

COGNITIVE PROCESSES UNDERLYING THE WEAPON IDENTIFICATION TASK: A COMPARISON OF MODELS ACCOUNTING FOR BOTH RESPONSE FREQUENCIES AND RESPONSE TIMES

Ruben Laukenmann and Edgar Erdfelder
University of Mannheim

Daniel W. Heck
University of Marburg

Morten Moshagen
Ulm University

The weapon identification task (WIT) is a sequential priming paradigm designed to assess effects of racial priming on visual discrimination between weapons (guns) and innocuous objects (tools). We compare four process models that differ in their assumptions on the nature and interplay of cognitive processes underlying prime-related weapon-bias effects in the WIT. All four models are variants of the process dissociation procedure, a widely used measurement model to disentangle effects of controlled and automatic processes. We formalized these models as response time-extended multinomial processing tree models and applied them to eight data sets. Overall, the default interventionist model (DIM) and the preemptive conflict-resolution model (PCRM) provided good model fit. Both assume fast automatic and slow controlled process routes. Additional comparisons

Supplementary material is available at <https://osf.io/7vjrq>

Manuscript preparation was supported by the Research Training Group “Statistical Modeling in Psychology” (SMiP), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, GRK 2277), and by a Doctoral Research Fellowship from the Stiftung der Deutschen Wirtschaft gGmbH (sdw, Foundation of German Business), funded by the German Federal Ministry of Education and Research (awarded to RL).

The authors are very grateful to M. J. Amon, A. J. Bishara, J. Correll, A. J. Lambert, C. Madurski, and A. M. Rivers for sharing the raw data we used in our model comparisons. We also thank Karl C. Klauer and an anonymous reviewer for helpful comments on an earlier version of this manuscript.

The authors have no conflicts of interest to disclose.

This article was accepted under the editorship of Jeffrey W. Sherman.

Address correspondence to: Ruben Laukenmann or Edgar Erdfelder, Cognition and Individual Differences Lab, A5, B 118, University of Mannheim, D-68159 Mannheim, Germany. E-mail: ruben.laukenmann@psychologie.uni-mannheim.de or erdfelder@uni-mannheim.de

avored the former model. In line with the DIM, we thus conclude that automatically evoked stereotype associations interfere with correct object identification from the outset of each WIT trial.

Keywords: weapon identification task, process dissociation procedure, dual process models, multinomial processing tree modeling, racial bias

Motivated by various incidents of police shootings of unarmed Black men, several studies have investigated the visual discrimination between weapons and perceptually similar innocuous objects as a function of person characteristics such as race (Payne, 2001; Rivers, 2017). To investigate the influence of racial stereotypes on object identification, Payne (2001) proposed the weapon identification task (WIT). The WIT is a sequential priming paradigm. Participants are instructed to identify visually presented objects as either weapons (e.g., guns) or innocuous objects (e.g., tools). In the standard version of the WIT, target objects are preceded by either a Black or a White male face (Rivers, 2017). Many studies found that innocuous objects are more often erroneously misidentified as weapons following a brief presentation of a Black male face compared to a brief presentation of a White male face. In addition, participants typically identify a weapon faster after seeing a Black male face prime and an innocuous object faster after seeing a White male face prime. This result pattern, as reflected in error rates and response times, is known as *weapon identification bias* (Payne, 2001; Rivers, 2017; Todd et al., 2021). According to the meta-analysis of Rivers (2017), the weapon identification bias has a large effect size for error rates ($\eta^2 = .204$) and a medium effect size for response times ($\eta^2 = .106$).

To investigate the cognitive processes underlying the weapon identification bias, Payne (2001) applied the process dissociation procedure (PDP), a measurement model that decomposes task performance into effects of controlled processes and automatic processes. Whereas the controlled process represents the ability to identify the target object correctly, the automatic process reflects a response bias towards weapons versus innocuous objects. The typical finding is that prime race does not affect the controlled process but it does affect the automatic process, with Black male face primes leading to a larger response bias towards weapons than White male face primes (e.g., Huntsinger et al., 2009; Ito et al., 2015; Payne, 2001; Payne et al., 2002).

Although the PDP model allows disentangling and measuring the contributions of automatic and controlled influences on performance, it is agnostic with respect to the exact specification and temporal order of these cognitive processes. In fact, as outlined by Klauer and Voss (2008), the PDP model is consistent with four different cognitive process models for the WIT. Identifying the empirically most adequate of these process models is important as different models suggest different intervention strategies to reduce weapon identification bias. Knowledge about the underlying processes enables a deeper understanding of the weapon identification bias and thus provides a better foundation for racial bias awareness programs

and for racial bias reduction interventions (e.g., Kahn & Martin, 2020; Klauer & Voss, 2008; Swencionis & Goff, 2017).

In this article, we first introduce the PDP and revisit the four psychological process models based on the PDP. We then formalize these process models in the framework of response time-extended multinomial processing tree (MPT-RT) models (Heck & Erdfelder, 2016) and compare them across eight different data sets of the weapon identification task. Including response latencies in the PDP by MPT-RT modeling allows us to compare the four models empirically and to identify the best-fitting process model. Hence, the current research goes beyond previous qualitative comparisons based on observed response times and accuracy patterns (Klauer & Voss, 2008) by providing quantitative assessments of the latent processing mechanisms underlying empirical data. Such an assessment is not possible when response frequencies are considered in isolation, which has been routine practice in PDP analyses of WIT performance so far. Finally, we discuss implications as well as remaining limitations of our model comparison.

PROCESS DISSOCIATION PROCEDURE

To measure the contributions of automatic processes and controlled processes to WIT performance, Payne (2001) proposed the process dissociation procedure. Since then, the PDP model has become a widely used dual-process model for analyzing performance in the WIT (e.g., Bishara & Payne, 2009; Huntsinger et al., 2009; Ito et al., 2015; Klauer et al., 2015; Payne, 2005). According to this model, the controlled process represents the latent ability to identify the target and thus always leads to a correct response if it succeeds. The controlled process is thought to be effortful and constrained by available cognitive resources (Klauer & Voss, 2008; Payne, 2001). In contrast, the automatic process reflects a latent response bias towards weapons rather than innocuous objects. This automatic bias is typically activated by a prime stimulus and driven by preexisting cognitive structures (e.g., stereotypical associations induced by a preceding face prime). The automatic process is assumed to be effortless and spontaneous (Conrey et al., 2005; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Based on the participants' task performance, the PDP model allows researchers to estimate the probabilities that responses are determined by controlled or automatic processes. Parameter C and its complement $(1 - C)$ denote the probabilities that the controlled process succeeds or fails, respectively. Analogously, parameters A and $(1 - A)$ represent the probabilities that automatic bias drives the response towards *gun* or *tool*, respectively.

Figure 1 illustrates the PDP measurement model in the form of two processing tree diagrams, one for each target condition (i.e., gun and tool). The left side indicates the to-be-identified target object, followed by the possible processing branches (and their associated probabilities) that lead to certain responses in the WIT, as indicated on the right side of the trees. For the gun target tree (we refer to the weapon target condition as gun target condition), for example, the controlled process succeeds with probability C , resulting in the correct response *gun*. When the controlled process fails to determine the response with probability $(1 - C)$,

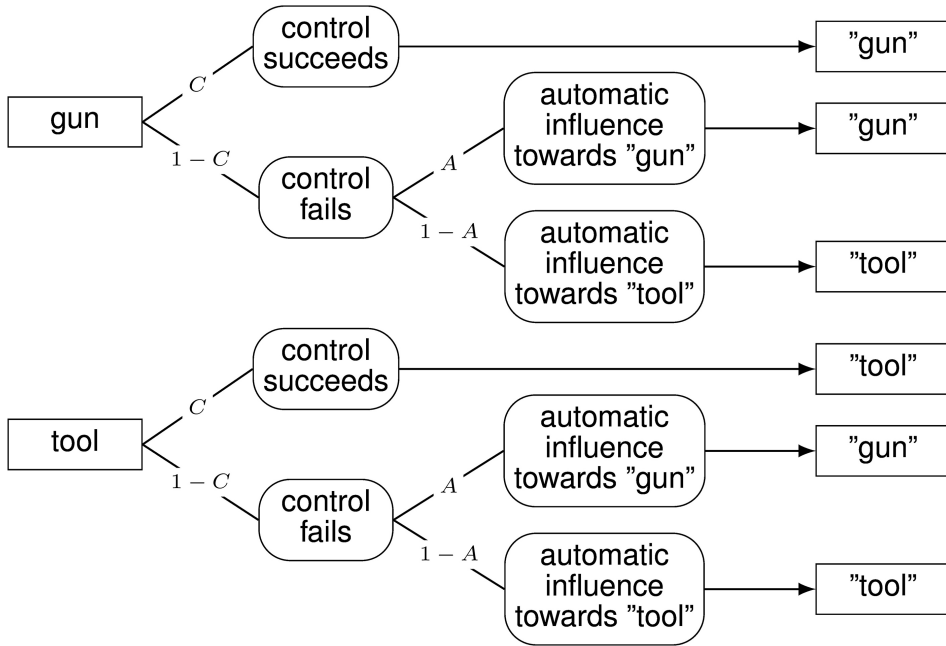


FIGURE 1. Generic PDP model for the weapon identification task. Parameters C and A denote probabilities of response determination by a controlled process and an automatic process, respectively. Note that A is conditional on a failure of the controlled process, that is, A represents the conditional probability of response determination by an automatic process given controlled process failure. Each of the parameters C and A may depend on whether the preceding prime is a Black or a White male face.

the automatic process may trigger the response *gun* with probability A and the response *tool* with the complementary probability $(1 - A)$. For the tool target tree, the processing branches are analogous, except that the controlled process results in the correct response *tool*.

Parameters of processing tree models such as those illustrated in Figure 1 can be estimated using multinomial processing tree (MPT) modeling (for a tutorial, see Schmidt et al., 2023). To enable parameter estimation in MPT models, expected response probabilities are first derived by calculating the product of the probability parameters along each branch of a tree and then summing up those branch probabilities that refer to the same response. For example, a correct response in a gun target trial can be due to the controlled process being successful (with probability C) or due to an automatic response bias towards *gun* when the controlled process fails (with probability $(1 - C) \cdot A$). The probability of a correct response thus equals the sum of these two probabilities:

$$p(\text{response gun} \mid \text{target object gun}) = C + (1 - C) \cdot A$$

Based on a system of such model equations for all response categories in each condition, parameters are then estimated by minimizing an appropriate distance

measure between observed and expected response frequencies (e.g., the log-likelihood ratio statistic G^2). MPT models have previously been used to estimate PDP parameters for the WIT (Bishara & Payne, 2009; Klauer et al., 2015). Compared to alternative estimation methods based on simple algebraic equations (Payne, 2001), MPT modeling enables not only conjoint estimation of all relevant PDP parameters along with their standard errors in one or more experimental conditions, but also, more importantly, the assessment of model fit whenever the number of parameters is less than the number of independent response categories.

NATURE OF COGNITIVE PROCESSES

Despite the clear-cut definition of the parameters C and A as probabilities of latent cognitive processes underlying responses in the WIT, their psychological nature has been a matter of debate (e.g., Bishara & Payne, 2009; Klauer & Voss, 2008; Klauer et al., 2015; Payne, 2005). Conrey et al. (2005) considered four potential cognitive determinants of performance in the WIT (see also Evans, 2007; Klauer & Voss, 2008; Sherman, 2006):

- The ability to discriminate between weapons and innocuous objects
- Activation of racial associations (e.g., Black males are associated with guns)
- The ability to resolve conflicts between racial associations and target identification (the so-called overcoming bias)
- The guessing tendency towards one of the response options when none of the other processes determines the response

These four processes closely relate to controlled and automatic processes in the PDP. Specifically, both the ability to discriminate between stimuli and the ability to overcome bias reflect controlled processes, whereas activation of racial associations and guessing tendencies reflect automatic processes (Klauer & Voss, 2008). For the PDP this implies that parameter C can reflect target discrimination, overcoming bias, or both, whereas parameter A may reflect activation of racial associations, guessing tendencies, or both. In addition, given that controlled and automatic processes may operate sequentially or in parallel, it is worthwhile to define more precisely how these processes interact and, if so, how they are temporally ordered.

As the PDP is silent about the psychological nature of these processes and their temporal order and duration, several psychological process models based on the PDP are conceivable, each of which relies on different assumptions about the nature and the interplay of automatic and controlled processes.

PROCESS MODELS FOR THE WIT

Based on Evans (2007), Klauer and Voss (2008) discussed four different psychological process models for the WIT that directly relate to the PDP. The preemptive conflict-resolution model (PCRM) assumes a preemptive decision whether the controlled process or the automatic process will determine the response. In

comparison, the default interventionist model (DIM) presupposes that a default automatic response is always activated first, which may or may not be successfully overcome by the controlled process subsequently. According to the parallel competitive model (PCM), in contrast, automatic and controlled processes run in parallel, sometimes requiring a subsequent conflict resolution process when the proposed responses disagree. Finally, the guessing model (GM) assumes that the controlled process determines the response, but if it fails to do so, the response is based on guessing.

As detailed below, all four psychological process models are based on different assumptions about the nature and the chronological order of the underlying cognitive processes. This notwithstanding, they are all instantiations of the PDP that cannot be distinguished based on the accuracy of responses across the four prime-target conditions alone. In other words, these process models are empirically indistinguishable when formalized as simple PDP models for response frequencies.

However, these models come along with different assumptions about the relative latencies of the processing branches of the PDP (Klauer & Voss, 2008) and can thus be tested against each other when predictions concerning relative latencies are considered in addition. To investigate the relative latencies of process branches in MPT models, Heck and Erdfelder (2016) proposed response time-extended multinomial processing tree (MPT-RT) models. In the MPT-RT framework, responses are subclassified by their response time, for example, by a dichotomous split into fast versus slow responses. To transform an MPT model into an MPT-RT model, an additional probability parameter L_j is simply added to each branch of the MPT model such that branch j terminates in a fast response with probability L_j and in a slow response with the complementary probability $(1 - L_j)$.¹ Importantly, MPT-RT models allow including equality constraints and order restrictions between latency parameters to test different hypotheses on the relative speed of processing branches.

In the following, we present the four process models that Klauer and Voss (2008) discussed for the WIT in more detail and show how they can be formalized as MPT-RT models. To foreshadow, the four process models are formalized as specific versions of the PDP that rely on different parametrizations (with A conditional on C for PCRM and GM, and C conditional on A for DIM and PCM), and they differ in their latency assumptions for the processing branches. The corresponding MPT-RT model equations are listed in the online supplementary material. Table 1 presents an overview of the process models and their key assumptions. In what follows, we use the uppercase letter M with a subscript indicating the specific model to denote the proposed MPT-RT implementations of the respective process models (i.e., M_{PCRM} , M_{DIM} , M_{PCM} , and M_{GM}).

Specifically, according to the preemptive conflict-resolution model (M_{PCRM}) (Figure 2), an initial binary decision determines whether the response to a stimulus

1. The latency parameter L_j may vary as a function of prime race, target object presented (i.e., G = gun; T = tool), and the previous process parameter in the branch the latency parameter belongs to (i.e., C , A , $(1 - C)$, or $(1 - A)$). Subscripts $+$ and $-$ indicate that the parameter is conditional on A or $(1 - A)$, respectively.

TABLE 1. Overview of Key Assumptions for Cognitive Process Models of the Weapon Identification Task

Model	Process Structure	Expected Latency Parameter Pattern
M_{PCRM} : Preemptive conflict-resolution model	An initial preemptive decision results in a controlled process branch with slow responses or in automatic process branches with fast responses.	$L_A = L_{(1-A)} > L_C$
M_{DIM} : Default interventionist model	The default automatic process results in a fast response unless the controlled process corrects the response, resulting in a slow correct response.	$L_{(1-C)+} = L_{(1-C)-} > L_{C+} = L_{C-}$
M_{PCM} : Parallel competitive model	Controlled and automatic processes operate in parallel and result in faster congruent response branches compared to slower conflicting response branches.	$L_{G,A} > L_{G,C} = L_{G,(1-C)}; L_{T,(1-A)} > L_{T,C} = L_{T,(1-C)}$
M_{CM} : Guessing model	Successful discrimination results in fast responses, discrimination failure results in additional response time due to guessing.	$L_A = L_{(1-A)} < L_C$

Note. Latency parameters L_i may vary between prime races, target objects (i.e., G = gun; T = tool), and the previous process parameter in the branch the latency parameter belongs to (i.e., C , $(1-C)$, A , or $(1-A)$). Subscripts + and – indicate that the parameter is conditional on A and $(1-A)$, respectively. Concerning the latency patterns, larger L_i parameters indicate faster relative latencies.

is driven by controlled processes with probability C or by automatic processes with probability $(1-C)$. Importantly, the M_{PCRM} assumes that participants adjust their reliance on the controlled versus the automatic process preemptively on a trial-by-trial basis. For example, participants might try to respond more carefully after committing an error by relying more on the controlled process than the automatic process in the upcoming trial. Conversely, participants might choose to rely more often on automatic processes if cognitive resources are scarce or if they are under time pressure (Klauer & Voss, 2008). To reiterate, the controlled process necessarily results in a correct response, whereas the automatic process corresponds to the automatically activated associations of the prime that trigger the response *gun* with probability A and the response *tool* with the complementary probability $(1-A)$ for both target conditions.

Regarding processing times, the M_{PCRM} assumes that automatic processes are equally fast, irrespective of the activated associations. Controlled processes are effortful and thus need more time to determine the response than spontaneous, prime-triggered automatic processes. When formalized in the MPT-RT framework, both automatic processing branches are assumed to be equally fast within each prime-target condition ($L_A = L_{(1-A)}$). Furthermore, latency parameters of the automatic processing branches (L_A and $L_{(1-A)}$) are expected to be larger than the latency parameter of the controlled processing branch (L_C).

In contrast to the M_{PCRM} , the default interventionist model (M_{DIM}) lacks a preemptive conflict-resolution process. According to the M_{DIM} (Figure 3), the automatic process suggests, based on activated associations, a default response towards *gun*

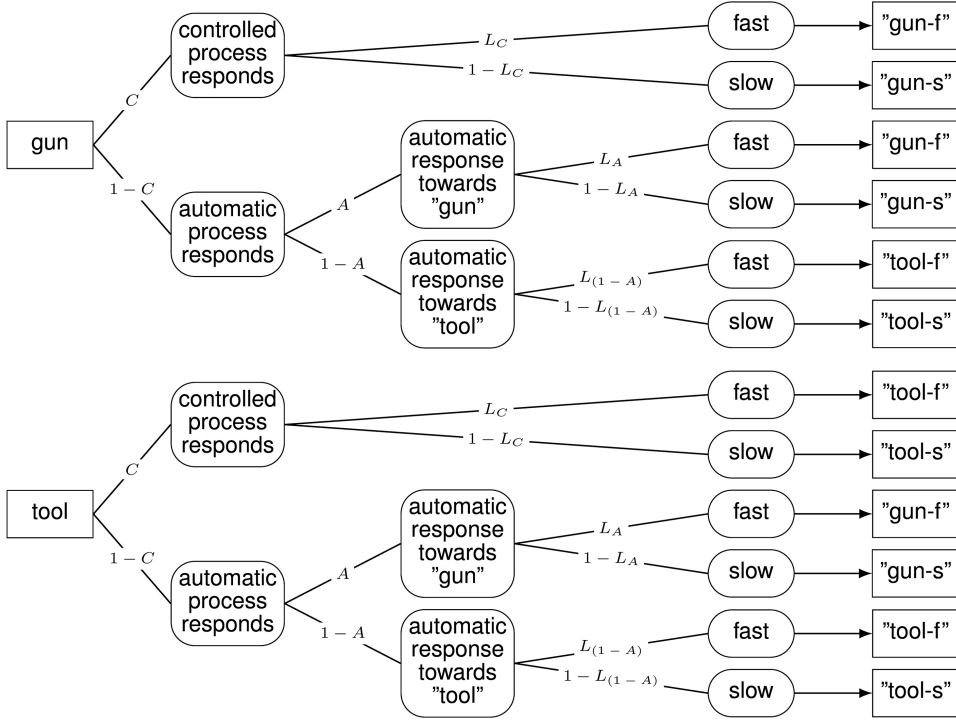


FIGURE 2. MPT-RT model version of the preemptive conflict-resolution model M_{PCRM} . The controlled process parameter C and the automatic process parameter A may vary between prime races. The latency parameters L_i may vary as a function of prime races, target objects, and the previous process parameter in the branch the latency parameter belongs to (i.e., C , A , or $(1 - A)$). The category labels “-f” and “-s” indicate responses categorized as fast and slow, respectively.

with probability A or towards *tool* with probability $(1 - A)$. This default response may then be overcome by the controlled process with probability C , resulting in a correct response. In contrast, if the default response is not overcome by the controlled process with probability $(1 - C)$, the default response suggested by the automatic process is given. The DIM presupposes that the overcoming probabilities of controlled processes do not depend on whether the automatically suggested default response is *gun* (overcoming probability $C+$) or *tool* (overcoming probability $C-$). Hence, there is only a single process parameter for each prime condition $C = C+ = C-$ that characterizes controlled processes in the M_{DIM} .

Regarding processing times, the M_{DIM} assumes that overcoming the default response is effortful and thus takes more time. In contrast, when the controlled process fails to intervene, the response follows a fast, spontaneous, prime-driven automatic process. In the MPT-RT framework, both processing branches that lead to the default response are assumed to be equally fast within each prime-target condition ($L_{(1-C+)} = L_{(1-C-)}$), and the same holds for both controlled processing branches ($L_{(C+)} = L_{(C-)}$). Furthermore, latency parameters of the default response

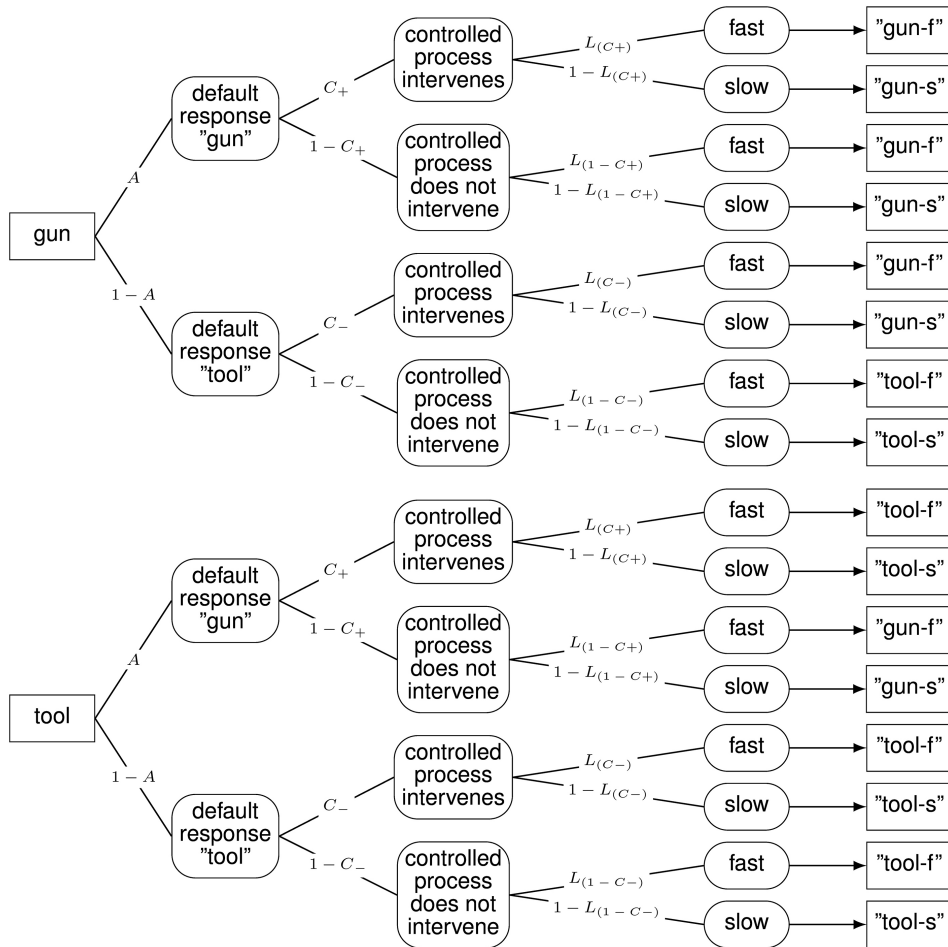


FIGURE 3. MPT-RT model version of the default interventionist model M_{DIM} . The controlled process parameter C and the automatic process parameter A may vary between prime races. Subscripts $+$ and $-$ indicate that the parameter is conditional on the default response “gun” A versus “tool” $(1 - A)$ suggested by the automatic process, respectively. The controlled process parameter C and the automatic process parameter A may vary between prime races. The latency parameters L_i may vary as a function of prime races, target objects and the previous process parameter in the branch the latency parameter belongs to (i.e., C or $(1 - C)$). The category labels “-f” and “-s” indicate responses categorized as fast and slow, respectively.

branches ($L_{(1-C+)}$ and $L_{(1-C-)}$) are expected to be larger than those corresponding to the controlled processing branches ($L_{(C+)}$ and $L_{(C-)}$).

The parallel competitive model (M_{PCM}) (Figure 4) resembles the M_{DIM} in that the automatic process suggests, based on activated associations, a default response towards *gun* with probability A or towards *tool* with probability $(1 - A)$. Other than in the previously discussed models, however, the controlled process runs in parallel to the automatic process in the M_{PCM} . When either of the two parallel processes terminates, it suggests a response. If the automatic process proposes a response

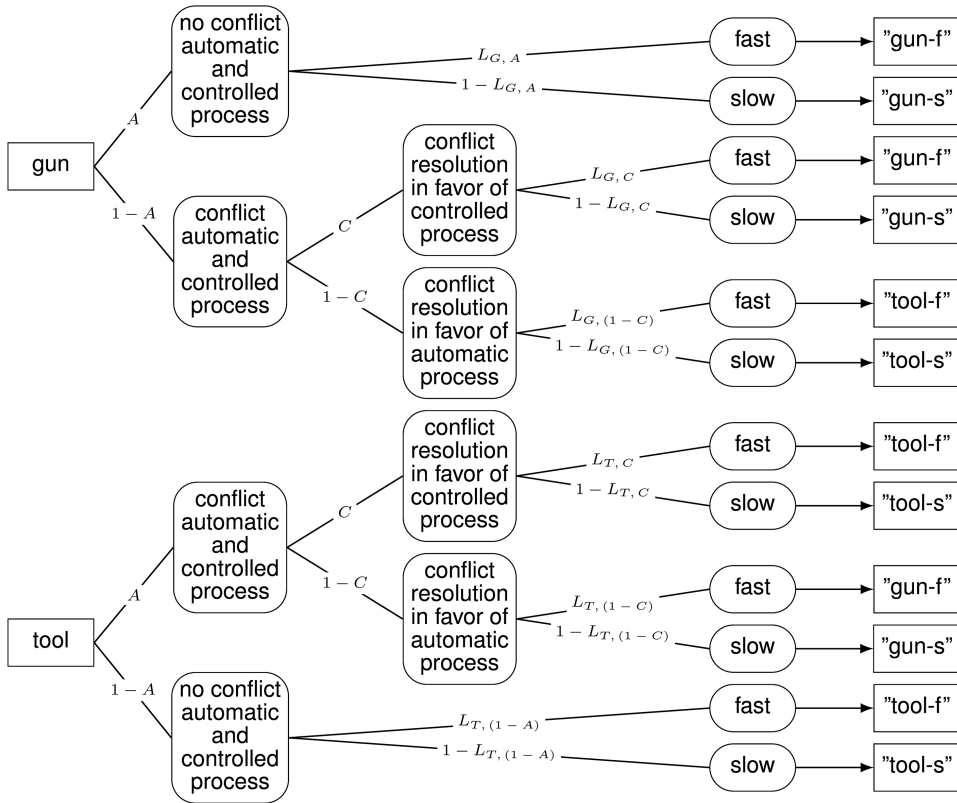


FIGURE 4. MPT-RT model version of the parallel competitive model M_{PCM} . The controlled process parameter C and the automatic process parameter A may vary between prime races. The latency parameters L_i may vary as a function of prime races, target objects ($G = \text{gun}$; $T = \text{tool}$), and the previous process parameter in the branch the latency parameter belongs to (i.e., C , $(1 - C)$, A , or $(1 - A)$). The category labels “-f” and “-s” indicate responses categorized as fast and slow, respectively.

congruent with the target condition (with probability A for the gun target and $(1 - A)$ for the tool target condition), it leads to a correct response in congruence with the controlled process. If the automatic process proposes a response incongruent with the target condition (with probability $(1 - A)$ for the gun target and A for the tool target condition), the conflict is resolved in favor of the controlled process with probability C and in favor of the automatic process with probability $(1 - C)$.

Regarding processing times, the M_{PCM} assumes that processing branches characterized by congruency of suggested responses are fast (i.e., $L_{G,A}$ and $L_{T,(1-A)}$) whereas processing branches characterized by conflict between processes are equally slow (i.e., $L_{G,C} = L_{G,(1-C)}$ and $L_{T,C} = L_{T,(1-C)}$).

Finally, according to the guessing model (M_{GM}), the response to a stimulus is driven by either a controlled or an automatic process with probability C and $(1 - C)$, respectively. The basic processing-tree structure of the M_{GM} is equivalent

to the M_{PCRM} (Figure 2) and thus does not require a new figure for purposes of illustration. However, both the psychological interpretation and the restrictions imposed on the parameters differ between the M_{GM} and the M_{PCRM} , making them distinguishable in terms of their empirical predictions. In the M_{GM} , the controlled process succeeds in determining the correct response with probability C . If the controlled process fails with probability $(1 - C)$, the response is guessed as *gun* or *tool* with probabilities A and $(1 - A)$, respectively, depending on stereotypes activated by the prime. Regarding processing times, the M_{GM} assumes that the controlled process results in faster responses than those of the automatic processing branches because additional guessing is involved in the latter. In the MPT-RT framework, both automatic processing branches are assumed to be equally fast within each prime-target condition ($L_A = L_{(1-A)}$), and latency parameters of the automatic processing branches (L_A and $L_{(1-A)}$) are expected to be smaller than the latency parameter of the controlled processing branch (L_C).

In sum, although the M_{GM} and the M_{PCRM} are equivalent in their MPT model equations, they differ with respect to the order restrictions imposed on the latency parameters. Specifically, automatic processing branches are faster in the M_{PCRM} ($L_C < L_A = L_{(1-A)}$) whereas controlled processing branches are faster in the M_{GM} ($L_C > L_A = L_{(1-A)}$).

OVERVIEW OF MODEL COMPARISONS

As explained above, all four psychological process models are instantiations of the PDP but make different assumptions about the nature and the chronological order of the underlying cognitive processes. Hence, these models cannot be distinguished based on the accuracy of responses across the four prime-target conditions alone. However, because these process models impose different constraints on the latency parameters for each branch within the MPT-RT framework, they can be tested against each other, provided that both response frequencies and response latencies are available for all prime-target conditions of the WIT.

An appropriate framework for testing the four process models is response time-extended multinomial processing tree modeling (MPT-RT; Heck & Erdfelder, 2016). MPT-RT models enable the joint estimation of the core parameters C and A as well as the relative branch latencies L_j . Moreover, they allow researchers to test different sets of equality constraints and order restrictions on the parameters (Knapp & Batchelder, 2004), such as those implied by the four process models. In contrast, previous work had to rely on comparisons of observed overall response times for correct and false responses across conditions (Klauer & Voss, 2008), that is, response times aggregated across different processing branches that terminate in the same response.

In what follows, we first evaluate the MPT-RT model fit of the psychological process models by re-analyzing eight previously published WIT data sets (Model Comparison 1). This allows us to identify whether there are specific process models that are in line with the available data originating from different samples, stimuli, and researchers applying the WIT. Second, we compare different sets of additional restrictions on the branch latencies and try to identify more parsimonious model

versions that describe the observed data with fewer parameters (Model Comparison 2). Based on a simplifying restriction pattern that is in line with our data, we propose two extended MPT-RT models that discriminate between the best-fitting models of the previous comparison—the M_{PCRM} and the M_{DIM} —using plausible auxiliary assumptions (Model Comparison 3). For the sake of brevity, we focus on goodness-of-fit measures of the estimated models only. Group level parameter estimates and their corresponding descriptive statistics are provided in the online supplementary material for Model Comparisons 1, 2, and 3 (<https://osf.io/7vjrj>).

MODEL COMPARISON 1: TESTING THE FIT OF THE FOUR PROCESS MODELS

In Model Comparison 1, we estimated MPT-RT implementations of the four psychological process models M_{PCRM} , M_{DIM} , M_{PCM} , and M_{GM} (see Figures 2 to 4) with order restrictions on the parameters as listed in Table 1. The equations for each model can be found in the online supplementary materials. We performed model comparisons across eight data sets of the WIT as described in the next section.

METHOD

Study Search and Inclusion Criteria. Starting from the set of studies included in the meta-analysis by Rivers (2017), we selected North American studies that made use of the standard four-condition WIT paradigm, that is, investigated cross-classified White male faces and Black male faces as primes with guns and tools as target objects in 2×2 within-subject designs. We obtained a total of eight relevant data sets originating from five publications. The respective data sets were either available through the Open Science Framework (i.e., Madurski & LeBel, 2015; Rivers, 2017), or obtained by contacting the first or second author (i.e., Amon & Holden, 2016; Correll, 2008; Lambert et al., 2003). Table 2 provides an overview of the main characteristics of each data set such as the number of participants ($N_{\text{sample}} = 22\text{--}50$), the percentage of Caucasian participants (17%–83%), the number of trials ($N_{\text{trials}} = 200\text{--}1,100$), prime duration (200 ms–300 ms), and the response time limit (500 ms–1,000 ms). All data sets were obtained in the U.S. or in Canada and include responses from primarily non-Black university students as participants. We used the data sets in the same way as they were used in the original publications and provided by the authors; no further participants were excluded. The single exception was the study by Lambert et al. (2003) for which only a part of the data set included response times in addition to response frequencies. If a data set contained additional experimental manipulations (Amon & Holden, 2016; Correll, 2008; Lambert et al., 2003; Madurski & LeBel, 2015) or additional prime stimulus materials (Rivers, 2017) compared to the standard procedure of the WIT (Payne, 2001), we excluded these trials from data analysis.

MPT-RT Modeling Procedure. We used the hierarchical latent-trait model of Klauer (2010) as the statistical framework for all analyses. This model family

TABLE 2. Characteristics of the Eight WIT Data Sets Used in the Present Analyses

Article	Study	N_{sample}	% of Caucasian participants	N_{trials}	Prime duration [ms]	Response time limit [ms]	Subset of the data for reanalysis
Amon & Holden (2016)	1a	32	83% (93% non-Black)	1,100	200	1,000	only control group
Amon & Holden (2016)	1c	32	83% (93% non-Black)	1,100	300	500	only control group
Correll (2008)	2	24	not reported	200	200	1,000	only control group
Lambert et al. (2003)	2	22	all non-Black	384	200	550	only private response group
Madurski & LeBel (2015)	1a	48	62% (98% non-Black)	200	200	1,000	only control group
Madurski & LeBel (2015)	1b	50	60% (98% non-Black)	200	200	1,000	only control groups
Rivers (2017)	1a	40	17% (98% non-Black)	216	200	500	neutral faces excluded
Rivers (2017)	1b	42	17% (98% non-Black)	216	200	1,000	neutral faces excluded

Note. For the data from Study 2 of Lambert et al. (2003), individual accuracy and response-time data were available for only 22 of 61 participants.

allows for variability of all parameters across individuals such that the probit-transformed person parameters follow a joint multivariate normal distribution at the group level. In contrast to alternative hierarchical MPT models, the latent-trait model not only provides both individual and group-level parameters (i.e., parameter means) but also variances and correlations of person parameters across participants.

For assessing goodness of fit, we used the Bayesian posterior predictive p value based on the test statistics $T1$ and $T2$ (Klauer, 2010; Klauer et al., 2015). $T1$ summarizes how well the model accounts for the average response frequencies across participants. It resembles the Pearson goodness-of-fit statistic χ^2 often used in categorical data analysis (Klauer et al., 2015; Riefer & Batchelder, 1991). The test statistic $T2$, in contrast, summarizes how well the model accounts for the variances and correlations of the response frequencies across participants. The posterior predictive p value represents the comparison of the calculated test statistics obtained for observed and predicted response frequencies. A value of $p > .05$ is typically seen as evidence that model assumptions are in line with the data (Klauer, 2010; Klauer et al., 2015).

We used the R package TreeBUGS (Heck et al., 2018) to fit hierarchical versions of the four MPT-RT models to the eight data sets of interest. The Markov Chain Monte Carlo (MCMC) algorithm was run for three independent estimation chains with 750,000 iterations each, of which 250,000 were removed as a burn-in period. Every 50th iteration was retained to compute summary statistics. For models M_{PCRM} , M_{DIM} , M_{PCM} , and M_{GM} , the parameters A and C were allowed to vary between primes (i.e., Black vs. White male face primes), and latency parameters L_j were allowed to vary between primes, target object conditions (i.e., guns vs. tools),

and the previous process parameter in the branch the latency parameter belongs to (i.e., C , A , $(1 - C)$, or $(1 - A)$). The Rubin-Gelman statistic $\hat{R}$ was smaller than 1.05 for all parameter estimates across all models, showing an acceptable convergence of MCMC sampling.

Response-Time Data Preparation. Regarding response times, we included all trials provided by the authors of the data sets, that is, all trials with response times faster than the response-time limit of the corresponding data set (between 500 ms and 1000 ms). To estimate MPT-RT models, we divided trials into fast and slow responses for each participant based on whether a specific response was faster or slower than the individual mean of their log-transformed response times across response categories (see Heck & Erdfelder, 2016). This gives us two discrete RT bins of fast and slow responses for every correct and false response across conditions, separately for each individual. Although categorization of response times results in a loss of information compared to analyzing continuous response times, participant-specific categorization in fast and slow responses is advantageous because estimation is simpler, more robust, and does not need any a priori assumptions about the distribution of response times (Heck & Erdfelder, 2016).

RESULTS

Table 3 shows the goodness-of-fit measures for the four psychological process models M_{PCRM} , M_{DIM} , M_{PCM} , and M_{GM} including the respective order restrictions on the latency parameters for each data set. Regarding model fit statistics, models M_{PCRM} and M_{DIM} show acceptable fit in both test statistics for almost all data sets ($p(T1)s > .218$; $p(T2)s > .177$). The single exception is Study 1a of Amon and Holden (2016) with misfit in the $T2$ statistic ($p(T1) < .006$). In stark contrast, models M_{PCM} and M_{GM} show blatant misfit in both test statistics for all data sets (all $p(T1)s < .001$; all $p(T2)s < .008$).

DISCUSSION

The data clearly favor the preemptive conflict-resolution model M_{PCRM} and the default interventionist model M_{DIM} . In contrast, the parallel competitive model M_{PCM} and the guessing model M_{GM} are clearly rejected by the data. The reason for the good fit of M_{PCRM} and M_{DIM} is the typically observed data pattern that WIT error latencies are faster than correct response latencies. While this pattern can be accommodated by M_{PCRM} and M_{DIM} , it conflicts with M_{PCM} and M_{GM} . Hence, the joint modeling of response frequencies and response times results in the same substantive conclusions as the model-free analysis of response times by Klauer and Voss (2008), namely, that automatic and controlled processing routes are fast and slow, respectively, in line with the PCRM and the DIM. Specifically, automatic processes lead to both fast incorrect and fast correct responses as a consequence of stereotypical associations induced by the prime's race. In contrast, controlled process branches involve more time and effort to identify the target object correctly

TABLE 3. Fit Statistics for the Four Process Models

Article	Study	M _{PCRM}		M _{DIM}		M _{PCM}		M _{GM}	
		<i>p</i> (T1)	<i>p</i> (T2)	<i>p</i> (T1)	<i>p</i> (T2)	<i>p</i> (T1)	<i>p</i> (T2)	<i>p</i> (T1)	<i>p</i> (T2)
AH	1a	.232	.006	.218	.005	< .001	< .001	< .001	< .001
AH	1c	.503	.370	.500	.374	< .001	< .001	< .001	< .001
C	2	.451	.425	.447	.415	< .001	.008	< .001	.008
L	2	.458	.200	.462	.200	< .001	< .001	< .001	< .001
ML	1a	.510	.245	.512	.239	< .001	< .001	< .001	< .001
ML	1b	.525	.194	.521	.177	< .001	< .001	< .001	< .001
R	1a	.528	.338	.519	.348	< .001	< .001	< .001	< .001
R	1b	.507	.236	.502	.246	< .001	.003	< .001	.007

Note. References are abbreviated as follows: AH = Amon & Holden (2016), C = Correll (2008), L = Lambert et al. (2003), ML = Madurski & LeBel (2015), R = Rivers (2017). Models are abbreviated as follows: M_{PCRM} = preemptive conflict-resolution model, M_{DIM} = default interventionist model, M_{PCM} = parallel competitive model, M_{GM} = guessing model.

(i.e., target discrimination for the PCRM and target discrimination with default response inhibition for the DIM).

Importantly, despite different theoretical assumptions underlying each model, M_{PCRM} and M_{DIM} are mathematically equivalent if independence of automatic and controlled processes is assumed, that is, if controlled process parameters (and their corresponding branch latency parameters) do not differ between congruent and incongruent default responses in the M_{DIM} (i.e., C = C+ = C–; see Figure 3), as is the case here. Hence, the latencies of the C+ and C– branches cannot be estimated separately. In other words, to distinguish the influence of congruent versus incongruent automatic responses on branch latencies in M_{DIM}, additional assumptions about the latency parameters need to be incorporated.

To gain more degrees of freedom for tests between the M_{PCRM} and the M_{DIM}, we opted to search for a more parsimonious model (Model Comparison 2). Using this approach, influences of congruent versus incongruent automatic responses on branch latencies can be estimated separately (Model Comparison 3), which in turn allows for a test between M_{PCRM} and M_{DIM}.

MODEL COMPARISON 2: CONSTRAINING PROCESS LATENCIES ACROSS CONDITIONS

The M_{PCRM} and the M_{DIM} previously analyzed in Model Comparison 1 provide for separate controlled and automatic branch latency parameters for each prime-target condition. Here, we consider the possibility that this assumption is unnecessarily complex because corresponding branch latencies may not differ between prime or target conditions or even both.

For automatic process routes, we assume that latency parameters do not differ between prime-target conditions. This is reasonable because the automatic

associations are instantly triggered by the prime and thus do not depend on the target object that follows later. Moreover, even if relative response latencies of automatic process branches should differ slightly between prime conditions, we expect these differences to be negligible compared to the relatively slow controlled process branch latencies.

For the controlled process routes, we assume that response latencies may differ between gun and tool targets while being independent of prime condition (i.e., whether a Black or White face was displayed before). This assumption is plausible for two reasons. First, we expect the speed of controlled processes to vary between target object conditions because previous WIT research has shown that participants' correct responses tend to be faster for gun than for tool targets (Payne, 2001; Rivers, 2017). Second, equality of latency parameters within target condition is in line with the invariance assumption of the PDP for the controlled process parameters. This invariance assumption presupposes that target discrimination is not affected by prime race (Payne et al., 2005). For a full discussion of the validity and possible criticism of the invariance assumption, see the General Discussion section below.

To test this restriction pattern in the MPT-RT framework, we propose model M_{Basic} as a starting point. Model M_{Basic} includes the mathematically equivalent models M_{PCRM} and M_{DIM} as submodels (see online supplementary material). More precisely, the MPT-RT model equations of M_{Basic} are equivalent to those submodels but do not impose any order restrictions on the latency parameters (as those listed in Table 1). Starting from M_{Basic} , the more parsimonious nested model M_{Pars} imposes additional equality restrictions on automatic and controlled processing branches as outlined above. Specifically, for the automatic processing branches, latency parameters are equated across all four prime-target conditions. This refers to parameters L_A and $L_{(I-A)}$ for M_{PCRM} (see Figure 2) and latency parameters $L_{(I-C+)}$ and $L_{(I-C-)}$ for M_{DIM} (see Figure 3), respectively. For the controlled processing branches, latency parameters are equated within target conditions. This refers to parameter L_C for M_{PCRM} and latency parameters $L_{(C+)}$ and $L_{(C-)}$ for M_{DIM} , respectively.

METHOD

We again fitted the hierarchical MPT models using TreeBUGS (Heck et al., 2018). For comparisons between models taking model complexity into account, we used the deviance information criterion (DIC; Klauer et al., 2015; Spiegelhalter et al., 2002) and the widely applicable information criterion (WAIC; Vehtari et al., 2017; Watanabe, 2010). The DIC is a hierarchical modeling generalization of the Akaike information criterion (AIC). The DIC consists of a term quantifying lack of model fit and a term penalizing model complexity (Klauer, 2010; Klauer et al., 2015; Spiegelhalter et al., 2002, 2014). The WAIC can be seen as an improved version of the DIC, as it is fully Bayesian and based on the entire posterior distribution of participants and not solely on a point estimate. The WAIC consists of a term summed across the estimated log pointwise predictive density and a penalizing term summed across the estimated variance of the log pointwise predictive density (Vehtari et al., 2017).

Like AIC, the model with the lowest DIC or WAIC value represents the best compromise between fit and complexity, with an absolute Δ DIC (Δ WAIC) difference larger than 2 usually interpreted as evidence for one model compared to the other (Burnham & Anderson, 2004; Klauer et al., 2015; Spiegelhalter et al., 2002).

We used the same estimation settings as in Model Comparison 1. The Rubin-Gelman statistic $\hat{R}$ was smaller than 1.05 for all parameter estimates across all models. To obtain stable goodness-of-fit measure statistics, we estimated each model 10 times and calculated means and standard errors for the model fit statistics $p(T1)$, $p(T2)$, DIC, and WAIC.

RESULTS

Table 4 shows the model fit statistics $p(T1)$ and $p(T2)$ of M_{Basic} and M_{Pars} , as well as DIC and WAIC differences between models. Model M_{Basic} fits all data sets (all $p(T1)s > .473$; all $p(T2)s > .163$). The parsimonious model M_{Pars} fits almost all data sets (all $p(T1)s > .144$; all $p(T2)s > .084$) except Study 1a of Amon and Holden (2016) and Study 1b of Rivers (2017) ($p(T1)s < .037$).

Δ DIC (Δ WAIC) was calculated as the difference of the DIC (WAIC) measures between M_{Pars} and M_{Basic} . Hence, a negative Δ DIC (Δ WAIC) supports M_{Pars} relative to M_{Basic} . As evident from Table 4, for five data sets DIC favors M_{Pars} more than M_{Basic} (Δ DICs < -5.35). For three data sets no model is favored (Δ DICs $< |1.52|$). With respect to WAIC, two data sets favor M_{Pars} more than M_{Basic} (Δ WAICs < -9.77), while preferences are reversed for five data sets (Δ WAICs > 17.05), and for one data set no model is favored (Δ WAIC $< |0.29|$).

In sum, both M_{Basic} and M_{Pars} fit almost all data sets. While model comparisons using DIC revealed a preference for M_{Pars} , those based on WAIC showed the reversed pattern.

DISCUSSION

Overall, the parsimonious model M_{Pars} fits the data almost as well as the more general model M_{Basic} . Although model comparison using DIC and WAIC diverged slightly, the posterior predictive p values for $T1$ and $T2$ showed that the nested model M_{Pars} does a good job in describing both the observed frequencies and the covariance structure of the responses for nearly all data sets.² For the weapon identification task, this implies that responses based on automatic processes tend to be equally fast irrespective of elicited automatic response and prime race. However, responses based on the controlled process may depend on the to-be-identified target. This may be due to differences between stimuli in terms of visual discriminability. Alternatively, it may also reflect participants' strategy to respond to a threatening weapon target as soon as they identified it but be more hesitant when

2. Additional model comparisons based on nested latency parameter restrictions for the automatic process routes (across prime conditions and prime-target conditions) and for the controlled process routes (across target conditions and prime-target conditions) are reported in the online supplementary materials.

TABLE 4. Fit Statistics for Process Model Comparison

Article	Study	M_{Basic}		M_{Pars}		Comparison	
		$p(T1)$	$p(T2)$	$p(T1)$	$p(T2)$	ΔDIC	ΔWAIC
AH	1a	.498 [.001]	.451 [.002]	.037 [.001]	.557 [.002]	0.36 [0.10]	61.91 [0.23]
AH	1c	.492 [.001]	.354 [.002]	.541 [.002]	.379 [.001]	−37.18 [0.08]	−9.77 [0.17]
C	2	.487 [.002]	.357 [.001]	.594 [.001]	.230 [.001]	−16.88 [0.09]	28.45 [0.12]
L	2	.485 [.002]	.232 [.001]	.144 [.002]	.336 [.002]	−23.09 [0.09]	−13.08 [0.10]
ML	1a	.498 [.002]	.220 [.002]	.370 [.001]	.381 [.002]	−32.40 [0.20]	0.29 [0.20]
ML	1b	.480 [.002]	.163 [.002]	.174 [.001]	.086 [.001]	−5.35 [0.14]	20.63 [0.18]
R	1a	.519 [.001]	.370 [.002]	.347 [.001]	.084 [.001]	−0.69 [0.17]	17.05 [0.11]
R	1b	.473 [.002]	.244 [.002]	< .001 [.001]	.437 [.002]	1.52 [0.21]	30.49 [0.17]

Note. $p(T1)$, $p(T2)$, ΔDIC , and ΔWAIC are means with standard errors in brackets across 10 independent estimations. Negative values for ΔDIC and ΔWAIC indicate a preference for model M_{Pars} over model M_{Basic} . References are abbreviated as follows: AH = Amon & Holden (2016), C = Correll (2008), L = Lambert et al. (2003), ML = Madurski & LeBel (2015), R = Rivers (2017). Models are abbreviated as follows: M_{Basic} = basic model, M_{Pars} = parsimonious model.

they believed they identified a tool, primarily to avoid missing a weapon target. This is in line with previous findings showing that participants tend to respond faster to threat-related targets (Payne, 2001; Rivers, 2017).

For follow-up model comparisons between M_{PCRM} and M_{DIM} (Model Comparison 3), we rely on the latency-parameter restrictions across conditions imposed by M_{Pars} as auxiliary assumptions. The additional auxiliary assumptions provide a more parsimonious model with additional degrees of freedom. This allows us to test between different order restrictions for the controlled processing branches implied by the PCRM and the DIM.

MODEL COMPARISON 3: TESTING EXTENDED VERSIONS OF THE PCRM AND THE DIM

Models M_{PCRM} and M_{DIM} are mathematically equivalent if controlled process parameters (and their corresponding branch latency parameters) do not differ between congruent and incongruent default responses in model M_{DIM} , that is, if $C = C+ = C-$ (Figure 3). To distinguish between these two models, further assumptions about the latency parameters are necessary. According to Klauer and Voss (2008), the DIM is likely to be more appropriate for the WIT than the PCRM because the DIM can explain faster identification of weapons after Black face primes, even when the success probability of controlled processes is high. This data pattern can be explained by assuming that controlled processes are faster following default responses congruent with the target condition compared to incongruent default responses. As outlined by Klauer and Voss (2008), the PCRM lacks a plausible mechanism to explain this outcome.

To test this idea, we combined the order constraints of the two well-fitting models of Model Comparison 1 (M_{PCRM} and M_{DIM}) with the equality constraints of the parsimonious model M_{Pars} from Model Comparison 2. To reiterate, according to M_{Pars} the relative speed of automatic processing branches does not differ between different prime-target conditions, and the relative speed of controlled process branches does not differ within target conditions. By combining the different sets of order constraints of M_{PCRM} and M_{DIM} (Table 1) with the equality constraints of M_{Pars} , two extended model versions $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ are obtained. In contrast to the model version considered in Model Comparison 1, $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ have fewer parameters than degrees of freedom available in the data. Goodness-of-fit differences between both models can therefore be assessed.

Model $M_{\text{PCRM,ext}}$ mirrors the MPT-RT structure of M_{PCRM} model (Figure 2) but imposes the additional constraints that automatic processing branch latencies do not differ between prime-target conditions and that controlled processing branch latencies do not differ within target conditions. Regarding order restrictions for latency parameters, $M_{\text{PCRM,ext}}$ assumes that automatic processing branches are generally faster than controlled processing branches in each target condition (i.e., $L_A > L_{G,C}$ and $L_A > L_{T,C}$).

Analogously, model $M_{\text{DIM,ext}}$ has the same MPT-RT structure as M_{DIM} (Figure 3) but imposes the additional constraints that latencies of automatic processing branches do not differ between prime-target conditions and that latencies of controlled processing branches do not differ within target conditions. Regarding order restrictions for latency branches, recall that model M_{DIM} assumes that controlled processing branches are equally fast ($L_{C+} = L_{C-}$), regardless of whether they follow a default response that is congruent or incongruent with the target condition. In contrast, the extended model $M_{\text{DIM,ext}}$ assumes that controlled processes following congruent default responses are faster than those following incongruent default responses. In sum, according to $M_{\text{DIM,ext}}$, automatic processing branches are generally faster than controlled processing branches for congruent responses in each target condition. These, in turn, are faster than controlled processing branches for incongruent responses in each target condition (i.e., $L_A > L_{G,C+} > L_{G,C-}$ and $L_A > L_{T,C-} > L_{T,C+}$).

METHOD

We used the same methods and estimation settings as in Model Comparisons 1 and 2. $\hat{R}$ was smaller than 1.05 for all parameter estimates across all models. Again, to obtain stable goodness-of-fit statistics, we estimated each model 10 times and calculated means and standard errors for the model fit statistics $p(T1)$, $p(T2)$, DIC, and WAIC.

RESULTS

Table 5 shows the model fit statistics $p(T1)$ and $p(T2)$ of $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ as well as DIC and WAIC differences between models. Model $M_{\text{PCRM,ext}}$ fits almost

TABLE 5. Fit Statistics and Model Selection Measures for the Extended Model Versions of PCRM and DIM

Article	Study	$M_{\text{PCRM,ext}}$		$M_{\text{DIM,ext}}$		Comparison	
		$p(T1)$	$p(T2)$	$p(T1)$	$p(T2)$	ΔDIC	ΔWAIC
AH	1a	.004 [.001]	.031 [.001]	.004 [.001]	.031 [.001]	−10.12 [0.20]	−21.62 [0.23]
AH	1c	.548 [.001]	.408 [.001]	.520 [.002]	.388 [.001]	2.71 [0.11]	−0.14 [0.17]
C	2	.618 [.001]	.233 [.001]	.519 [.002]	.379 [.001]	−3.38 [0.08]	−3.20 [0.15]
L	2	.146 [.001]	.330 [.001]	.209 [.001]	.460 [.002]	−10.31 [0.10]	−13.08 [0.10]
ML	1a	.329 [.001]	.336 [.002]	.337 [.001]	.371 [.002]	−1.35 [0.11]	−9.31 [0.20]
ML	1b	.186 [.001]	.087 [.001]	.349 [.002]	.157 [.001]	−8.38 [0.16]	−17.94 [0.18]
R	1a	.346 [.001]	.082 [.001]	.496 [.002]	.309 [.001]	−8.26 [0.11]	−16.84 [0.14]
R	1b	< .001 [.001]	.432 [.001]	.230 [.002]	.440 [.002]	−18.59 [0.16]	−27.91 [0.20]

Note. $p(T1)$, $p(T2)$, ΔDIC , and ΔWAIC are means with standard errors in brackets across 10 independent estimations. Negative values for ΔDIC and ΔWAIC indicate a preference for model $M_{\text{DIM,ext}}$ relative to $M_{\text{PCRM,ext}}$. References are abbreviated as follows: AH = Amon & Holden (2016), C = Correll (2008), L = Lambert et al. (2003), ML = Madurski & LeBel (2015), R = Rivers (2017). Models are abbreviated as follows: $M_{\text{PCRM,ext}}$ = extended version of preemptive conflict-resolution model, $M_{\text{DIM,ext}}$ = extended version of default interventionist model.

all data sets (all $p(T1)$ s > .146; all $p(T2)$ s > .082), except for Study 1a of Amon and Holden (2016; $p(T1)$ = .004 and $p(T2)$ = .031) and Study 1b of Rivers (2017; $p(T1)$ < .001). Similarly, model $M_{\text{DIM,ext}}$ fits almost all data sets (all $p(T1)$ s > .209 and all $p(T1)$ s > .157), except for Study 1a of Amon and Holden (2016; $p(T1)$ = .004 and $p(T2)$ = .031).

The ΔDIC (ΔWAIC) was calculated as the difference of the DIC (WAIC) measures between $M_{\text{DIM,ext}}$ and $M_{\text{PCRM,ext}}$. Hence, a negative ΔDIC (ΔWAIC) supports $M_{\text{DIM,ext}}$ relative to $M_{\text{PCRM,ext}}$. Note that we chose to not interpret ΔDIC and ΔWAIC for Study 1a of Amon and Holden (2016) as neither $M_{\text{PCRM,ext}}$ nor $M_{\text{DIM,ext}}$ fits the data well. For the remaining seven data sets, DIC favors $M_{\text{DIM,ext}}$ more than $M_{\text{PCRM,ext}}$ in five cases (ΔDIC s < −3.38) and WAIC even in six cases (ΔWAIC s < −3.20). While DIC preferences were reversed for one data set (ΔDIC s = 2.71), we observed no preference reversals for WAIC. None of the models is favored for one data set, and this holds with respect to both DIC (ΔDIC s = −1.35) and WAIC (ΔWAIC s = −0.14).

DISCUSSION

In sum, both models $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ fit almost all data sets. This notwithstanding, there is a general preference for $M_{\text{DIM,ext}}$, irrespective of the model selection measure used. The only exceptions are studies by Amon and Holden (2016), which show either no fit (Study 1a) or no preference for $M_{\text{DIM,ext}}$ (Study 1b). These discrepancies are perhaps due to a significantly larger number of trials (N_{trials} = 1,100) in Amon and Holden’s studies than in the other studies (all N_{trials} ≤ 384). The total

number of trials may affect the general strategy of how participants perform in the WIT.

The preference for $M_{\text{DIM,ext}}$ indicates that automatically activated default responses always interfere with weapon identification in every trial. According to $M_{\text{PCRM,ext}}$, in contrast, only some trials are influenced by automatic processes. Regarding controlled responding, $M_{\text{DIM,ext}}$ maintains that two controlled processes play an important role in weapon identification, namely (1) target discrimination and (2) overcoming the default response when the latter turns out to be incongruent with the to-be-identified target.

GENERAL DISCUSSION

The weapon identification task (WIT) is a widely used sequential priming procedure to assess the influence of racial threat stereotypes on weapon identification (Payne, 2001; Rivers, 2017). The process dissociation procedure (PDP) has routinely been used as measurement model for the WIT to estimate the influence of controlled and automatic processes on task performance and to assess how these processes are affected by different factors. For example, participants show less controlled processing for shorter response-time windows (Payne, 2001; Payne et al., 2002) when more anxious because they expect public accountability for their own performance (Lambert et al., 2003) and when their cognitive resources are limited (Govorun & Payne, 2006; Ito et al., 2015; Payne, 2005). Furthermore, the automatic association of Black male face primes with guns is enhanced when race is made salient (Payne et al., 2002) and participants are in a positive mood (Huntsinger et al., 2009) or can be diminished when participants use implementation intentions to think “safe” for Black face primes (Stewart & Payne, 2008).

Although the proportion of responses determined by controlled or automatic processes might be influenced by these factors, their nature, relationship, and temporal order remains underdetermined in the standard PDP framework. For example, the controlled process can entail target discrimination and overcoming of automatic associations, while the automatic process can entail activation of racial associations or a prime-informed guessing tendency. In fact, the PDP is consistent with four process models for the WIT: the preemptive conflict-resolution model (PCRM), the default interventionist model (DIM), the guessing model (GM), and the parallel competitive model (PCM).

In this article, we formalized these process models in the framework of response time-extended multinomial processing tree (MPT-RT) models (Heck & Erdfelder, 2016). This approach has several advantages. First, MPT-RT models combine information on response frequencies and response times in the same model. Second, MPT-RT models provide model fit and model selection statistics to test between (or select among) models. Third, MPT-RT models provide for equality and order restrictions on both the structural and the latency parameters of the respective MPT-RT model. Most importantly in the present context, the MPT-RT framework allows testing the four process models of the WIT against each other. This is not

possible when analyzing response frequencies in isolation (as in the standard PDP). Essentially, the different process models become empirically distinguishable after extending the PDP model with latency parameters for fast and slow responses in different processing branches. The major advantage is that model fit can be assessed for response frequencies and response latencies conjointly rather than focusing on each dependent variable in two separate steps.

TESTING BETWEEN PSYCHOLOGICAL PROCESS MODELS

Our results support the default interventionist model (DIM) and the preemptive conflict-resolution model (PCRM), but not the guessing model (GM) nor the parallel competitive model (PCM) for the WIT. For all eight data sets considered here, the DIM and PCRM fitted the data very well. Based on the parameter estimates we observed, both models predict controlled process branches to be slower than automatic process branches. This is in line with the result pattern typically observed for the WIT (i.e., fast errors and slow correct responses).

Further model comparisons based on additional invariance assumptions deemed the DIM as a more appropriate process model for the WIT than the PCRM. This is in line with the reasoning of Klauer and Voss (2008), who argued that the DIM can explain that, in studies with high accuracy rates, response bias due to prime race persists in response latencies. According to the DIM, overcoming the default response by controlled processing leads to slower responses if automatic and controlled processes disagree than if both agree. The PCRM cannot explain this pattern easily because the latency of the controlled process is assumed to be independent of the outcome of the automatic process.

Additional support for the DIM comes from research on brain region activation after seeing Black and White faces (Cunningham et al., 2004). Specifically, a short presentation (30 ms) of faces leads to higher activation of the amygdala for Black than for White faces, a brain region typically responsive to the emotionality of a stimuli (e.g., induced threat). However, a longer presentation time (525 ms) of Black and White faces leads to no difference in amygdala activation. This lack of difference in amygdala activation for longer presentation times of faces is associated with increased activity of the dorsolateral prefrontal cortex and the anterior cingulate, brain regions associated with controlled processing and executive function. In line with the DIM, this suggests that, in a first step, automatic associations and emotional reactions are instantly activated when processing a face stimulus, while in a second step (about half a second later) these associations can be inhibited and modulated by controlled processes.

Adopting the DIM as the best-fitting model of cognitive processes underlying performance in the weapon identification task thus implies that automatic threat-stereotype associations interfere with object identification throughout each trial. Hence, controlled processes are required for both object discrimination *and* conflict resolution when the automatic process elicits an incongruent default response. In other words, object discrimination abilities and cognitive control for overcoming default responses *both* play a crucial role in WIT performance.

PSYCHOLOGICAL IMPLICATIONS

The assumption of automatically generated default responses in the DIM nicely fits findings regarding the malleability of automatic racial stereotyping depending on participants' latent focus on race at the beginning of each trial. For example, race salience (Jones & Fazio, 2010; Payne et al., 2002) strengthens the association of Black males with guns, whereas focusing on other face characteristics (e.g., age) diminishes the association of Black males with guns (Jones & Fazio, 2010; Todd et al., 2021). Similarly, implementation intentions (i.e., previously determined if-then action plans) allow adjusting associations (e.g., as safe) for specific face primes (Stewart & Payne, 2008). Hence, even if face primes elicit a default response at the beginning of each trial, this response is malleable if the associations tied to the face primes are adjusted in advance.

Similarly, findings regarding reduced capabilities of participants for correct responding are easily explained by failures of the controlled process in target identification and overcoming bias. For example, shorter response deadlines, stress, and diminished cognitive capabilities are known to compromise target discrimination as well as conflict resolution to overcome bias (Govorun & Payne, 2006; Ito et al., 2015; Lambert et al., 2003; Payne, 2001, 2005; Payne et al., 2002). Furthermore, this is in line with the findings that participants with higher cognitive inhibitory capabilities (Ito et al., 2015) or with higher internal motivation to control their prejudice (Volpert-Esmond et al., 2020) show more controlled responding.³ This suggests that more practice in resolving conflicts induced by automatic default responses results in more correct responding. Hence, strategy training and interventions that strengthen object discrimination *and* cognitive conflict resolution should improve overall performance in the WIT.

MPT-RT implementations of WIT process models such as the DIM provide a number of advantages that can guide future research. For example, participants of different social groups or participants working under different experimental conditions can be compared not only with respect to their cognitive process parameters but also with respect to the relative speed of automatic and controlled processing branches. Moreover, given the hierarchical nature of the MPT-RT approach proposed here, personality traits or cognitive abilities can be studied as external covariates and potential predictors of cognitive process and latency parameters. Both of these approaches are helpful not only for basic social cognition research but also for designing and evaluating training programs and interventions that counteract maladaptive effects of social stereotypes.

LIMITATIONS

One limitation when using MPT-RT models of WIT performance is the necessity to constrain model parameters to achieve model identifiability. Regarding

3. As pointed out by a reviewer, these results also align with the PCRM. From the perspective of the PCRM, individuals with better inhibitory capacities or higher prejudice-control motivation may simply be the ones who choose more often to preempt conflict by engaging in controlled processing.

latency parameters, models M_{PCRM} , M_{DIM} , and M_{GM} assume that relative latencies do not differ between the two automatic processing branches suggesting a *gun* or a *tool* response. Similarly, according to model M_{PCM} relative latencies do not differ between the two processing branches that represent conflicts between incongruent default responses and successful target discrimination. Since the models M_{PCRM} , M_{DIM} , M_{PCM} , and M_{GM} already have as many free parameters as there are independent response categories in the data (i.e., $df = 0$), they do not allow testing these underlying assumptions.

In a similar vein, the comparison of models $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ is based on auxiliary assumptions for latency parameter restrictions across conditions. The parsimonious model M_{Pars} showed acceptable model fit, but model selection indices DIC and WAIC do not unequivocally support M_{Pars} relative to the unrestricted model M_{Basic} . Hence, it should not be overlooked that the model comparison of the extended models $M_{\text{PCRM,ext}}$ and $M_{\text{DIM,ext}}$ is conditional on the auxiliary assumptions implemented in M_{Pars} .

Furthermore, $M_{\text{DIM,ext}}$ in Model Comparison 3 is endowed with the capacity to account for stereotype-congruency effects of correct response times by latency parameters whereas $M_{\text{PCRM,ext}}$ cannot. More specifically, $M_{\text{DIM,ext}}$ differentiates between latencies of controlled process branches resulting in congruent versus incongruent default responses whereas the $M_{\text{PCRM,ext}}$ does not have this feature. Alternatively, the stereotype-congruency effect in response times might be due to a variation in latencies of automatic process branches, meaning that the speed of stereotype-congruent automatic process branches is faster than the speed of incongruent automatic process branches. For example, the association of a Black face with a gun occurs faster than the association of a Black face with a tool (Klauer & Voss, 2008). This alternative view allows explaining stereotype-congruency effects based on the latency parameters of the automatic process branches. It can also account for the reversed congruency effect reported by Klauer and Voss (2008), that is, that errors for tool targets are faster if they followed Black face primes and errors for gun targets are faster if they followed White face primes.

Model comparisons implementing this alternative specification for M_{PCRM} and M_{DIM} (i.e., stereotype-congruent automatic process branches are faster than incongruent ones) are listed in the online supplemental materials as Model Comparisons 4 and 5 (<https://osf.io/7vjrq>). Importantly, these alternative model comparisons overall result in the same conclusion as reported above, namely, that the DIM is more closely aligned with the available data than the PCRM. Hence, the specific assumptions implemented in our Model Comparison 3 are not crucial for the conclusion that the DIM provides the best account of WIT response frequencies and response times when considered jointly.

Beyond the constraints on the latency parameters, the PDP itself has been criticized because of the assumption that the controlled process parameter C is invariant across target conditions. Payne et al. (2005) as well as Klauer and Voss (2008) critically discussed this invariance assumption. For example, a metal tube as part of a target object might more easily be interpreted as part of a gun after seeing a Black face prime and more easily as part of a screwdriver after seeing a

White face prime. Consequently, the prime would be expected to modulate the controlled process of discriminating the object (Klauer & Voss, 2008; Klauer et al., 2015; Payne et al., 2005). Experimental tests of the invariance assumption have resulted in mixed outcomes so far (Klauer et al., 2015; Payne et al., 2005). Based on an extended version of the PDP, Klauer et al. (2015) observed results inconsistent with the invariance assumption. More recently, however, Todd et al. (2021) analyzed the WIT using diffusion modeling as an alternative analytic approach to the PDP. Their diffusion analysis suggests that the discrimination process (as mirrored in the drift rate) was not influenced by an interaction between primes and targets, a result that is in line with the invariance assumption. Clearly, further research addressing the validity of the invariance assumption is needed.

ALTERNATIVE MODELING APPROACHES

As alternatives to MPT-RT modeling, other cognitive modeling approaches can be used to gain a deeper understanding of underlying processes. As already mentioned, Todd et al. (2021) used the diffusion model (DM), a specific version of evidence-accumulation models, to analyze WIT performance. The DM decomposes decisions into four components: an initial response bias towards guns or tools, the quality of information extracted from a stimulus (called drift rate), non-decision time (e.g., encoding and motor response time), and the amount of evidence required to make a decision (called threshold separation). In their DM analysis, they found that the initial response bias for Black face primes was closer to the gun response than for White face primes. The drift rate, however, did not vary as a function of prime race. This suggests that the prime biases the decision in the WIT from the outset of each trial, again a result that is in accordance with the DIM.

As an alternative to MPT-RT and DM approaches, RT-MPT models (Klauer & Kellen, 2018) could also be used to incorporate response times in multinomial models. RT-MPT models, in contrast to MPT-RT models, allow estimation of single-process latencies instead of relative latencies of entire process branches. To our knowledge, RT-MPTs have not been applied to the WIT so far. They could be a valuable alternative to gain more fine-grained insights into the processes underlying WIT performance. To conclude, future research on the WIT can benefit from using multiple modeling approaches to gain a deeper understanding of the underlying cognitive processes, to combine the strengths of the different approaches, and eventually achieve converging evidence across modeling approaches.

CONCLUSION

We reanalyzed eight data sets of the weapon identification task (WIT) to evaluate four process models of racial bias in weapon identification after formalizing them as response time-extended multinomial processing tree models. Our process model comparison clearly supported the default interventionist model (DIM) and the preemptive conflict-resolution model (PCRM). Both models share the assumption that responses based on automatic processes are faster than those based on

controlled target identification. Follow-up model comparisons revealed a preference for the DIM vis-à-vis the PCRM. According to the DIM, automatic associations elicit response bias from the outset of each trial. Therefore, controlled processes—beyond their importance for target identification—are required to resolve possible response conflicts whenever default responses deviate from the correct response.

In addition to their significance for basic social cognition research, our results are highly relevant also for applied fields such designing and implementing police training. Our results suggest that interventions that increase both the capacity for controlled processing to discriminate the target object *and* the ability to overcome default automatic responses increase the accuracy of weapon identification. In line with this conclusion, previous studies showed that increasing the capacity for controlled processing, for example, by providing sufficiently large response-time windows, by increasing cognitive resources, and by training to resolve conflicts induced by default responses (Govorun & Payne, 2006; Ito et al., 2015; Klauer & Voss, 2008; Kleiman et al., 2013; Lambert et al., 2003; Payne, 2001, 2005; Payne et al., 2002; Rivers, 2017) generally improves performance in the WIT. Under standard WIT conditions, however, people tend to be biased by race (or other social characteristics) early on in identifying innocuous or threatening objects. Nevertheless, by developing well-designed interventions, this bias likely can be overcome or at least reduced effectively. This is an important result, especially for bias awareness programs and police trainings that aim at reducing racial bias.

REFERENCES

- Amon, M. J., & Holden, J. G. (2016). Fractal scaling and implicit bias: A conceptual replication of Correll (2008). In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1553–1558). Cognitive Science Society. https://cognitivesciencesociety.org/wp-content/uploads/2019/03/cogsci2016_proceedings.pdf
- Bishara, A. J., & Payne, B. K. (2009). Multinomial process tree models of control and automaticity in weapon misidentification. *Journal of Experimental Social Psychology*, 45(3), 524–534. <https://doi.org/10.1016/j.jesp.2008.11.002>
- Burnham, K. P., & Anderson, D. R. (2004). Multi-model inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. <https://doi.org/10.1177/0049124104268644>
- 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>
- Correll, J. (2008). 1/f noise and effort on implicit measures of bias. *Journal of Personality and Social Psychology*, 94(1), 48–59. <https://doi.org/10.1037/0022-3514.94.1.48>
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, 15(12), 806–813. <https://doi.org/10.1111/j.0956-7976.2004.00760.x>
- Evans, J. St. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, 13(4), 321–339. <https://doi.org/10.1080/13546780601008825>
- Govorun, O., & Payne, B. K. (2006). Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, 24(2), 111–136. <https://doi.org/10.1521/soco.2006.24.2.111>
- 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>
- Huntsinger, J. R., Sinclair, S., & Clore, G. L. (2009). Affective regulation of implicitly measured stereotypes and attitudes: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 45(3), 560–566. <https://doi.org/10.1016/j.jesp.2009.01.007>
- Ito, T. A., Friedman, N. P., Bartholow, B. D., Correll, J., Loersch, C., Altamirano, L. J., & Miyake, A. (2015). Toward a comprehensive understanding of executive cognitive function in implicit racial bias. *Journal of Personality and Social Psychology*, 108(2), 187–218. <https://doi.org/10.1037/a0038557>
- Jones, C. R., & Fazio, R. H. (2010). Person categorization and automatic racial stereotyping effects on weapon identification. *Personality and Social Psychology Bulletin*, 36(8), 1073–1085. <https://doi.org/10.1177/0146167210375817>
- Kahn, K. B., & Martin, K. D. (2020). The social psychology of racially biased policing: Evidence-based policy responses. *Policy Insights from the Behavioral and Brain Sciences*, 7(2), 107–114. <https://doi.org/10.1177/2372732220943639>
- 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., Dittrich, K., Scholtes, C., & Voss, A. (2015). The invariance assumption in process-dissociation models: An evaluation across three domains. *Journal of Experimental Psychology: General*, 144(1), 198–221. <https://doi.org/10.1037/xge0000044>
- 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.
- 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>
- Kleiman, T., Hassin, R. R., & Trope, Y. (2014). The control-freak mind: Stereotypical biases are eliminated following conflict-activated cognitive control. *Journal of Experimental Psychology: General*, 143(2), 498–503. <https://doi.org/10.1037/a0033047>
- Knapp, B. R., & Batchelder, W. H. (2004). Representing parametric order constraints in multi-trial applications of multinomial processing tree models. *Journal of Mathematical Psychology*, 48(4), 215–229. <https://doi.org/10.1016/j.jmp.2004.03.002>
- Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R. (2003). Stereotypes as dominant responses: On the “social facilitation” of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, 84(2), 277–295. <https://doi.org/10.1037/0022-3514.84.2.277>
- Madurski, C., & LeBel, E. P. (2015). Making sense of the noise: Replication difficulties of Correll’s (2008) modulation of 1/f noise in a racial bias task. *Psychonomic Bulletin & Review*, 22(4), 1135–1141. <https://doi.org/10.3758/s13423-014-0757-4>
- 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. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology*, 89(4), 488–503. <https://doi.org/10.1037/0022-3514.89.4.488>
- Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, 38(4), 384–396. [https://doi.org/10.1016/S0022-1031\(02\)00006-9](https://doi.org/10.1016/S0022-1031(02)00006-9)

- Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. *Journal of Experimental Social Psychology*, 41(1), 36–47. <https://doi.org/10.1016/j.jesp.2004.05.001>
- Riefer, D. M., & Batchelder, W. H. (1991). Statistical inference for multinomial processing tree models. In J.-P. Doignon & J.-C. Falmagne (Eds.), *Mathematical psychology: Current developments* (pp. 313–335). Springer. https://doi.org/10.1007/978-1-4613-9728-1_18
- Rivers, A. M. (2017). The weapons identification task: Recommendations for adequately powered research. *PLoS One*, 12(6), e0177857. <https://doi.org/10.1371/journal.pone.0177857>
- Schmidt, O., Erdfelder, E., & Heck, D. W. (2023). How to develop, test, and extend multinomial processing tree models: A tutorial. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000561>
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. detection, search, and attention. *Psychological Review*, 84(1), 1–66. <https://doi.org/10.1037/0033-295X.84.1.1>
- Sherman, J. W. (2006). On building a better process model: It's not only how many, but which ones and by which means? *Psychological Inquiry*, 17(3), 173–184.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127–190. <https://doi.org/10.1037/0033-295X.84.2.127>
- 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>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3), 485–493. <https://doi.org/10.1111/rssb.12062>
- Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, 34(10), 1332–1345. <https://doi.org/10.1177/0146167208321269>
- Swencionis, J. K., & Goff, P. A. (2017). The psychological science of racial bias and policing. *Psychology, Public Policy, and Law*, 23(4), 398–409. <https://doi.org/10.1037/law0000130>
- Todd, A. R., Johnson, D. J., Lassetter, B., Neel, R., Simpson, A. J., & Cesario, J. (2021). Category salience and racial bias in weapon identification: A diffusion modeling approach. *Journal of Personality and Social Psychology*, 120(3), 672–693. <https://doi.org/10.1037/pspi0000279>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Volpert-Esmond, H. I., Scherer, L. D., & Bartholow, B. D. (2020). Dissociating automatic associations: Comparing two implicit measurements of race bias. *European Journal of Social Psychology*, 50(4), 876–888. <https://doi.org/10.1002/ejsp.2655>
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.