 1998), false memory (Jacoby et al., 2005), moral judgment (Gawronski et al., 2017), stereotyping (Krieglmeyer & Sherman, 2012), and prejudice (Conrey et al., 2005; Meissner & Rothermund, 2013). In the context of social cognition specifically, MPT models have been applied to a variety of tasks, such as the affect misattribution procedure (Payne et al., 2010), the go/no-go association task (Nadarevic & Erdfelder, 2011), the extrinsic affective Simon task (Stahl & Degner, 2007), the Weapon Identification Task (Payne, 2001), and the Implicit Association Test (Greenwald et al., 1998). Traditionally, MPT modeling relies solely on response frequencies (e.g., correct versus incorrect responses; see Schmidt et al., 2023 for an excellent tutorial on this approach). However, two recent methods have been introduced to incorporate response latencies into MPT modeling (Heck & Erdfelder, 2016; Klauer & Kellen, 2018), which have the potential to greatly extend the utility of the MPT modeling method. This article provides a user-friendly step-by-step tutorial for the response-time extended MPT methods developed by Heck and Erdfelder (2016) and Klauer and Kellen (2018) with annotated code and data. Given the outsized influence of the IAT in social cognition research, we use it as the working example in this tutorial. We begin with an illustration of how modeling IAT responses using MPT methods can provide greater insight compared to alternative scoring approaches, discuss some theoretical contributions of extending MPT models to account for response times, then demonstrate the two major approaches for response-time extended MPT modeling with working examples.

MPT Modeling versus Summary Statistics

Implicit measures, such as the IAT, were developed to overcome individuals' inability and/or unwillingness to accurately report their evaluations of social groups. To do so, implicit measures are typically configured with task conditions aimed to constrain the influence of deliberative processing, social desirability biases, and introspection failures – thereby facilitating the expression of relatively automatic associations linking social groups (e.g., Black versus White people) with evaluations (e.g., good versus bad). Traditionally, IAT responses are quantified in terms of the D-score (Greenwald et al., 2003) based on differences in response latencies to blocks of trials in which social groups share one set of attribute pairings (e.g., White/positive, Black/negative) versus another set of pairings (White/negative, Black/positive). Because such measures were initially assumed to primarily reflect the influence of automatic associations, the latency-based D-score summary statistic is often interpreted to reflect the relative strength of mental associations. Consequently, the dominance of the D-score as the primary method of quantifying IAT responses has influenced the interpretation of research findings in this domain. For example, research exploring age differences in responses to the race IAT has found that older people's D-scores reflect greater pro-White/anti-Black bias compared to younger people's D-scores (Nosek et al., 2002). This result aligns with historical negative portrayal and treatment of Black people and suggests that older people demonstrate stronger implicit biases because they learned more racially biased associations than did younger people who grew up in an ostensibly more egalitarian society.

Though this interpretation of age-related differences in implicit bias is reasonable, research using MPT models suggests a different interpretation (Calanchini, 2020; Hütter & Klauer, 2016). Recognizing that D-scores can quantify the direction of bias (e.g., a relative, evaluative preference for Black versus White people) but cannot provide insight into the cognitive processes underlying this bias, researchers used MPT models to quantify the joint contributions of multiple processes to responses on implicit measures to better understand this apparent age-related difference in implicit bias. The quad model (Conrey et al., 2005) is an MPT model that proposes that IAT responses reflect not only mental associations between social groups and attributes, but also an inhibitory cognitive process that can constrain the expression of associations, among other processes. D-scores cannot distinguish between the contributions of associations and inhibition to IAT responses. Research applying the quad model to the race IAT responses of older and younger people found that racial associations did not vary between age groups, but instead that older people's inhibitory responses were weaker than those of younger people (Gonsalkorale et al., 2009). Thus, the quad model suggests that deficits in inhibition, rather than differences in associations, can explain IAT responses across age groups. As this example illustrates, MPT models can provide more theoretically-precise insight into the cognitive processes underlying social cognitive behaviors than traditional summary statistics. However, MPT models are also limited because, until recently, they relied solely on response frequencies (e.g., correct, incorrect) to the exclusion of response latencies.

Response-Time Extended MPT Models

To date, most research using MPT models has relied on categorical data (e.g., frequencies of correct/incorrect responses), excluding latency data. Consequently, neither D-scores nor MPT models make full use of the available frequency and latency data that is typically collected in social cognitive research. However, recent methodological advances extend MPT models to account for response times (Heck & Erdfelder, 2016; Klauer & Kellen, 2018). Such response-time extended MPT models not only produce estimates of the extent to which multiple cognitive processes influence responses, but they also produce estimates of the latent processes' completion times or the relative speed of processing paths – the sequence of processes that lead to a response. Whereas summary statistics like the IAT D-score rely primarily on response latencies and MPT modeling relies solely on response frequencies, response-time extended MPT models account for both bases of information. In doing so, they are more psychologically comprehensive than other analytic approaches, thereby opening up numerous avenues for exploration and tests of the kinds of dual-process theory that underlie much of the social cognitive literature.

Rooted in traditional dual-process theory, MPT models always assume that judgments and behaviors are driven by multiple latent processes that sometimes work in tandem but sometimes compete against one another. One core tenet of dual-process theory is that automatic processing is faster than controlled processing – an assumption that can be readily modeled and tested using response-time extended MPT models. Furthermore, researchers can use response-time extended MPT models to test different assumptions about the nature and temporal interplay of the latent processes specified in the model structure. For example, (Klauer & Voss, 2008) propose multiple interpretations of the cognitive processes that contribute to responses on the

Weapon Identification Task (WIT; Payne, 2001), which is a sequential priming task designed to assess the effects of racial priming on the correct identification of weapons and tools. Each interpretation relies on distinct assumptions about the time course by which the processes unfold, and analyses based on response frequencies alone cannot distinguish among these interpretations. Consequently, (Laukenmann et al., 2023) used response-time extended MPT models to investigate the relative speed of each processing path that contributes to WIT responses. They found that the model in which automatically-evoked associations interfere with correct object identification from the outset of each trial best fit the observed data, thereby advancing our understanding of the cognitive processes that contribute to responses on the WIT. When theories or assumptions that are relevant to a researcher's objectives cannot be distinguished using the traditional MPT modeling framework, response-time extensions of MPT models serve as a viable alternative approach.

This tutorial focuses on two prominent approaches for extending MPT models with response time data. *RT-MPT* models (Klauer & Kellen, 2018) estimate the response times for each individual cognitive process within the processing paths of a model. In contrast, *MPT-RT* models (Heck & Erdfelder, 2016) estimate the relative speed of a sequence of cognitive processes in the model's paths by splitting response time data into discrete bins. We begin with a discussion of the major differences between these two model classes. For each approach, we formalize the well-validated quad model (Conrey et al., 2005) as a response-time extended MPT model in the context of the IAT. Then, we provide guidance to (a) specify a model using appropriate syntax, b) prepare the IAT data for modeling, c) fit the model to the IAT response frequency and response time data, d) interpret the model output, and e) test hypotheses involving process times. All data and code are freely available and can be accessed at https://osf.io/rnbhd/.

Selecting RT-MPT versus MPT-RT Models

Your decision to rely either on an RT-MPT or MPT-RT model should be theoretically grounded and will likely depend on your research design and objectives. Table 1 summarizes some major differences between both models. One prominent difference between the two model types is that each integrates RTs differently. MPT-RTs are situated within the framework of traditional frequency-based MPT models, in that observed RTs are not treated as continuous variables but instead transformed into bins for each individual and then used as additional discrete response categories. In contrast, RT-MPTs jointly model response frequencies and their respective latencies, which ultimately provides the researcher with estimates of process probabilities and average process completion times.[1] Thus, the first major difference between response time-extended MPT modeling methods is that MPT-RT models do not produce process completion times for each process specified in the model, whereas RT-MPT models do. As such, MPT-RT models could be used to instead investigate the relative speed of different processing paths. Another difference between the two model types is that RT-MPT models produce estimates of encoding and motor execution times, but MPT-RT models do not. This feature may be particularly valuable if a researcher suspects that encoding or motor responses may vary across different stimuli or experimental conditions and would like to account for that source of variance within the model.

Another difference between model types is the assumptions each makes about the distributions of RTs. RT-MPT models assume that the completion time for each process is

[1] RT-MPT models assume independence of processing times within participants at the level of individual trials, such that the time required for completing one quad process is uncorrelated with completion time of another quad process (e.g., overcoming bias takes the same amount of time irrespective of how fast activated associations took previously). In contrast, due to the hierarchical structure of the model, correlations between the average process completion times of two processes estimated per participant can be accommodated and estimated by RT-MPT models.

exponentially distributed and that the motor plus execution time is normally distributed with a truncation from below at zero. However, this assumption may not be tenable across all experimental designs and paradigms and can potentially result in distorted results if the distributional assumptions do not hold. In contrast, MPT-RT models avoid imposing assumptions on the shape of RT distributions, which is potentially an advantage over RT-MPT models to the extent that it avoids concerns about violations of the distributional assumptions of the underlying process and motor times. That said, RT-MPT models assume that processing paths proceed serially, such that process completion times add up along the processing path. This assumption may be useful for testing the sequential order of the processes, but it may be too restrictive or theoretically fail to account for the observed data. For example, this assumption of seriality may not be compatible with other reasonable perspectives on cognition, such as race models or parallel competitive models in which two different processes run in parallel. MPT-RT models make no such assumption about whether the processes within a processing path unfold serially or in parallel and, thus, are relatively more flexible than RT-MPT models on this dimension.

Because MPT-RT models are situated within the framework of frequency-based MPT models, researchers can rely on a relatively thriving methodological toolbox that has already been developed. Software for frequency-based MPT modeling include the MPTinR (Singmann & Kellen, 2013) and TreeBUGS (Heck, Arnold, et al., 2018) *R* packages, as well as the stand-alone multiTree (Moshagen, 2010) and HMMTree (Stahl & Klauer, 2007) programs. Thus, MPT-RT models may be particularly user-friendly for researchers who are already familiar with this software and their associated file types and statistical output. Additionally, Hartmann and colleagues (2020) recently introduced the rtmpt package to facilitate RT-MPT modeling in the R environment.

All MPT models – both traditional and RT-extended – must be able to infer unique parameter values from observed data (i.e., be identifiable)[2]. An advantage of the RT-MPT approach is that incorporating response time data can produce an identified model that would have otherwise not been identified on the basis of frequency data alone.[3] In contrast, MPT-RT models are typically not identified if the underlying MPT model is not identified. In addition, MPT-RT models involve additional parameters that may require the researcher to impose theoretically-grounded model constraints to achieve identifiability, and researchers may find it difficult to justify their proposed constraints. We recommend Heck and Erdfelder (2016) for further reading on MPT-RT model identifiability, specifically, as well as Schmittmann et al. (2010) for further reading on MPT model identifiability more generally.

In summary, researchers may favor RT-MPT models when their research objective involves testing assumptions about the order of serial processes, when they are interested in process completion times for individual processes, or when the underlying MPT model would not be identified on the basis of frequency data alone. Researchers may favor MPT-RT models when they have assumptions about the serial nature of processes, when the research objective centers on processing paths rather than individual processes, or when they want to avoid imposing distributional assumptions.

[2] The identifiability of a model can be heuristically investigated by fitting the model to simulated data with known underlying parameter values and evaluating if the model produces parameter estimates sufficiently close to the true values (Schmittmann et al., 2010).

[3] RT-MPT models possess this feature of identifiability because they not only predict frequencies and mean response latencies, but also variance, skewness, etc. of response-time distributions for each response. Each of these estimates are non-redundant and effectively add an independent model equation, resulting in a system of equations that can in most cases be solved uniquely for a given set of parameter values.

Table 1

A Comparison of RT-MPT and MPT-RT Models

	Model Type	
Major Difference	RT-MPT	MPT-RT
1. Process Completion Times	Provides estimates of average process completion times.	Provides estimates of the relative speed of processing paths.
2. Encoding Plus Motor Execution Time	Provides estimates of the time required for encoding task stimuli and motor execution of the response.	Does not estimate encoding and motor execution time.
3. Distributional Assumptions	Assumes the completion time for each process is exponentially distributed. Assumes the motor plus execution time is normally distributed with a truncation from below at zero.	Avoids distributional assumptions by subclassifying response times into bins.[4]
4. Seriality vs. parallelism	Assumes serial processes within paths.	Avoids assumptions about the serial or parallel nature of processes.
5. Identifiability	Models that are not identified for frequency-based MPT models (including MPT-RT) may be rendered identified.	Typically not identified if the underlying MPT model is not identified.

RT-MPT Models

RT-MPT models (Klauer & Kellen, 2018) belong to a class of formal mathematical models that link latent cognitive processes to response latencies and frequencies on tasks like the

[4] Heck, Erdfelder, and Kieslich (2018) proposed generalized processing tree models, an extension of MPT-RT models that rely on distributional assumptions for path completion times. To the best of our knowledge, this approach has not yet been developed for use within the Bayesian hierarchical framework.

IAT. Like MPT models, they provide estimates of the likelihood that each process specified in the model influenced IAT responses. Additionally, RT-MPTs provide process completion times – the time it takes for a process to get to either of two outcomes (e.g., success, failure) – for each process specified in the model, as well as encoding and motor execution times. This tutorial walks through the RT-MPT model class that assumes exponentially distributed process completion times and truncated normally distributed encoding and motor execution times,[5] which is implemented in the *R* software package 'rtmpt' (Hartmann et al., 2020). For more information about the auxiliary assumptions on process times and motor times, see Klauer and Kellen (2018) and Klauer et al. (2024). We begin by specifying the quad model, pre-processing and transforming IAT data for modeling, fitting the quad model, and interpreting the output.

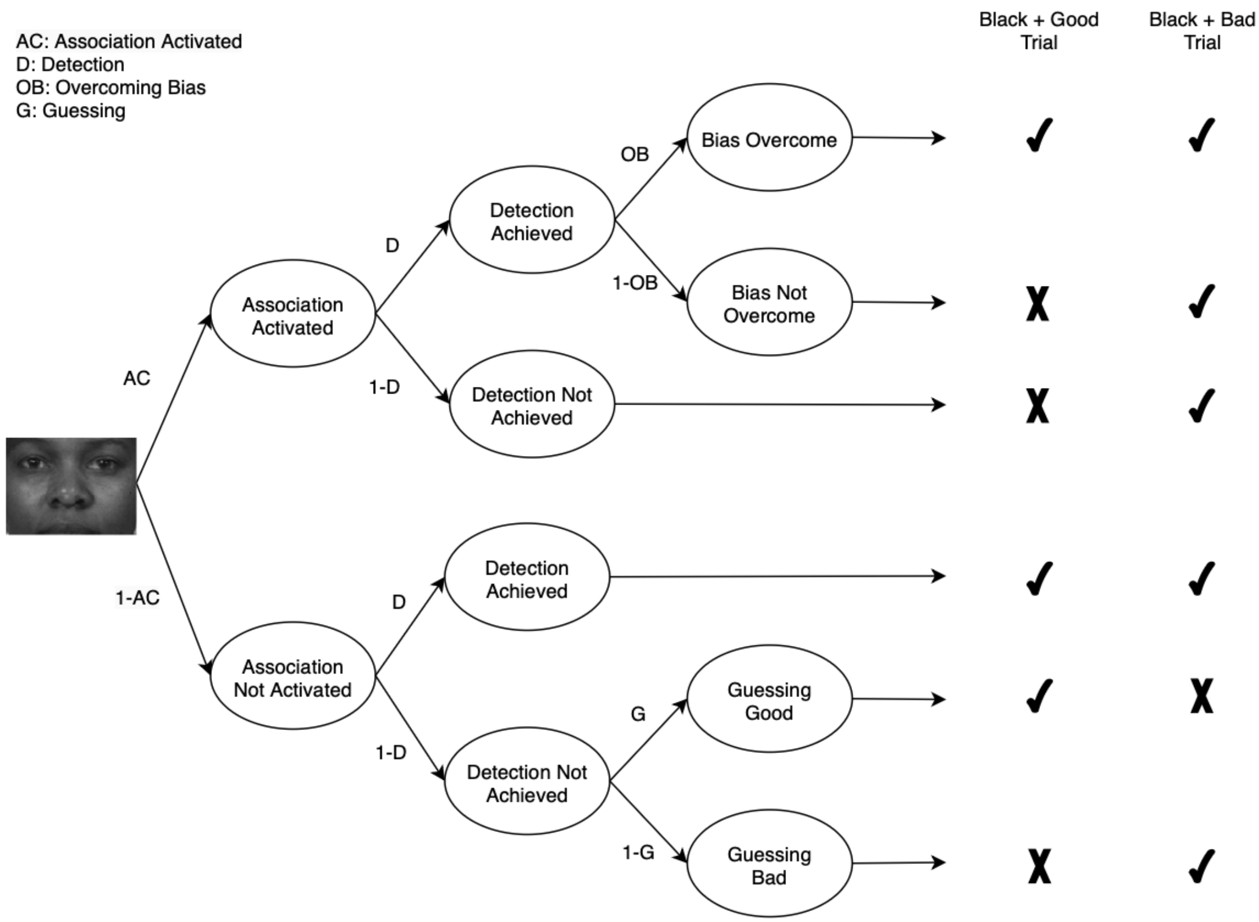
The Quad Model

The quad model (Conrey et al., 2005) proposes that four qualitatively distinct processes influence performance on the IAT. The proposed interplay among the processes specified in the quad model can be illustrated in a processing tree (Figure 1).

[5] RT-MPT models can also be specified using an alternative distributional assumption of cognitive processes following a Wiener-Diffusion process. This approach assumes that the process completion times follow a first-passage time distribution of a Wiener-Diffusion model and the encoding and motor execution times follow a truncated t-distribution (Klauer et al., 2024).

Figure 1

A Portion of the Quad Model



Note. The table on the right illustrates correct (✓) and incorrect (X) responses across different trial types. Parameters with lines leading to them are conditional upon the preceding parameters.

The *Associations* parameter refers to the degree to which mental associations are activated when responding to a stimulus. All else equal, the stronger the mental link between the target (e.g., White) and the attribute (e.g., good), the more likely the Association is to be activated and drive responses in an association-consistent direction. The quad model estimates two different Associations parameters and is typically configured such that one parameter reflects an association between White and positive evaluations, and the other parameter reflects an association between Black and negative evaluations. As illustrated in Figure 1, the Associations parameter is situated at the root of the processing tree. The *Detection* of correct responses parameter is conceptualized as an accuracy-oriented process and reflects the likelihood that the participant can discern the correct response. Sometimes activated associations conflict with the detected correct response. For example, on trials in which White faces appear and the categories 'White' and 'bad' share a response key, to the extent that a participant associates 'White' with 'good' then activated associations would conflict with the detected correct response. The quad model proposes an *Overcoming Bias* parameter to resolve such a conflict between Associations and Detection. Overcoming Bias refers to an inhibitory process that prevents associations from influencing behaviors when they conflict with detected correct responses. As illustrated in Figure 1, the Overcoming Bias parameter is situated along the top two paths of the processing tree. Finally, the *Guessing* parameter does not represent a specific process, per se, but instead reflects the tendency to respond with 'good' versus 'bad' in the absence of influence from Associations, Detection, and Overcoming Bias.

Specifying RT-MPT Models (Klauer & Kellen, 2018 Approach)

For the RT-MPT approach (Klauer & Kellen, 2018), we begin by specifying the quad model using the well-established EQN syntax in which each line defines a single path probability.[6] An annotated version of the quad EQN file ('quad.eqn') and an equation key ('EquationKey.txt') can be found in our OSF project page. The first few lines of the quad model are:

```
1   ;   t01   ;   ACwg1
1   ;   t01   ;   (1-ACwg1)*D1
1   ;   t01   ;   (1-ACwg1)*(1-D1)*G1
```

[6] Researchers may also specify the MPT model using MDL format (Singmann & Kellen, 2013) in which all path probabilities are specified within the same line of code separately for each response category. For a more detailed description of MDL format, see Singmann and Kellen (2013) and Hartmann et al., (2020).

```
1   ;   t02   ;   (1-ACwg1)*(1-D1)*(1-G1)
```

The left column specifies the IAT trial type ('1'): in this case, trial type '1' refers to responses to White stimuli when 'White' and 'good' share a response key. The middle column specifies categories for correct ('t01') and incorrect ('t02') responses. The right column specifies the product of quad parameters theorized to underlie responses for each processing path. Each column is separated by semicolons. The first three lines describe how a correct response can be produced by: activated White-good Associations alone; successful Detection in the absence of activated Associations; or a response bias towards 'good' in the absence of activated Associations and Detection. In contrast, the fourth line describes how an incorrect response can be produced by a response bias towards 'bad' in the absence of activated Associations and Detection. For each line, the parameters must be specified in the order they appear along the processing path, starting at the left with the root of the path extending to the end of the path on the right. For example, in the fourth line of the quad EQN, the processes must be specified as '(1-ACwg1)*(1-D1)*(1-G1)' and not any other order. Once the researcher specifies the MPT model, it must be translated to a model class that can be used to fit RT-MPT models. The quad model specification can be saved as an *R* object:

```
eqn="
1   ;   t01   ;   ACwg1
1   ;   t01   ;   (1-ACwg1)*D1
1   ;   t01   ;   (1-ACwg1)*(1-D1)*G1
1   ;   t02   ;   (1-ACwg1)*(1-D1)*(1-G1)
…
8   ;   t15   ;  (1-ACbb1)*(1-D1)*(1-G1)
"
```

Then, the researcher can use the `to_ertmpt_model()` function to translate the EQN to a format that can be read and used in subsequent modeling functions:

```
quad_eqn <- to_ertmpt_model(eqn_file = eqn)
```

Alternatively, the researcher can provide a path to the text file:

```
quad_eqn <- to_ertmpt_model(eqn_file = “./quad.eqn”)
```

Data Source

The example IAT data that forms the basis of this tutorial was configured to assess race evaluations and consists of stimuli reflecting two target categories (Black, White) and two attribute categories (good, bad). This data was collected online from 138 undergraduate students and can be downloaded from our OSF project page. We include three versions of this data, two of which are relevant to RT-MPT models specifically:

1. iatData.csv: contains the raw IAT trial-level data describing the trial type, response time, and accuracy of each response for all 138 participants (63.77% female; age: M = 19.91 years, SD = 3.25). This sample consists of undergraduate students at a public university in California, and a subset of these data were reported in Elder et al. (2023). All data used in this tutorial were collected online using Millisecond’s Inquisit software, which can achieve the millisecond-precision timing that is crucial for rt-extended modeling.
2. rtmptData.csv: reflects responses that have been cleaned and formatted for RT-MPT modeling. This data is in long format such that each line describes a unique IAT trial.
3. mptrtData.csv: reflects responses that have been cleaned and formatted for MPT-RT modeling. This data is in wide format such that each line summarizes a participant’s total number of responses for each response category.

Data Preprocessing and Transformation

Response latency outliers often lead to convergence problems for RT-MPT models, especially when the outliers arise due to extraneous processes unrelated to the processes of interest specified in the model, such as distracted responding.[7] Consequently, researchers may choose to apply an outlier-exclusion criterion to address unreasonably slow or fast responses. For researchers interested in applying these exclusions to their own data, our OSF project page includes annotated R code that excludes very slow and very fast responses in each individuals' RT distribution using Tukey's outlier criterion. The Gibbs sampler used in this RT-MPT implementation is sensitive to outliers, so these preprocessing steps may lead to a quicker and more efficient model estimation.

Here we describe how to transform raw IAT response data ('iatData.csv') to a format appropriate for model fitting ('rtmptData.csv'). The necessary variables in the data frame should include the subject identifier (subj), the experimental group the subject was assigned to (group), the tree of the current trial (tree), the observed response category (cat), and the observed response time in milliseconds (rt) as outlined below:

subj	group	tree	cat	rt
52641	1	3	t05	579
52641	1	1	t01	1316
52641	1	3	t05	762
52641	1	1	t01	678
52641	1	4	t08	440

The 'tree' variable corresponds to the trial type categories found in the left column of the EQN file, and the 'cat' variable corresponds to the response categories found in the middle

[7] Excluding outliers that arise from the distribution of response times predicted by the modeled processes is likely to lead to systematic bias (Ulrich & Miller, 1994). However, simulation studies exploring the size of such bias and the conditions under which bias arises are rare and would be a fruitful direction for future research.

column of the EQN file. The 'group' variable can be used to specify separate experimental groups for which the researcher wishes to estimate separate parameters. For example, the researcher could assign 0 to a control group and 1 to an experimental group, which would produce separate group-level parameters for each group. The model fitting function requires data to be formatted such that all variables, besides response times, start at 0. The code below carries out this necessary transformation on the IAT data.

```
data <- to_ertmpt_data(raw_data = MPT_DF, model = quad_eqn)
```

After running this code, the subset of data shown above would now be formatted properly for model fitting and would look like the lines of data below. For example, the observed response category variable would be recoded such that t01 is now 0, t02 is now 1, t03 is now 2, etc.

subj	group	tree	cat	rt
0	0	2	4	579
0	0	0	0	1316
0	0	2	4	762
0	0	0	0	678
0	0	3	7	440

Model Fitting

To fit RT-MPT models to data, researchers can use the model-fitting function `fit_ertmpt()`. This function samples from the posterior distribution using a Metropolis-Gibbs sampler and returns the posterior samples saved as an mcmc.list. The code below outlines the function arguments to fit an RT-MPT model to data. Two of these arguments must be specified by the researcher: model and data. The other nine arguments can be omitted, in which case the algorithm will rely on default values and settings. For this Bayesian estimation method, the default settings assume weakly noninformative priors. For detailed descriptions of these other arguments, see the supplementary materials as well as the package vignette of 'rtmpt'. Researchers seeking additional information about Bayesian inference and estimation in R may find (Kruschke, 2015) to be helpful in providing accessible information and step-by-step instructions.

```
rtmpt_out <- fit_ertmpt(model = quad_eqn,
                                data = data,
                                n.chains = 4,
                                n.iter = 5000,
                                n.burnin = 200,
                                n.thin = 1,
                                Rhat_max = 1.05,
                                Irep = 1000,
                                prior_params = NULL,
                                indices = FALSE,
                                save_log_lik = FALSE,
                                old_label = FALSE)
```

Once the researcher has run the model-fitting function, sampling will continue until the Rhat value is less than the value specified in the `Rhat_max` argument. When the model converges, researchers can then use the code below to save the fitted model. If the model fails to converge or converges slowly, the problem may be a poor starting value for that particular estimation, or one or more of the chains has gotten lost exploring a different area of the posterior distribution compared to the other chains. In either case, the solution may be to stop and restart the estimation with a new starting value. Additionally, a model that fails to converge may be poorly specified, in which case you may need to consider alternative model specifications.

```
save(rtmpt_out, file="rtmpt_out.Rdata")
```

Interpreting Model Output

Model Fit

To evaluate model fit, researchers can rely on the T_1 statistic (Klauer, 2010), which summarizes how well the model accounts for the pattern of observed response latencies and frequencies (e.g., the number of correct and incorrect IAT responses) aggregated across participants. The posterior predictive p-values (PPP) for T_1 quantify the discrepancy between the expected and observed data, with PPP < .05 indicating poor model fit. In our example, the observed and expected values were 17.49 and 7.62, respectively, $p = .03$. The code below produces T_1 for the frequency and latency data.

```
rtmpt_out$diags$PostPredCheck_Frequencies

rtmpt_out$diags$PostPredCheck_Latencies
```

T_1 for the frequency data corresponds to the goodness-of-fit chi-square statistic used in traditional modeling approaches (Batchelder & Riefer, 1999). Consequently, T_1 depends on the number of observations, with more observations providing greater power to detect even a small degree of model misfit. Though our p-value would indicate model misfit, we have little insight into its degree. For this reason, we recommend also calculating W using T_1observed for the RT-extended frequencies. W scales T_1 by the number of observations according to the equation below. Based on (Cohen, 1992) effect size recommendations, $W < 0.1$ is heuristically interpreted as indicating acceptable model fit. In our example, we compute a W-statistic of 0.03, suggesting that the model fit was acceptable.

$$W = \sqrt{\frac{T_1 observed}{n\ participants * n\ trials\ per\ participant}}$$

Group-Level Parameter Estimates

The code below produces the posterior mean and median of the group-level parameters, as well as 95% Bayesian credible intervals for each parameter specified in the RT-MPT model.

The output of this code is illustrated in Figure 2. For the full summary output from this model, see our project OSF page.

```
rtmpt_out$summary$transformed_pars
```

Figure 2

R Output of RT-MPT Group-level Parameter Estimates

	Mean	SD	2.5%	50%	97.5%	Naive SE	Time-series SE	n.eff	Rhat	R_95%
theta_ACbb1	0.061	0.018	0.027	0.061	0.098	0.000	0.000	1469.992	1.004	1.012
theta_ACwg1	0.065	0.022	0.025	0.064	0.109	0.000	0.001	1216.746	1.026	1.077
theta_D1	0.909	0.013	0.883	0.909	0.933	0.000	0.000	2231.733	1.006	1.016
theta_G1	0.602	0.048	0.509	0.600	0.698	0.000	0.001	1731.898	1.002	1.006
theta_OB1	0.806	0.103	0.567	0.817	0.969	0.001	0.004	762.691	1.036	1.093
E(tau_minus_ACbb1)	27.810	8.824	14.538	26.423	48.704	0.044	0.421	457.946	1.051	1.137
E(tau_minus_ACwg1)	39.397	9.396	23.695	38.505	60.110	0.047	0.428	491.342	1.046	1.130
E(tau_minus_D1)	25.752	20.119	8.241	18.340	82.880	0.101	2.315	154.119	1.535	2.694
E(tau_minus_G1)	40.392	24.289	12.173	33.606	101.884	0.121	1.888	202.578	1.575	2.516
E(tau_minus_OB1)	64.861	50.694	17.593	50.406	197.407	0.253	2.552	420.995	1.037	1.043
E(tau_plus_ACbb1)	211.149	43.478	135.207	207.684	307.092	0.217	1.294	1158.042	1.015	1.036
E(tau_plus_ACwg1)	171.996	45.421	98.242	166.810	277.045	0.227	1.770	715.561	1.010	1.030
E(tau_plus_D1)	187.772	19.427	152.650	186.775	229.434	0.097	0.737	710.383	1.002	1.006
E(tau_plus_G1)	79.519	42.359	18.459	74.377	169.229	0.212	3.410	218.619	1.676	2.620
E(tau_plus_OB1)	47.399	23.554	17.290	42.266	106.229	0.118	0.879	717.875	1.007	1.020
E(delta_R0)	474.197	9.086	455.445	474.429	491.586	0.045	0.321	854.392	1.032	1.089

In the summary output, theta values reflect group-level parameter estimates transformed to a probability scale ranging from zero to one, such that the group-level means correspond to the probability that each process influences IAT responses. For example, the mean estimate for the Detection parameter (‘theta_D1) is about .91, which suggests that Detection plays a relatively large role in IAT responding. In contrast, the parameter estimates for White-good (‘theta_ACwg1’) and Black-bad (‘theta_ACbb1’) activated Associations are far lower (~.06), which suggests that they play a relatively small role in IAT responding. The output also contains information about the standard deviation of the posterior distributions for the parameters, the median and 95% Bayesian credible interval, the standard error, the standard error adjusted for autocorrelation,[8] the effective sample size, and the Rhat values.

In addition to process probabilities, the output includes process times for positive outcomes (tau_plus), negative outcomes (tau_minus), and encoding the IAT stimuli plus motor execution of the response (delta). In the context of the activated Associations, Detection, and Overcoming Bias, positive outcomes correspond to the success of each process, and negative outcomes correspond to the failure of each process. However, the Guessing parameter is coded such that the positive outcome is a response bias for ‘good’ and a negative outcome is a response bias for ‘bad.’ The average process time linked to each path can be calculated by summing up the process times for all parameters plus the encoding and motor execution time.

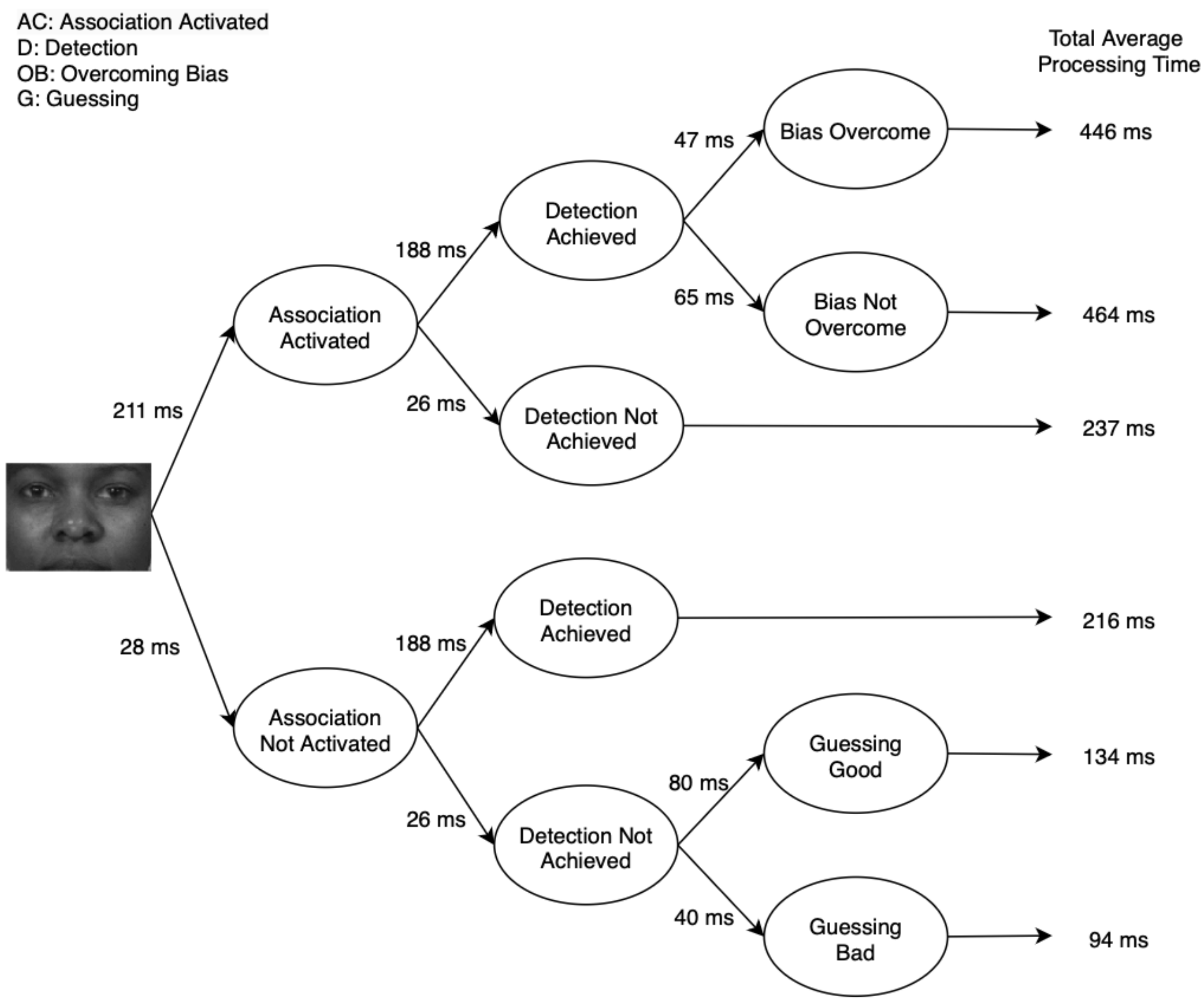
Testing Process Time Hypotheses

Researchers often have testable hypotheses about the speed of cognitive processes. One such hypothesis is rooted in a core tenet of dual-process theory – that automatic processing is faster than controlled processing – and can be readily tested using RT-MPT models. To explore this question, we began by computing the total average processing times for each path by summing the independent processing times (Figure 3). Descriptively, we see that the paths that include the successful influence of the control-oriented processes detection and overcoming bias are the slowest paths. Furthermore, the fastest branches are the bottom two, which reflect a general response bias driving IAT responding either in one direction or another following a failed activation of associations and a failed detection attempt.

[8] The standard errors refer to computational error only (i.e., how well the posterior mean was approximated by MCMC sampling).

Figure 3

Total Average Processing Time for Paths in RT-MPT Quad Model



Using the current example on a trial with a Black stimulus, we test for credible differences in processing times in the top path of the quad model (i.e., AC*D*OB) reflecting both the control-oriented processes of successful Detection and Overcoming Bias, compared to the bottom path of the quad model (i.e., [1-AC]*[1-D]*[1-G]) reflecting a response bias towards 'bad' in the absence of influence from the other processes. We predict that the top path should be credibly slower than the bottom path. Using the code below, we take the posterior samples of latencies corresponding to the bottom path and subtract them from the posterior samples of latencies corresponding to the top path to produce a distribution of differences. Then, we compute the 95% highest density interval (HDI) to determine if there are credible differences in processing speed between the two paths, which can be interpreted similarly to classic confidence intervals where an HDI containing zero indicates no credible differences.

```
diff <- coda::as.mcmc(times1 - times2)
coda::HPDinterval(obj = diff, prob = 0.95)
```

Indeed, we find that average process times for the top path (M = 446ms) are credibly slower than process times for the bottom path (M = 93ms; mean difference = 353ms, 95% HDI [235ms, 475ms]). Of note, the average process times for the bottom path are unexpectedly quite fast. However, given that all the branches of the quad model contain controlled processes (completing either successfully or failing), this approach does not provide a clear test of the question of whether automatic processes are faster than controlled ones for the current model.

Another approach for testing process time hypotheses involves comparing the mean processing time estimates and 95% BCIs for control-oriented processes vs. automatic processes. Neither the process completion times for the successful activation of White-good associations (M = 172ms, 95% BCI [98ms, 277ms]) or Black-bad associations (M = 211ms, 95% BCI [135ms, 307ms]) appear to be different from process completion time for the control-oriented process of Detection (M = 188ms, 95% BCI [153ms, 229ms]). Furthermore, the process completion time for the success of Overcoming Bias (M = 47ms, 95% BCI [17ms, 106ms]) does not overlap with the 95% BCI for the success of activated Black-bad associations and appears to be reliably faster. In contrast to the testing approach outlined in the above paragraph, this approach does not provide evidence that the control-oriented processes are slower than the automatic process in the RT-MPT version of the quad model. Note also that an average processing time of 47 ms for successful overcoming of bias (see Figure 2) seems relatively fast for a controlled process. Other models that include response times, such as diffusion models, also often attribute the bulk of the reaction time to motor processes and only attribute relatively short amounts of time to cognitive processing. Nevertheless, short process times can also indicate the parallel operation of two or more processes – in violation of the seriality assumption underlying RT-MPT models. For example, detection processes may overlap in time with attempts to overcome bias, such that the sum of the two process times should be interpreted to reflect their joint completion. Consequently, the completion time estimated for overcoming bias is an underestimate due to the overlap with the detection process.

MPT-RT (Heck & Erdfelder 2016 Approach)

MPT-RT models (Heck & Erdfelder, 2016) integrate response time data into the traditional MPT modeling framework without making assumptions about the shape of the observed RT distribution (Heck & Erdfelder, 2016). Rather than relying on distributional assumptions, MPT-RTs categorize each individual's RT distribution into multiple bins.

We prepared the raw IAT data ('iatData.csv') to be fit to an MPT-RT version of the quad model by relying on a mean split to categorize response times. For the purpose of exposition, we did not collect an additional set of data to be used for informing the RT boundaries of bins – but as a best practice, researchers should use data collected in an independent pilot study to identify category boundaries for assigning RTs to bins. The data that has been cleaned and prepared for model fitting can be found in the OSF repository ('mptrtData.csv'). We calculated the mean log-transformed response time for each participant and categorized each response as either fast or slow depending on whether it was lower or higher than the mean, respectively. This method produces two discrete RT bins of fast and slow responses for each correct and incorrect response across all IAT trial types for each participant. Though dichotomizing logarithmized response times according to a mean split results in a loss of information compared to analyzing continuous response times, this strategy speeds up parameter estimation, creates comparable RT bins across participants, and reduces the possibility of zero-count bins that may produce misleading results. The following code splits each participants' response times into fast and slow bins based on a mean split.

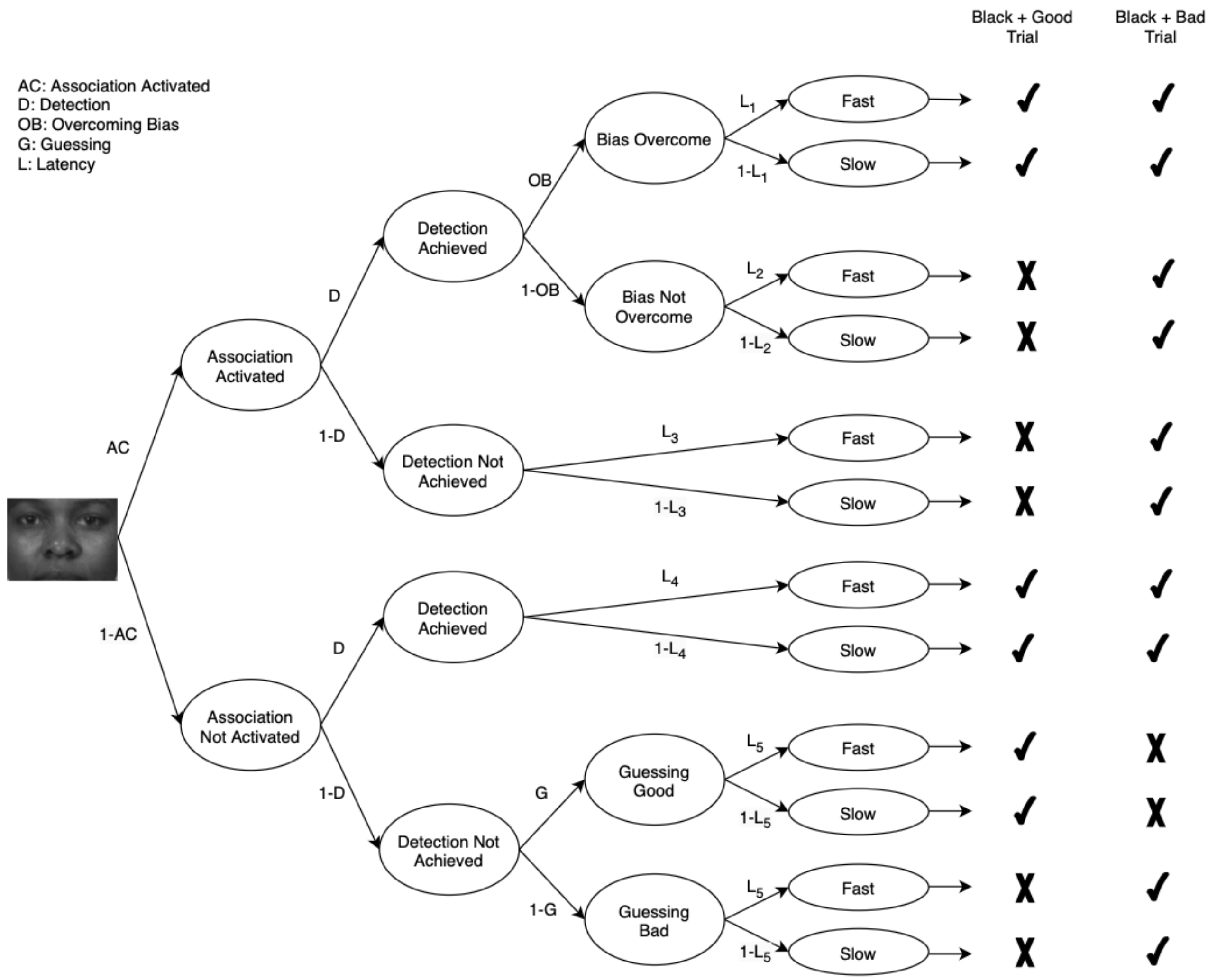
```
df.rt.bin <- df %>%
  dplyr::group_by(participant.id, .drop = FALSE) %>%
  dplyr::summarise(rt.logmean = exp(mean(log(rt), na.rm =
T)))

df.rt.bin <- df.rt.bin %>%
  dplyr::mutate(rt.bin = case_when(rt.logmean <= rt ~ 0,
                                   rt.logmean > rt ~ 1))
```

Figure 4 illustrates the cognitive architecture of this MPT-RT version of the quad model in a processing tree. At the end of each processing path, an additional parameter *L* (for latency) reflects the probability of the path terminating in a fast response with probability *L* and in a slow response with the complementary probability (1-*L*). In the context of the MPT-RT version of the quad model, we specify six separate latency parameters for different types of processing pathways. L1 is the latency parameter for the Associations, Detection, and Overcoming Bias pathway, L2 is the latency parameter for the Associations, Detection, and failure to Overcome Bias pathway, L3 is the latency parameter for the Association pathway in the absence of Detection, and L4 is the latency parameter for the pathway involving Detection in the absence of Associations. We specify two separate L4 parameters, one for face stimuli (L4face), and another for words (L4word). Last, L5 is the latency parameter for the path involving Guessing good or bad in the absence of influence from other processes.

Figure 4

A Portion of the MPT-RT model version of the Quad Model



Note. The table on the right illustrates correct (✓) and incorrect (X) responses across different trial types. Parameters with lines leading to them are conditional upon the preceding parameters. We estimate two different L_4 parameters – one for face stimuli and another for word stimuli – which are not reflected in this simplified model depiction.

Model Fitting

We outline here a modeling approach that relies on Bayesian estimation for hierarchical latent-trait MPT models (Klauer, 2010). This approach is implemented in the '*TreeBUGS'* R package (Heck et al., 2018) and draws posterior samples of the parameters using Markov chain Monte Carlo (MCMC) methods. The code below outlines the function arguments.

```
mptrtModel <- traitMPT(eqnfile="mptrtQuad.eqn",
                 data = mptrtData,
                 n.adapt=10000,
                 n.iter=100000,
                 n.burnin=5000,
                 n.thin=5,
                 n.chains=4)
```

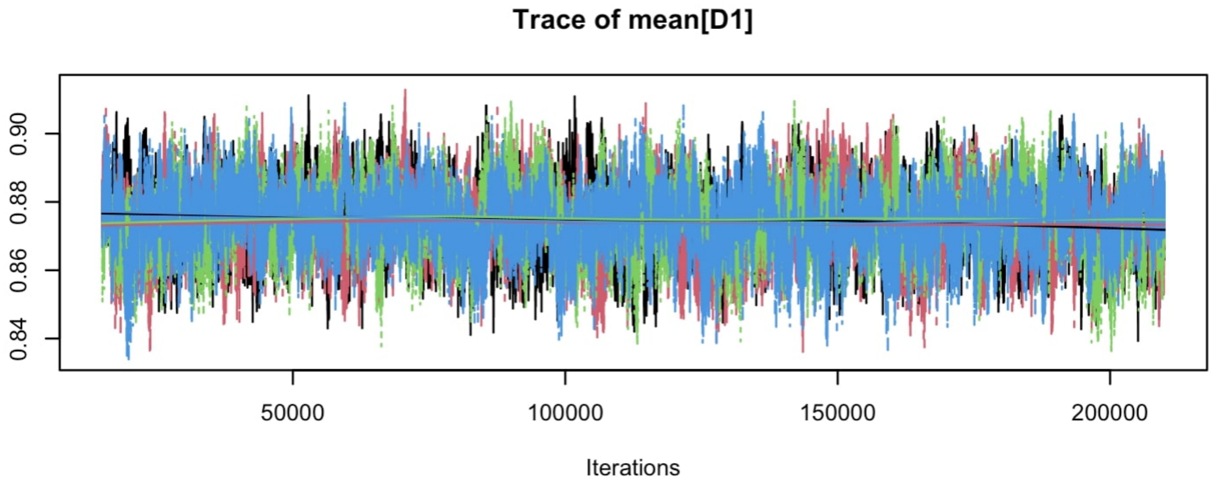
Like the RT-MPT approach described above, the estimation described here also relies on hierarchical Bayesian inference and uses similar arguments for the model fitting function: `n.chains`, `n.iter`, `n.burnin`, and `n.thin`. However, this approach differs from the RT-MPT modeling approach because it does not include a Rhat_max argument. Rather, sampling will continue for however many iterations are specified in the `n.iter` argument. After the estimation has completed, researchers can evaluate the Rhat convergence statistic in the summary output of the fitted model.

Additionally, researchers can visually check for model convergence by inspecting trace plots. Figure 5 illustrates the trace plot for the Detection parameter, which shows the values that Detection took during the sampling of the chain, with different colors representing different chains. Ideally, the trace plot should resemble a “fat, hairy caterpillar” with large amounts of overlap among the chains, which indicates that the sampler has adequately explored the full range of the posterior distribution. We generated this plot using the code below.

```
plot(mptrtModel, parameter = "mean")
```

Figure 5

Trace Plot for Quad Detection Parameter



If either the Rhat convergence statistic or the visual check for convergence suggests that the model did not sufficiently converge, then researchers can begin a new estimation with an increased burnin period and/or number of iterations. However, given that MPT estimation can often be time intensive, researchers can instead use the `extendMPT()` function to add more MCMC samples to the fitted model while retaining all of the previously sampled posterior values.

```
summary(mptrtModel)

mptrtModelExtended <- extendMPT(mptrtModel, n.iter = 50000,

n.adapt = 10000, n.thin =5)
```

Interpreting Model Output

As with the RT-MPT model, researchers can use the T_1 statistic (Klauer, 2010) to evaluate model fit. Additionally, researchers using MPT-RT models can use the T_2 statistic, which summarizes how well the model accounts for the variances and covariances of the RT-extended frequencies computed across participants and quantifies how well the model accounts for individual differences between participants' RT-extended response frequencies (Klauer, 2010). Similar to the RT-MPT models, we recommend reporting the W statistic because the group-level T_1 statistic is highly powered to detect even a small degree of misfit. Relative to the group-level tests, individual-level tests necessarily have less power to detect misfit. Consequently, we also recommend exploring the median individual-level p-value for T_1. In our example, the group-level T_1 observed and predicted values were 4.21 and .17, respectively, $p <$ .001. However, at the group level, the model shows adequate fit after adjusting for the number of observations (W = 0.016). In addition, the median individual T_1 p-value was .20. The code below produces group-level T_1, group-level T_2, and individual-level T_1.

```
PPP(mptrtModel)
```

Figure 6 shows the summary output for MPT-RT models, which contains information about the mean process probabilities and 95% Bayesian credible intervals, the standard error adjusted for autocorrelation, effective sample size, and Rhat values. The basic parameter estimates in the summary output for MPT-RT models can be interpreted very similarly to the RT-MPT models described above, such that the mean parameter estimates for MPT-RT models quantify the probability that each quad process influences IAT responses. For example, the mean estimate for the Detection parameter ('mean_D1') is about .88, suggesting that Detection plays a large role in IAT responding. Importantly, the probability parameter estimates produced from this MPT-RT implementation of the quad model may differ somewhat from the probability parameter estimates from the RT-MPT implementation of the quad model due to the substantive differences in the assumed prior distributions and other underlying model assumptions for each modeling method. That said, they should ideally match relatively closely; indeed, in the present analyses the BCIs of parameters estimated by both methods overlap. However, major discrepancies between parameters estimated by the two methods may indicate model misfit.

Figure 6

R Output of MPT-RT Group-level Parameter Estimates

```
Group-level medians of MPT parameters (probability scale):
              Mean    SD  2.5%   50% 97.5% Time-series SE n.eff  Rhat R_95%
mean_ACbb1   0.019 0.008 0.005 0.019 0.036          0.000  3169 1.001 1.003
mean_ACwg1   0.037 0.010 0.018 0.037 0.057          0.000  3193 1.003 1.008
mean_D1      0.875 0.010 0.855 0.875 0.893          0.000  1492 1.002 1.002
mean_G1      0.569 0.019 0.532 0.568 0.608          0.000  3775 1.004 1.010
mean_L1      0.026 0.050 0.000 0.013 0.122          0.001  5298 1.021 1.027
mean_L2      0.381 0.281 0.009 0.332 0.943          0.004  5155 1.001 1.004
mean_L3      0.171 0.083 0.032 0.164 0.351          0.002  1355 1.009 1.026
mean_L4face  0.835 0.012 0.812 0.835 0.858          0.000  2146 1.001 1.003
mean_L4word  0.616 0.010 0.596 0.616 0.637          0.000  3164 1.002 1.004
mean_L5      0.032 0.010 0.014 0.032 0.054          0.000  3802 1.001 1.002
mean_OB1     0.974 0.038 0.881 0.986 1.000          0.001  3973 1.044 1.052
```

Testing Process Time Hypotheses

Unlike RT-MPT models, MPT-RT models do not provide estimates of process times but instead provide insight into the relative speed of the processing paths. By specifying equality constraints and order restrictions between latency parameters, researchers can test hypotheses about the relative speed of different processing paths. For example, the core tenet of dual-process theory that automatic processing is faster than controlled processing can be readily tested in the context of the MPT-RT version of the quad model, which specifies six separate latency parameters for different types of processing pathways. Some processing pathways, such as L3, solely reflect parameters assumed to capture automatically-activated Associations (though the assumption that the failure of Detection has no time cost is an empirical question). Other processing pathways, such as L1, reflect parameters assumed to capture control-oriented Detection and Overcoming Bias. Based on dual-process theory, we should expect processing routes characterized by automatic processing without the successful influence of control-oriented processing to be faster than processing routes characterized by successful control-oriented processing.

To test this hypothesis using the MPT-RT framework, we fit a model with constraints on two latency parameters, such that $L_3 \geq L_1$, which reflects the assumption that the automatic processing pathway is never slower than the control-oriented processing pathway. We also fit the basic MPT-RT model with no constraints on the relative speed of different pathways. To test our assumption about automatic- versus control-oriented processing speeds, we will compare the models using two Bayesian-oriented model selection indices: the deviance information criterion (DIC: Spiegelhalter et al., 2002) and the widely applicable information criterion (WAIC: Watanabe, 2010). Both of these model selection indices quantify model fit while penalizing for model complexity. Like the better-known model selection indices AIC and BIC, the model with the lowest DIC or WAIC value represents the preferred model. Differences of 2 are interpreted to reflect weak evidence for one model over the other and differences greater than 10 are interpreted to reflect strong evidence for one model over the other (Spiegelhalter et al., 2002). If the model selection indices suggest that the quad MPT-RT with constraints better accounts for observed IAT responses than the basic quad MPT-RT model, then we can conclude that automatic processing happens faster than control-oriented processing. The code below computes DIC and WAIC for the fitted model.

```
mptrtModel$summary$dic <-
runjags::extract(mptrtModel$runjags,"dic")

mptrtModel$summary$waic <- WAIC(mptrtModel)
```

We observe adequate model fit for both the MPT-RT model with constraints ($W = 0.016$) and the basic MPT-RT model ($W = 0.016$). Model selection indices for the MPT-RT model with constraints (DIC = 9272; WAIC = 9392) versus the basic MPT-RT model (DIC = 9274; WAIC = 9397) offer moderate support for the constrained model (ΔDIC = 2, ΔWAIC=5).

Model-selection analyses evaluate whether a model with constraints imposed on each individual's data provides a better account than the unconstrained model. We can also test whether such a constraint is satisfied in terms of group-level means estimated by the unconstrained model. For this purpose, we computed the proportion of samples in which the group-level estimate of L_3 is smaller than L_1. The resulting Bayesian *p*-value from this test may be interpreted as support for our hypothesis of $L_3 \geq L_1$ at the level of group means if the *p*-value is small. The code below computes this proportion and summarizes the results.

```
tp <- transformedParameters(mptrtModel, transformedParameters =
                                  list("diff= L3 < L1"))

summary(tp)
```

Indeed, this test returns a Bayesian p-value of .036, offering additional support for our hypothesis that the automatic processing pathway is faster than the control-oriented processing pathway.

Testing Correlations Between MPT-RT Parameters and Covariates

Researchers may want to explore correlations between model parameter estimates and external covariates, such as personality traits, attitudes, test scores, or demographic variables. The TreeBUGS package can perform such tests by including covariates in the model fitting function. Covariate testing is not an advantage of the MPT-RT modeling approach over the RT-MPT modeling approach, per se, but instead is a feature that is available in the TreeBUGS package but not the rtmpt package. Here we explore the sample correlation of quad MPT-RT parameters with self-reported attitudes towards Black people, measured with a feeling thermometer ranging from 1 to 10. The code below shows how we adjust arguments in the fitting function to carry out this analysis.

```
mptrtCovariate <- traitMPT(eqnfile="mptrtQuad.eqn",
                  data = rtmptData,
                  n.adapt=10000,
                  n.iter=200000,
                  n.burnin=5000,
                  n.thin=5,
                  n.chains=4,
                                  covData = subset(covariate,
                  select = "att_Black"),
                                  predStructure = list("ACwg1
                  ACbb1 D1 G1 OB1; att_Black"),
                  predType = list("c"))
```

The argument `covData = subset(covariate, select = "att_Black")` provides a data frame containing the self-reported attitude scores for each participant. Importantly, the order of participants in the covariate data frame must be identical to the order of participants in the IAT frequency data provided in the mptrtData.csv file. Additionally, we include the name of the covariate we wish to use to predict specific MPT parameters in the argument

```
predStructure = list("ACwg1 ACbb1 D1 G1 OB1; att_Black")
```

In this argument, we separate the list of MPT parameters of interest from the covariate with a semicolon. After fitting the model to the data, the summary of the results can be reviewed using this function:

```
summary(mptrtCovariate)
```

From our summary output, we find that neither participant's White-good Associations (β = -.14, 95% BCI [-.37, .10]) nor Black-bad Associations (β = -.07, 95% BCI [-.21, .35]) are correlated with their self-reported attitudes towards Black people.

Conclusion and Limitations

Response-time extended MPT models are still relatively novel in social cognition, but researchers have already begun to apply these methods to study a wide variety of topics, including weapon bias (Laukenmann et al., 2023), heuristic decision making (Heck & Erdfelder, 2017), memory (Brainerd et al., 2019; Greene & Naveh-Benjamin, 2024; Gutkin et al., 2024), speed-accuracy tradeoff (Heck & Erdfelder, 2020; Starns, 2018), and selective attention (Li & Deng, 2024). We hope our tutorial contributes to this growing body of work and provides researchers with a toolbox for implementing response-time extended MPT models within their own studies. Please note that this tutorial should not be taken as empirical evidence of the validity of the RT-extended versions of the quad model, so please refrain from drawing inferences based on model results we reported here. For example, our quad model implementations via RT-MPT and MPT-RT modeling methods differ in their estimated order of fast to slow process paths. Such divergence may arise due to differences in response time data curation and/or different assumptions for response time parameters (e.g., RT distribution assumptions, latency parameter restrictions). These and other issues merit further empirical investigation. In addition to adequate model fit, an RT-extended MPT model must undergo a program of validation studies intended to provide support for the psychological interpretation of each parameter.[9] This tutorial does not provide evidence for the validity of either of the response time-extended models we describe here, but instead highlights how both modeling approaches hold great promise for investigating questions related to the nature and interplay of the cognitive processes underlying IAT responses. Ultimately, RT-extended MPT models can position

[9] MPT models – traditional and response-time extended – are validated through a series of selective influence studies in which experimental manipulations are introduced that are intended to facilitate or constrain the specific cognitive processes specified in the MPT model. For further reading on MPT validation, we recommend Klauer (2024) and Hütter and Klauer (2016).

researchers to test novel hypotheses and provide greater theoretical precision than other formal modeling approaches.

That said, these approaches are not without their limitations. MPT-RT models, and especially RT-MPT models, are computationally resource-intensive, and sufficient convergence for IAT data often requires lengthy estimation times and a computer with adequate working memory. In the present example, fitting the MPT-RT version of the quad model to the IAT data took up to 30 minutes per estimation. Fitting the RT-MPT version of the quad model took longer and varied more substantially, with estimations ranging from 12 hours to multiple days. Due to the length of the estimation time, the feasibility of this approach is limited when modeling large samples of data, fitting complex models, and modeling without sufficient computing resources.

Given the practical obstacles and complexity of RT-extended MPT models, we do not recommend that they be used in all circumstances. If a research objective can be achieved with a traditional MPT model, then researchers can avoid the drawbacks of response-time extended approaches, such as the distributional assumptions or parameter constraints.

One final limitation is that, anecdotally, our colleagues often report feeling intimidated or confused by seemingly-arcane MPT analytic methods – response time-extended or otherwise. We wrote this user-friendly tutorial specifically to address this limitation. By combining step-by-step instructions with free open-source software, we hope that more researchers will use response-time extended MPT models for the IAT or other social cognitive tasks in their own work.

References

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*(4), 548–564. https://doi.org/10.1037/0033-295X.97.4.548

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, *6*(1), 57–86. https://doi.org/10.3758/BF03210812

Brainerd, C. J., Nakamura, K., Chang, M., & Bialer, D. M. (2019). Verbatim editing: A general model of recollection rejection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1776–1790. https://doi.org/10.1037/xlm0000684

Calanchini, J. (2020). How multinomial processing trees have advanced, and can continue to advance, research using implicit measures. *Social Cognition*, *38*(Supplement), s165–s186. https://doi.org/10.1521/soco.2020.38.supp.s165

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality and Social Psychology*, *89*(4), 469–487. https://doi.org/10.1037/0022-3514.89.4.469

Elder, J., Wilson, L., & Calanchini, J. (2023). Estimating the reliability and stability of cognitive processes contributing to responses on the Implicit Association Test. *Personality and Social Psychology Bulletin*, *50*(10), 1451–1470. https://doi.org/10.1177/01461672231171256

Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 387–414. https://doi.org/10.1037/0278-7393.24.2.387

Gawronski, B., Armstrong, J., Conway, P., Friesdorf, R., & Hütter, M. (2017). Consequences, norms, and generalized inaction in moral dilemmas: The CNI model of moral decision-making. *Journal of Personality and Social Psychology*, *113*(3), 343–376. https://doi.org/10.1037/pspa0000086

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*(2), 410–414. https://doi.org/10.1016/j.jesp.2008.11.004

Greene, N. R., & Naveh-Benjamin, M. (2024). The time course of encoding specific and gist episodic memory representations among young and older adults. *Journal of Experimental Psychology: General*, *153*(6), 1671–1697. https://doi.org/10.1037/xge0001589

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Gutkin, A., Suero, M., Botella, J., & Juola, J. F. (2024). Benefits of multinomial processing tree models with discrete and continuous variables in memory research: An alternative modeling proposal to Juola et al. (2019). *Memory & Cognition*, *52*(4), 793–825. https://doi.org/10.3758/s13421-023-01501-8

Hartmann, R., Johannsen, L., & Klauer, K. C. (2020). rtmpt: An R package for fitting response-time extended multinomial processing tree models. *Behavior Research Methods*, *52*(3), 1313–1338. https://doi.org/10.3758/s13428-019-01318-x

Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods*, *50*(1), 264–284. https://doi.org/10.3758/s13428-017-0869-7

Heck, D. W., & Erdfelder, E. (2016). Extending multinomial processing tree models to measure the relative speed of cognitive processes. *Psychonomic Bulletin & Review*, *23*(5), 1440–1465. https://doi.org/10.3758/s13423-016-1025-6

Heck, D. W., & Erdfelder, E. (2017). Linking process and measurement models of recognition-based decisions. *Psychological Review*, *124*(4), 442–471. https://doi.org/10.1037/rev0000063

Heck, D. W., & Erdfelder, E. (2020). Benefits of response time-extended multinomial processing tree models: A reply to Starns (2018). *Psychonomic Bulletin & Review*, *27*(3), 571–580. https://doi.org/10.3758/s13423-019-01663-0

Heck, D. W., Erdfelder, E., & Kieslich, P. J. (2018). Generalized processing tree models: Jointly modeling discrete and continuous variables. *Psychometrika*, *83*(4), 893–918. https://doi.org/10.1007/s11336-018-9622-0

Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology*, *27*(1), 116–159. https://doi.org/10.1080/10463283.2016.1212966

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, *134*(2), 131–148. https://doi.org/10.1037/0096-3445.134.2.131

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*(1), 70–98. https://doi.org/10.1007/s11336-009-9141-0

Klauer, K. C., Hartmann, R., & Meyer-Grant, C. G. (2024). RT-MPTs: Process models for response-time distributions with diffusion-model kernels. *Journal of Mathematical Psychology*, *120–121*, 102857. https://doi.org/10.1016/j.jmp.2024.102857

Klauer, K. C., & Kellen, D. (2018). RT-MPTs: Process models for response-time distributions based on multinomial processing trees with applications to recognition memory. *Journal of Mathematical Psychology*, *82*, 111–130. https://doi.org/10.1016/j.jmp.2017.12.003

Klauer, K. C., & Meiser, T. (2000). A source-monitoring analysis of illusory correlations. *Personality and Social Psychology Bulletin*, *26*(9), 1074–1093. https://doi.org/10.1177/01461672002611005

Klauer, K. C., & Voss, A. (2008). Effects of race on responses and response latencies in the Weapon Identification Task: A test of six models. *Personality and Social Psychology Bulletin*, *34*(8), 1124–1140. https://doi.org/10.1177/0146167208318603

Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *Journal of Personality and Social Psychology*, *103*(2), 205–224. https://doi.org/10.1037/a0028764

Kruschke, J. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (2nd edition). Academic Press.

Laukenmann, R., Erdfelder, E., Heck, D. W., & Moshagen, M. (2023). Cognitive processes underlying the Weapon Identification Task: A comparison of models accounting for both response frequencies and response times. *Social Cognition*, *41*(2), 137–164. https://doi.org/10.1521/soco.2023.41.2.137

Li, J., & Deng, S. W. (2024). Attentional focusing and filtering in multisensory categorization. *Psychonomic Bulletin & Review*, *31*(2), 708–720. https://doi.org/10.3758/s13423-023-02370-7

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: The ReAL model for the IAT. *Journal of Personality and Social Psychology*, *104*(1), 45–69. https://doi.org/10.1037/a0030734

Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54. https://doi.org/10.3758/BRM.42.1.42

Nadarevic, L., & Erdfelder, E. (2011). Cognitive processes in implicit attitude tasks: An experimental validation of the Trip Model. *European Journal of Social Psychology*, *41*(2), 254–268. https://doi.org/10.1002/ejsp.776

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*(1), 101–115. https://doi.org/10.1037/1089-2699.6.1.101

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*(2), 181–192. https://doi.org/10.1037/0022-3514.81.2.181

Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). A process model of affect misattribution. *Personality and Social Psychology Bulletin*, *36*(10), 1397–1408. https://doi.org/10.1177/0146167210383440

Schmittmann, V. D., Dolan, C. V., Raijmakers, M. E. J., & Batchelder, W. H. (2010). Parameter identification in multinomial processing tree models. *Behavior Research Methods*, *42*(3), 836–846. https://doi.org/10.3758/BRM.42.3.836

Singmann, H., & Kellen, D. (2013). MPTinR: Analysis of multinomial processing tree models in R. *Behavior Research Methods*, *45*(2), 560–575. https://doi.org/10.3758/s13428-012-0259-0

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence. *Experimental Psychology*, *54*(2), 99–112. https://doi.org/10.1027/1618-3169.54.2.99

Stahl, C., & Klauer, K. C. (2007). HMMTree: A computer program for latent-class hierarchical multinomial processing tree models. *Behavior Research Methods*, *39*(2), 267–273. https://doi.org/10.3758/BF03193157

Starns, J. J. (2018). Adding a speed–accuracy trade-off to discrete-state models: A comment on Heck and Erdfelder (2016). *Psychonomic Bulletin & Review*, *25*(6), 2406–2416. https://doi.org/10.3758/s13423-018-1456-3

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General*, *123*(1), 34–80. https://doi.org/10.1037/0096-3445.123.1.34

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*(12), 3571–3594.