How multinomial processing trees have advanced, and can continue to advance, research using implicit measures

Jimmy Calanchini[1]

Manuscript accepted for publication in *Social Cognition* special issue on "25 Years of Research Using Implicit Measures"

Manuscript # 2019SOCOG000129R

[1]Department of Psychology, University of California, Riverside, Riverside, CA 92521

Correspondence to:

Jimmy Calanchini, Department of Psychology, University of California, Riverside

E-mail: jimmy.calanchini@ucr.edu

Word count: 6925

Abstract

Implicit measures were developed to provide relatively pure estimates of attitudes and stereotypes, free from the influence of processes that constrain true and accurate reporting. However, implicit measures are not pure estimates of attitudes or stereotypes but, instead, reflect the joint contribution of multiple processes. The fact that responses on implicit measures reflect multiple cognitive processes complicates both their interpretation and application. In this article, I highlight contributions made to research using implicit measures by multinomial processing trees (MPTs), an analytic method that quantifies the joint contributions of multiple cognitive processes to observed responses. I provide examples of how MPTs have helped resolve mysteries that have arisen over the years, examples of findings that were initially taken at face-value but were later re-interpreted by MPTs, and look to the future for ways in which MPTs seem poised to further advance research using implicit measures.

KEYWORDS: Implicit Measures, Implicit Bias, Multinomial Processing Trees, Dual-Process Models

ACKNOWLEDGMENTS: I would like to thank my unconscious because the outline for this manuscript came to me, fully formed, in a dream.

How multinomial processing trees have advanced, and can continue to advance, research using implicit measures

Implicit[1] measures were developed, in part, to provide relatively pure estimates of attitudes and stereotypes, free from the influence of processes that constrain true and accurate responses (e.g., Bargh, 1999; Fazio, Jackson, Dunton, & Williams, 1995). To do so, implicit measures implement task conditions assumed to make responses difficult to feign strategically (e.g., response deadlines; subliminal stimulus presentation). However, subsequent research revealed that implicit measures are not pure estimates of attitudes and stereotypes but, instead, reflect the joint contribution of multiple processes.

The fact that multiple cognitive processes influence responses on implicit measures raises important questions about social cognition. For example, if an intervention reduces implicit bias, how can we know whether the intervention changed cognitive process related to the attitude or something else? If unrelated implicit measures correlate, or related implicit measures do not correlate, is there a problem with our measures or with the construct we are attempting to measure? Questions like these pose problems for basic scientists and practitioners alike who seek to use implicit measures to advance theory or improve the human condition. In this article, I will highlight the role of multinomial processing trees (MPTs: Riefer & Batchelder, 1988) in answering these kinds of questions, and others, thereby helping to advance research using implicit measures.

Multinomial Processing Trees: Some Background

[1] I use the term 'implicit' to mean 'indirect' (Corneille & Hütter, 2020). Thus, an 'implicit measure' assesses mental contents indirectly, often based on the speed or accuracy of responses, rather than on the contents of responses, *per se.* Accordingly, I use the terms 'implicit bias' and 'implicit attitudes' synonymously to refer to behavioral responses indexed by an implicit measure.

MPTs are a class of methods that has been used extensively to quantify the contributions of multiple processes to responses on implicit measures. For those not already familiar with MPT modeling, it might be helpful to think of MPTs as analogous to the perhaps better-known method of Structural Equation Models (SEM), in that both SEMs and MPTs use observed responses to estimate the contributions of unobservable cognitive processes (Figure 1). The similarities between SEM and MPTs are largely conceptual: for example, the math underlying each method is very different. That said, I will not be delving into the math underlying MPTs in this article, so the SEM analogy might be a helpful scaffold for the ideas presented below.

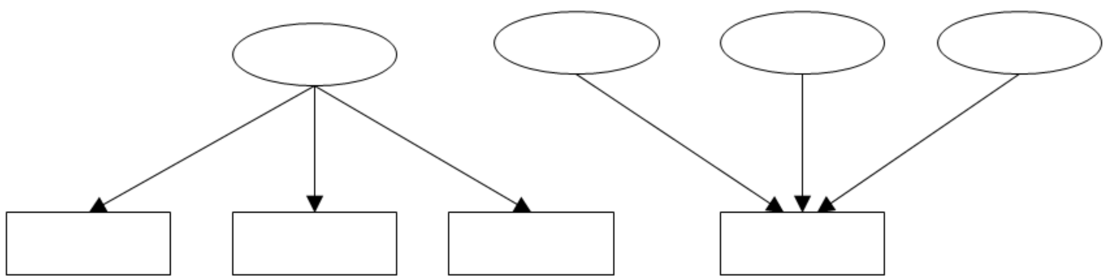

Figure 1. Simplified depictions of a structural equation model (SEM; left) and a multinomial processing tree (MPT; right). Rectangles reflect observed, manifest responses and ovals reflect unobserved, latent processes. SEM can estimate the contribution of a common latent process to multiple measures, whereas MPT can estimate the contribution of multiple latent processes to a single measure.

An MPT begins with a set of parameters that are hypothesized to represent the latent cognitive processes that influence discrete responses to a measure (e.g., correct/incorrect; yes/no). The parameters are operationalized on a probability scale, ranging from 0 to 1, such that the values reflect the likelihood of the cognitive process taking place. The relationships among the parameters can be represented in a tree-like structure that specifies how sequences of these processes (“branches”) jointly produce responses (“leaves”) to all possible conditions in the task, e.g., correct (✓) and incorrect (✕) responses (Figures 2-8). Parameters with lines leading to them are conditional upon all preceding parameters. The overall probability of each response can be represented as the sum of the products of the parameters (or their complements) in all of the branches that lead to that response. These probabilities form a system of equations that serves as the basis for parameter estimation: Participants' observed responses are entered as outcomes to the equations, and parameter values are estimated (e.g., through closed form, expectation maximization, or Markov Chain Monte Carlo methods) that most closely reproduce the observed responses. The degree to which the outcomes predicted by the parameter estimates align with participants' responses can be quantified using goodness-of-fit statistics, such as χ^2 or G^2, and the parameters interpreted as estimates of the influence of the hypothesized cognitive processes.

The manner in which the parameters of an MPT are specified to produce responses on an implicit measure is articulated based on theory. In this way, MPTs reflect a formalized statistical embodiment of theoretical assumptions about implicit measures. Consequently, an MPT can only provide meaningful results if the theory is correct, and model fit statistics provide a quantitative metric to gauge the accuracy of the theory underlying an MPT. Poor fit indicates that the theoretical assumptions operationalized in the equations of the model are insufficient to describe the observed outcomes. Thus, an MPT is not valid if it cannot provide adequate fit to data. However, model fit is only one of many steps in the development of an MPT. Additional evidence is necessary to demonstrate a relationship between each parameter and the psychological process it is assumed to reflect. Model developers must also conduct a series of selective-influence experiments to demonstrate the conceptual, convergent, discriminant, and predictive validity of each parameter in an MPT. For example, if a parameter is assumed to reflect evaluations of a racial group, then it should be sensitive to variations in the phenotypicality of the exemplars used to represent that racial group. MPTs have been validated on many implicit measures, such as the implicit association test (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Meissner & Rothermund, 2013), weapons identification task (Payne, 2001), affect misattribution procedure (Payne, Hall, Cameron, & Bishara, 2010), stereotype misperception task (Krieglmeyer & Sherman, 2012), go/no-go association task (Nadarevic & Erdfelder, 2011), and extrinsic affective simon task (Stahl & Degner, 2007).

The purpose of this article is not to provide a comprehensive review of the published literature of MPTs that have been applied to implicit measures. For that, see excellent reviews by Payne and Bishara (2009), and Hütter and Klauer (2016). Instead, the purpose of this article is to provide a brief summary of some of the contributions that MPTs have made to the literature on implicit measures. This summary is organized into three sections. The first provides examples of how MPTs have helped to resolve mysteries that have arisen over the years in research using implicit measures. The second section highlights findings in the literature that were initially taken at face-value – non-mysteries, as it were – but were later re-interpreted by MPTs. And finally, the third section discusses some areas in which MPTs are, in my view, poised to advance research using implicit measures.

Resolved Mysteries

Implicit Bias Change

When implicit measures were first developed, initial theorizing assumed that implicit biases operate automatically and invariantly and, thus, are largely impervious to control or change (e.g., Bargh, 1994; Devine, 1989; Dunton & Fazio, 1997; Fiske, 1998). However, evidence quickly emerged that implicit biases could be changed in a wide variety of ways (Blair, 2002). Because responses on implicit measures were largely assumed to primarily reflect associations[2] stored in memory (e.g., Fazio & Towles-Schwen, 1999), association-focused explanations for implicit bias change abounded (e.g., Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Smith & Zárate, 1992; Wilson, Lindsay, & Schooler, 2000). That said, other perspectives were beginning to suggest that responses on implicit measures did not reflect the unfettered expression of associations, as initially assumed. Instead, processes such as automatized egalitarian goals were proposed as candidates to constrain the expression of associations on implicit measures (e.g., Kawakami et al., 2000; Monteith, 1993; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). However, summary statistics of performance on an implicit measure, such as the *d*-score (Greenwald, Banaji, & Nosek, 2003) for the implicit association test (IAT: Greenwald, McGhee, & Schwarz, 1998), cannot distinguish between associative versus control-oriented explanations for implicit bias change. Changes in summary statistics indicate that *something* has changed, but provide no information on what *specifically* has changed. Consequently, the mystery of the processes underlying implicit bias change largely remained moot without a tool to investigate it.

Because MPTs quantify the joint contributions of multiple processes to responses, they are well-positioned to investigate the processes underlying implicit bias change. And they have. For example, implicit racial bias can be reduced by exposure to counter-stereotypical people (e.g., White serial killer Jeffrey Dahmer; Black actor Denzel Washington: Dasgupta & Greenwald, 2001), and MPT analysis using the Quad model (Figure 2) revealed that this change is mediated solely by an associative process (i.e., Association Activation: Gonsalkorale, Allen,

[2] Early association-focused theories shaped the terminology of the literature on implicit measures. In this manuscript I retain this terminology for the sale of clarity and continuity, and use the term "association" to refer to one of the mental constructs assessed by implicit measures. However, I make no strong assumptions about the representational nature – associative or otherwise – of the constructs assessed by implicit measures.

Sherman, & Klauer, 2010). Similarly, implicit preference for White relative to Black people can be reduced by portraying Black people in positive versus negative contexts (e.g., church versus ghetto: Wittenbrink, Judd, & Park, 2001), and the Quad model indicated that this change is mediated solely by a control process (i.e., Overcoming Bias: Allen, Sherman, & Klauer, 2010). And finally, implicit racial bias can be reduced through repeated practice affirming stereotype-incongruent trait pairings (e.g., Black-smart; White-violent: Kawakami et al., 2000), and the Quad model found that this change is mediated by both associative and control processes (i.e., Association Activation and Detection: Calanchini, Gonsalkorale, Sherman, & Klauer, 2013).

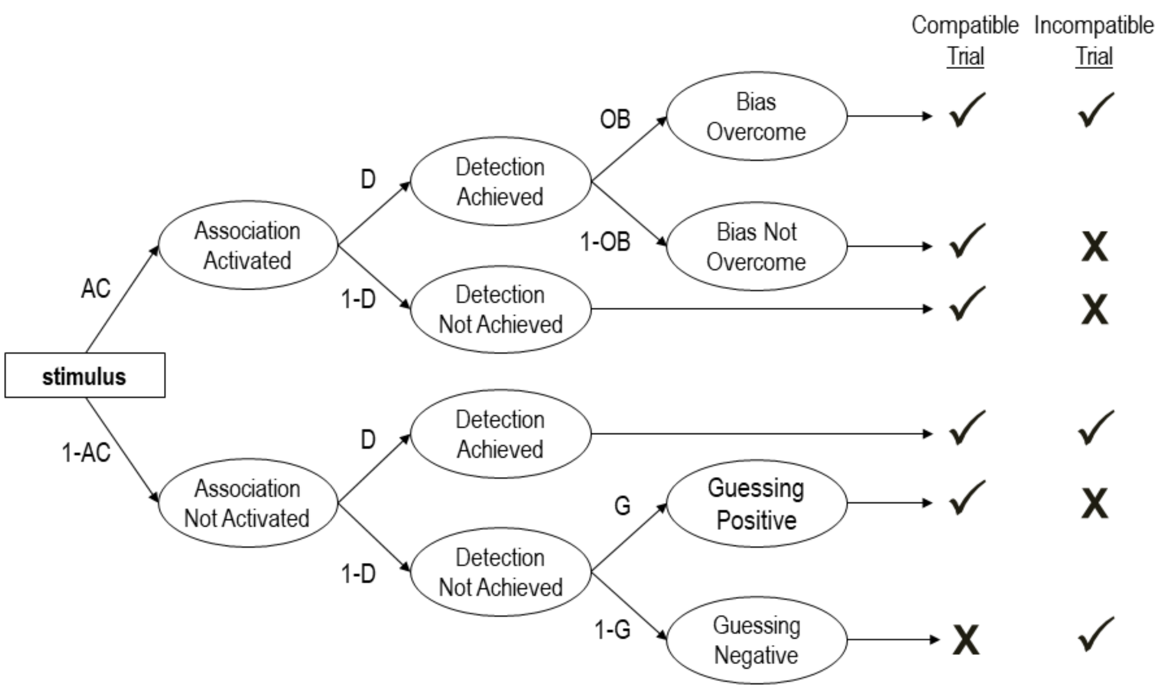

Figure 2. The quadruple process model (Quad model: Conrey et al., 2005) posits the influence of four qualitatively distinct processes. *Association Activation* (AC) reflects evaluations activated by the stimuli. *Detection* (D) reflects accurate responding to stimuli. *Overcoming Bias* (OB) reflects an inhibitory process that determines whether AC or D drives a response. *Guessing* (G) reflects any other processes that drive responses in the absence of influence from AC, D, and OB.

Taken together, MPTs helped to resolve the mystery of the processes underlying implicit bias change: Sometimes change reflects changes in associations; other times it reflects changes in control; and yet other times it reflects changes in both associations and control. That said, a more homogenous pattern of results (e.g., implicit bias change consistently reflecting changes in the same cognitive process) would seem more parsimonious. However, this disparate pattern of results is perhaps unsurprising, given the heterogeneity among the hundreds of implicit bias-reduction interventions developed over the years (Lai, Hoffman, & Nosek, 2013). In an effort to synthesize these different patterns of results, my colleagues and I (Calanchini, Lai, & Klauer, 2020) applied the Quad model to IAT data reflecting 18 different bias-reduction interventions. Using a taxonomy of intervention categories based on shared procedural and conceptual features, we meta-analyzed across categories and found that interventions based on the principles of evaluative conditioning influenced control processes (i.e., Overcoming Bias, Detection)[3], whereas interventions that relied on counterstereotypic exemplars or strategies to override biases influenced both associations and control processes (i.e., Association Activation, Detection). This taxonomy-based approach lays the foundation for an important advancement in research using implicit measures that will take some of the mystery out of implicit bias change: rather than researchers applying MPTs retroactively (and, perhaps, speculatively) to find out which cognitive process(es) were influenced by an intervention, this taxonomy may help researchers to more precisely develop interventions with the *a priori* intent of targeting specific cognitive processes.

Correlations Among Unrelated Implicit Measures

Another mystery that emerged relatively soon after implicit measures were first developed is that unrelated implicit measures sometimes correlate. For example, when implicit bias is operationalized in terms of latency-based summary statistics, an IAT configured to

[3] See also Hütter, Sweldens, Stahl, Unkelbach, and Klauer (2012) for further evidence from MPTs of the role of control processes in evaluative conditioning.

measure attitudes towards Black versus White people correlates with an IAT configured to measure attitudes towards things that are delicious versus happy (McFarland & Crouch, 2002; see also Mierke & Klauer, 2003). Improvements to the way in which latency-based summary statistics are calculated (Greenwald et al., 2003) reduced, but did not eliminate, the correlation between unrelated IATs (Cai, Sriram, Greenwald, & McFarland, 2004). Consequently, evidence of relationships between IATs configured to assess unrelated constructs indicated that something attitude-unrelated influences responses on IATs which, in turn, poses a problem for interpreting the magnitude of implicit bias summary statistics to reflect the strength of the attitude (e.g., Banaji, 2001; see also Mierke & Klauer, 2001).

MPTs can be used to disentangle the contributions of attitude-related versus -unrelated processes to responses on implicit measures. Most of the MPTs that have been applied to implicit measures share in common the dual-process framework of automaticity and control (Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Generally, the automatic processes reflected in MPTs are assumed to be attitude-related, whereas the control processes are assumed to be attitude-unrelated. However, this assumption conflates the circumstances under which the process can exert influence (i.e., its operating conditions) with the qualitative nature of the process (i.e., its operating principles): An automatic process is not necessarily attitude-related, and vice versa (Gawronski & Bodenhausen, 2009). To the extent that an MPT parameter reflects an attitude-related process, it should correlate across IATs that share conceptual content, but not correlate across dissimilar IATs. In contrast, to the extent that an MPT parameter reflects an attitude-unrelated (i.e., domain-general) process, it should correlate invariantly across IATs, regardless of conceptual overlap.

Like other MPTs for implicit measures, the Quad model is based on this dual-process theoretical framework. However, unlike other MPTs (i.e., Meissner & Rothermund, 2013; Nadarevic & Erdfelder, 2011; Payne, 2001; Payne et al. 2010; Stahl & Degner, 2007), the Quad model includes two parameters that are conceptualized to reflect qualitatively distinct automatic processes (i.e., Association Activation, Guessing), and two parameters that are conceptualized to reflect qualitatively distinct control processes (i.e., Detection, Overcoming Bias). Thus, the Quad model provides two opportunities each to examine the extent to which automatic versus control processes are attitude-related versus domain-general. Given that Association Activation is assumed to reflect valenced associations activated by specific stimuli (e.g., pictures of the ingroup activating positive associations), it should be expected to reflect an attitude-related process. However, even though Guessing is also conceptualized to reflect an automatic process, it could plausibly be either attitude-related or domain-general. Guessing is operationalized as the tendency to rely on one response over the other (e.g., pleasant versus unpleasant) in the absence of other guides to response. Such a response bias could be general (e.g., an optimistic outlook), but might also be specific (e.g., "I like people but hate spiders."). Both of the control processes would also seem to be equivocal in terms of their attitude-specificity versus domain-generality. One control process, Detection, is conceptualized as the likelihood of accurately determining the correct response, which could reflect a general skill (e.g., executive function), but could instead reflect specific expertise (e.g., a florist responding more accurately on a flower/insect IAT than on a race IAT). The other control process, Overcoming Bias, is conceptualized as the likelihood that a behavioral response produced by Association Activation that would lead to an incorrect response is replaced with a contextually correct response. Thus, Overcoming Bias could also reflect a general skill (e.g., executive function), or it could reflect specific motivations (e.g., egalitarianism). Consequently, even though previous evidence strongly suggested that *some* domain-general processes influence IAT responses (McFarland & Crouch, 2002; Mierke & Klauer, 2001, 2003), which processes specifically are domain-general remained an open question.

With this question in mind, my colleagues and I (Calanchini, Sherman, Klauer, & Lai, 2014) had participants complete pairs of IATs that varied in conceptual overlap. Quad model Association Activation parameters correlated most strongly across IATs that shared high conceptual overlap, and these correlations decreased as conceptual overlap decreased. Guessing parameters demonstrated the same pattern of results, which suggests that both Association Activation and Guessing are attitude-specific processes. In contrast, both Detection and Overcoming Bias correlated strongly with themselves across IATs, regardless of conceptual overlap, which suggests that these are domain-general processes. These findings illustrate how MPTs helped to resolve the mystery of correlations between unrelated implicit measures by quantifying the components of IAT responses that are attitude-related versus domain-general. Moreover, MPTs extended previous research on the contributions of attitude-unrelated processes to responses on implicit measures (Mierke & Klauer, 2001; 2003) by shedding light on the qualitative nature of these processes (i.e., accuracy-orientation; inhibition).

Lack of Correlations Between Related Implicit Measures

A third mystery that emerged relatively early in the history of implicit measures is that related implicit measures sometimes do not correlate. Though the IAT is arguably the flagship implicit measure, many other measures have been developed over the years: the evaluative priming procedure (EPP: Fazio et al., 1995), weapons identification task (WIT: Payne, 2001), affect misattribution procedure (AMP: Payne, Cheng, Govorun, & Stewart, 2005), go/no-go association task (GNAT: Nosek & Banaji, 2001), and extrinsic affective simon task (EAST: DeHouwer, 2003). To the extent that different implicit measures are configured to assess the same construct (e.g., racial bias), and to the extent that implicit measures are assumed to be relatively pure measures of that construct, then the measures should be expected to correlate strongly. But, sometimes, they do not (e.g., Bar-Anan & Nosek, 2014; Bosson, Swann, & Pennebaker, 2000; Nosek & Banaji, 2001).

Procedural demands determine which processes influence responses on an implicit measure (Payne, Burkley, & Stokes, 2008). Consequently, method-specific variance may obscure correspondence between different implicit measures configured to assess the same construct. Payne (2005) provides an example of using MPTs to separate construct and method variance. Participants completed an EPP and a WIT, both configured to assess racial bias. Summary statistics of performance bias (i.e., error rates) correlated poorly between the EPP and WIT. However, MPT analysis using the process-dissociation model (PD model: Payne, 2001; Figure 3) revealed a significant relationship between measures on the parameter assumed to reflect racial attitudes (i.e., Automatic). Thus, the apparent lack of correlation between related implicit measures turned out to be a problem of signal versus noise, and MPTs helped to resolve this mystery by isolating the process of interest (i.e., racial attitudes) from other processes (i.e., method variance).

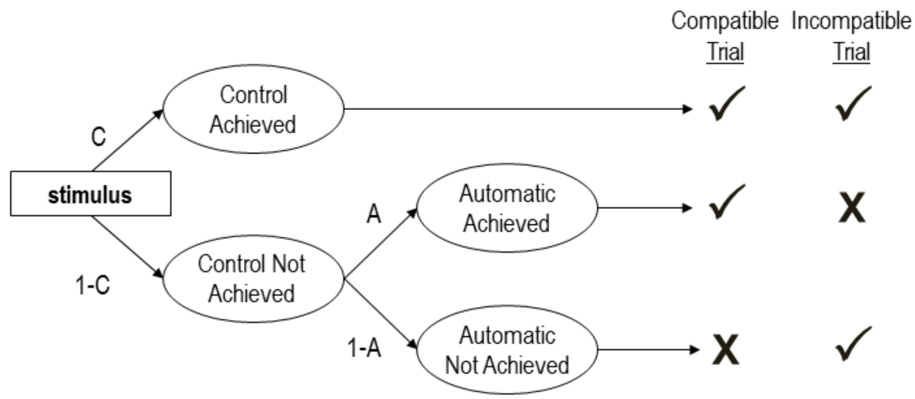

Figure 3. The process dissociation model (PD model: Payne, 2001) posits the influence of two qualitatively distinct processes. *Control* (C) reflects intentional responding in accordance with task instructions. *Automatic* (A) reflects associations activated by the stimuli.

Seemingly Nonsensical Implicit Biases

In one of the seminal demonstrations of implicit bias, participants responded to “flower + pleasant” and “insect + unpleasant” IAT pairings more quickly than they did to “flower + unpleasant” and “insect + pleasant” pairings (Greenwald et al., 1998). Greenwald and colleagues selected flowers and insects as stimuli because they assumed that participants evaluate them relatively uniformly – most people like flowers and dislike insects – which dovetails with an interpretation of the observed pattern of results to reflect an overall evaluative preference for flowers over insects. However, Brendl, Markman, and Messner (2001) threw a wrench into this otherwise straightforward example by replacing flowers with neutral non-words (e.g., “Larnist”). Surprisingly, participants responded to “insect + pleasant” and “non-word + unpleasant” IAT pairings more quickly than they did to “insect + unpleasant” and “non-word + pleasant” pairings. The evaluative preference for insects over nonwords suggested by this pattern of results is a mystery, given that participants are assumed to hold pre-existing negative evaluations of insects but no pre-existing evaluations (positive or negative) of novel non-words.

Given that competition between evaluative associations alone does sufficiently account for an implicit bias in favor of insects over non-words, Meissner and Rothermund (2015) proposed that participants were employing a task-simplifying strategy by which the double categorization task of the IAT is recoded into a single categorization task based on one feature shared among stimuli (Rothermund & Wentura, 2004). For example, on a flower/insect IAT, responses could be recoded into “things I like” and “things I don’t like” – but only on trials in which “flower + pleasant” (i.e., “things I like”) and “insect + unpleasant” (i.e., “things I don’t like”) share response keys. Thus, speedier responses in one IAT block versus the other does not only (or necessarily) reflect differences in underlying evaluations but, instead, the asymmetrical influence of recoding that simplifies responses on some trials but not others.

In the context of the insect/non-word IAT, Meissner and Rothermund (2015) proposed that participants employ recoding based on salience asymmetries among the response categories. Specifically, non-words are unfamiliar and, thus, are more salient than known words. Similarly, negative words are more salient than positive words (Rothermund & Wentura, 2004). Because non-words and negative words are more salient than insects and positive words, responses on “non-word + negative” and “insect + positive” trials can be recoded as “more salient” and “less salient”. To test this hypothesis, Meissner and Rothermund (2015) applied the ReAL model (Meissner & Rothermund, 2013: Figure 4) to participants’ responses on an insect/non-word IAT. As expected, the Recoding parameter differed significantly from zero, indicating the influence of this task-simplification strategy. Additionally, and importantly, the evaluation parameters (i.e., Associations) were negative for insects but neutral for non-words. Thus, MPTs helped to resolve a nonsensical implicit bias suggesting preference for insects over non-words: by separating out the influence of an additional process, the MPT revealed a theoretically-consistent pattern of evaluations.

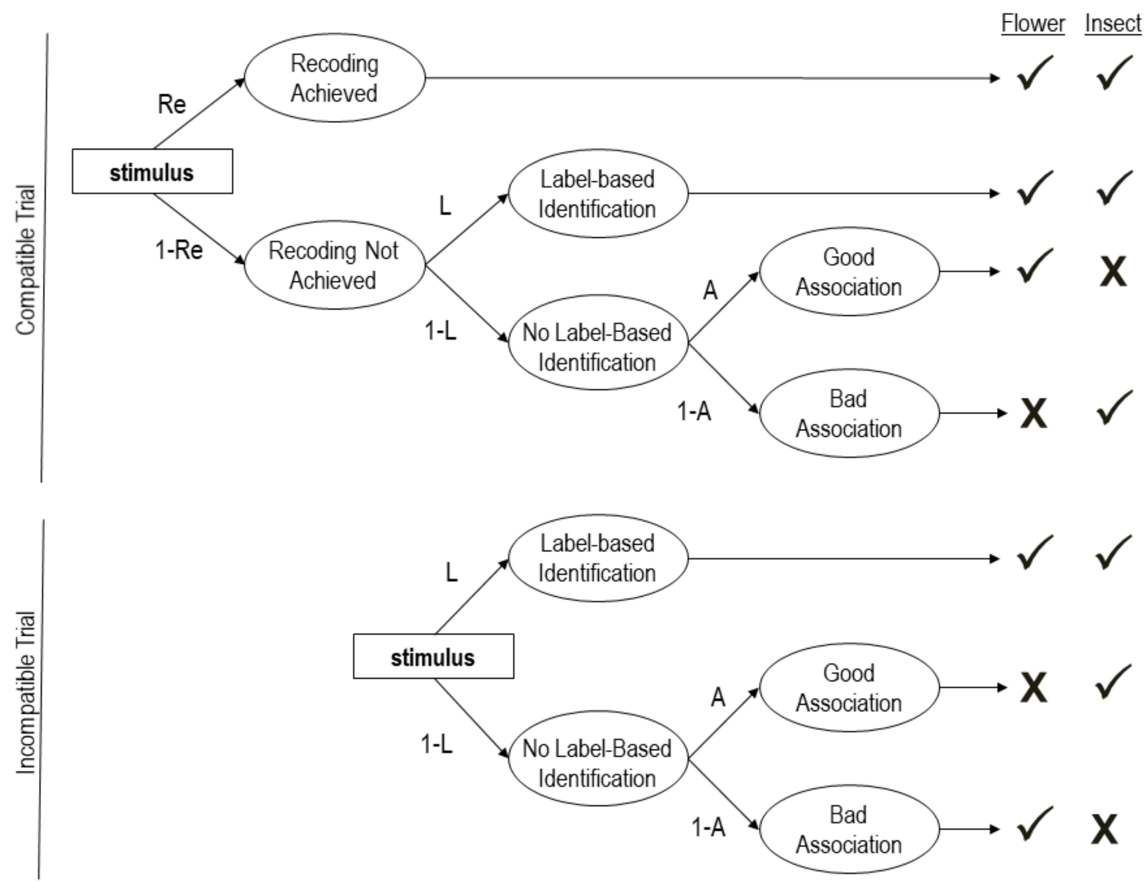

Figure 4. The ReAL model (Meissner & Rothermund, 2013) depicted for an evaluative IAT with flowers and insects as target categories. The ReAL model posits the influence of three qualitatively distinct processes. *Recoding* (Re) reflects a task simplification strategy by which the target and attribute categories are collapsed into a single category based on a common feature. *Label-based identification* (L) reflects the controlled search for a stimulus' category label and the response key to which it is assigned. *Evaluative associations* (A) reflect evaluations activated by the target stimuli.

The Mystery of Multiple MPTs

Given that multiple MPTs have been validated on implicit measures, the interested researcher might reasonably wonder which model is "right". The first, and most important, answer is: It depends on your research question. For example, the Quad model is the way to go if you are interested in inhibition, whereas the ReAL model is best if you are interested in recoding.

In other words, a validated model that includes parameters reflecting the processes that are relevant to your research is the right model for you. That said, many models for implicit measures have parameters in common, so more than one MPT might be suited to investigate your process(es) of interest. When faced with the choice between multiple models, the question of which model is best for a given set of data can be addressed quantitatively, through established model selection procedures.

The extent to which a given model fits data can be evaluated using familiar statistics such as χ^2 and G^2, which quantify the extent to which the outcomes predicted by a model align with the observed data. Consequently, competing models can be fit to the same data and the best one selected based on goodness of fit (i.e., the model with the lower χ^2 or G^2). However, a model with more parameters is more flexible and, all else equal, will provide better fit to data than a model with fewer parameters. Because χ^2 and G^2 do not account for the number of parameters in a model, they are insufficient for selecting between models with different numbers of parameters. For example, the PD model includes two parameters, whereas the ABC model (Stahl & Degner, 2007: Figure 5) includes three. Readers who are familiar with structural equation modeling may already know the model selection indices Akaike information criterion (AIC: Akaike, 1973) and Bayesian information criterion (BIC: Schwarz, 1978), which impose fit penalties based on the number of parameters in the model. That said, two models with the same number of parameters can still vary in equation complexity. For example, both the ABC model and the Trip model (Nadarevic & Erdfelder, 2011: Figure 6) include three parameters, but the equations of the Trip model are more complicated than those of the ABC model. A model with more complex equations is more flexible and, all else equal, will provide better fit to data than one with less complex equations. Minimum description length (MDL: Wu, Myung, &

Batchelder, 2010) is a model selection index that imposes fit penalties for both number of parameters and equation complexity. Thus, the decision between competing models should be guided, in part, by Occam's razor: Controlling for complexity, choose the model that provides best fit to the data.

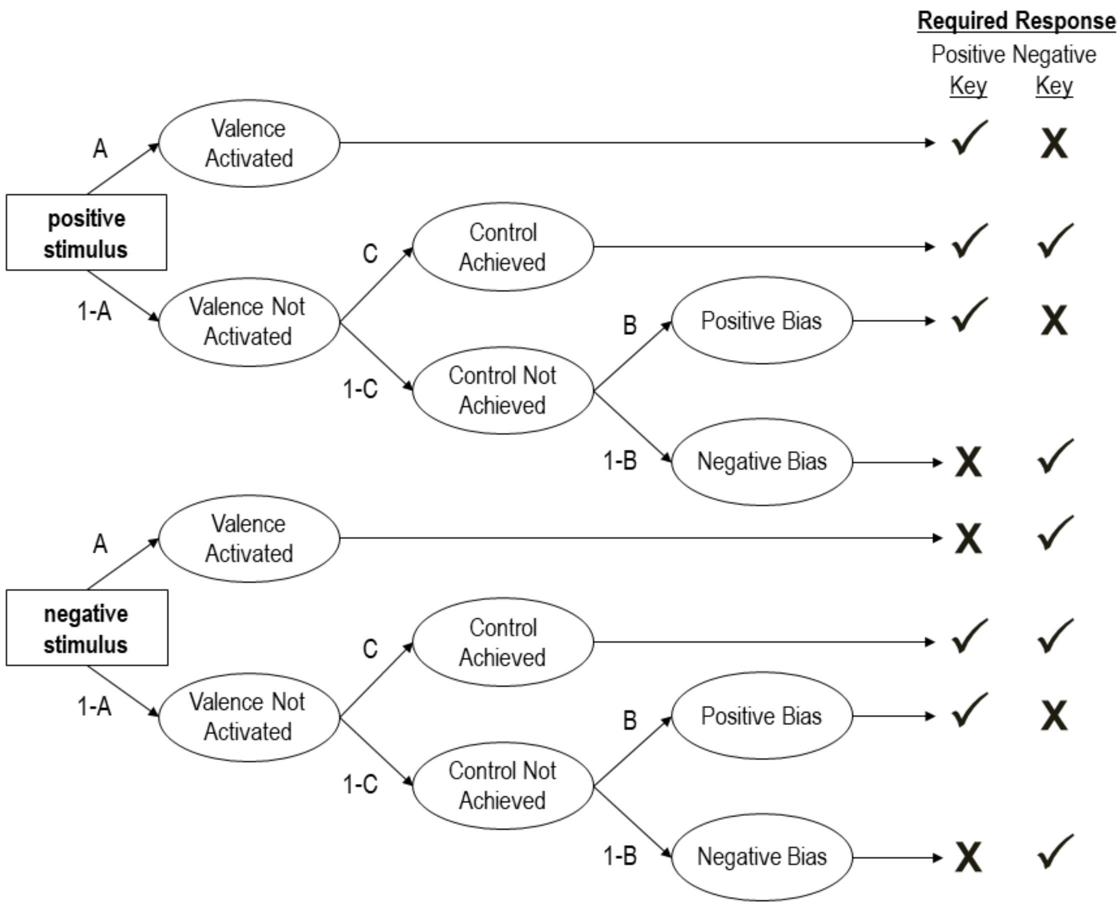

Figure 5. The ABC model (Stahl & Degner, 2007) posits the influence of three qualitatively distinct processes to the EAST (DeHouwer, 2003). *Control* (C) reflects controlled processing of task-relevant stimulus features. *Automatic* (A) reflects valence activated by the stimuli. *Bias* (B) reflects guessing when neither C nor A influence the response.

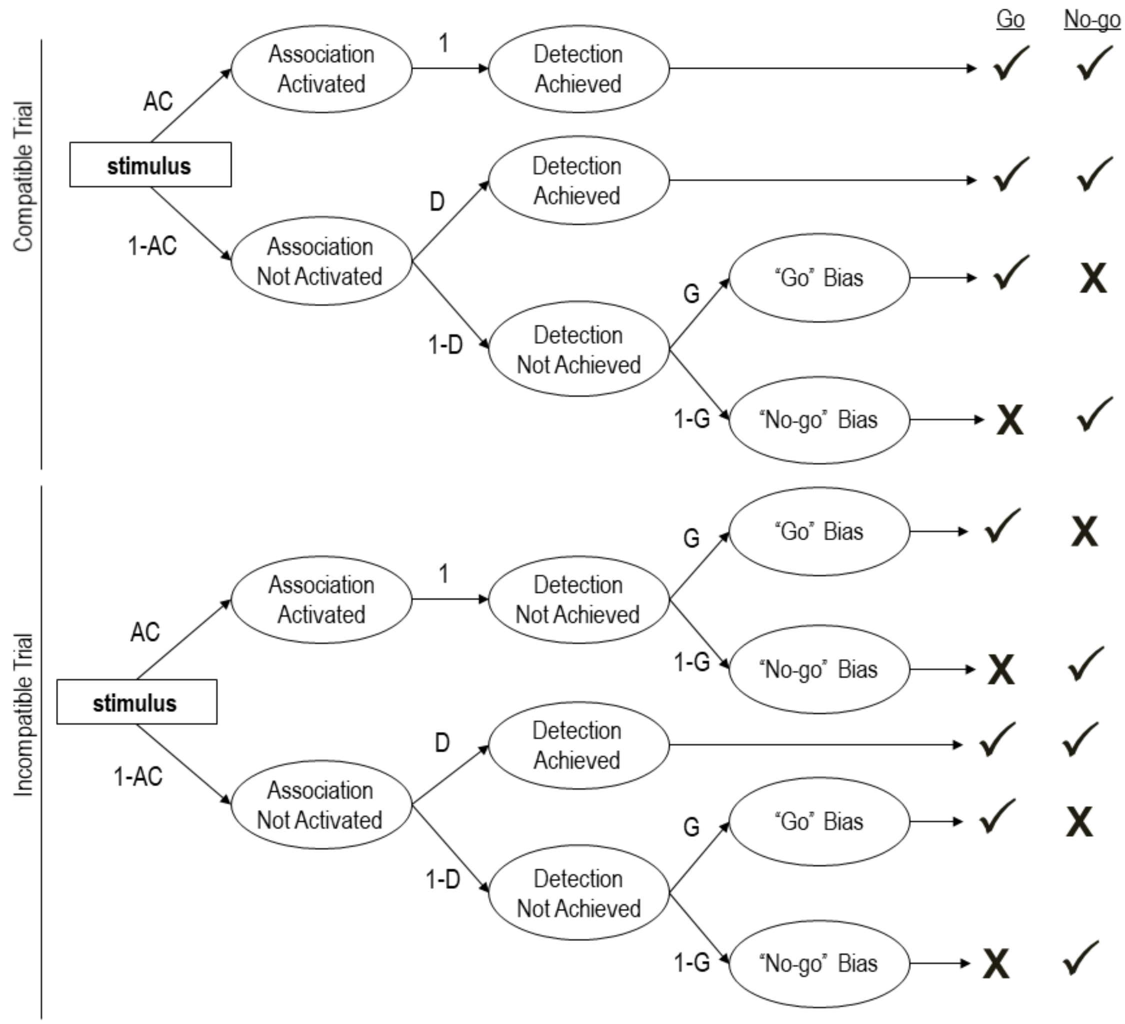

Figure 6. The Trip model (Nadarevic & Erdfelder, 2011) posits the influence of three qualitatively distinct processes to the GNAT (Nosek & Banaji, 2001). *Association Activation* (AC) reflects evaluations activated by the target stimuli. *Detection* (D) reflects accurate responding to stimuli. *Guessing* (G) reflects response biases that always determine responses when D is not achieved.

Selecting the one true model. Many MPTs for implicit measures share a common dual-process structure and are, thus, nested versions of one another. The PD model is the most basic of these MPTs, with two parameters reflecting one automatic and one control process. The ABC and Trip models build upon the PD model by including a third parameter reflecting a response bias or guessing process. And finally, the Quad model adds a second control parameter to the three parameters (i.e., automatic, control, and guessing) reflected in the ABC and Trip models.

Though the exact specifications of the shared parameters are not identical across models, the conceptual meaning of the automatic, control, and guessing parameters is consistent across these models. Thus, interested researchers can apply nested models to the same data to examine whether the model that includes a certain parameter provides better fit than the model that does not include that parameter.

Nadarevic and Erdfelder (2011) adopted this nested-model approach to evaluate the Trip versus Quad models in the context of both the IAT and the GNAT. In two studies, both the Trip and Quad models were fit to data from participants who completed an IAT and a GNAT. The Trip model consistently provided better fit to GNAT data, whereas the Quad model consistently provided better fit to IAT data. Given that the primary difference between the Trip and Quad models is that the latter includes a parameter reflecting inhibitory control (i.e., Overcoming Bias), this pattern of results indicates that inhibition of this kind plays a larger role in IAT responses than it does GNAT responses. Bishara and Payne (2009) and Sherman et al. (2008) report similar model selection exercises in the context of the IAT and sequential-priming type tasks, and converge on the conclusion that the PD model provides better fit to data from sequential priming-type tasks, whereas the Quad model provides better fit to data from the IAT. Taken together, these results not only highlight a fruitful approach to determine which model is best suited for a given implicit measure, but also underscore the broader point that task procedures facilitate versus constrain the influence of different cognitive processes (Payne et al., 2008).

Whereas Nadarevic and Erdfelder (2011), Bishara and Payne (2009), and Sherman et al. (2008) all compared two specific models against each other, Rees, Rivers, and Sherman (2019) adopted a multiverse approach to model selection (see also Rivers, Sherman, Rees, Reichardt, &

Klauer, 2020). Rees and colleagues' (2019) participants completed a sequential priming task designed to assess stereotypical judgments of Black versus White people (i.e., the Stereotype Misperception Task: SMT; Krieglmeyer & Sherman, 2012). In addition to applying the MPT that was developed in conjunction with the SMT (i.e., the SMT model: Figure 7), Rees and colleagues also applied six other MPTs to the same data: the PD model; a variant of the PD model that includes a guessing parameter; a variant of the PD model in which the automatic rather than the control process is dominant (i.e., the "Stroop" model: Lindsay & Jacoby, 1994); the automatic-dominant PD model with the addition of a guessing parameter (which corresponds to the ABC model); the AMP model (Payne et al., 2010: Figure 8); and an alternative specification of the SMT model in which the control rather than the automatic process is dominant. The model selection indices AIC, BIC, and MDL illustrate that the SMT model consistently provides better fit to these data than all other models, with the exception of the AMP model. In the case of the SMT versus AMP model, the evidence is relatively more mixed. In two of four experiments, the SMT model provides better fit than the AMP model on all three indices, but in one experiment the SMT model provides better fit in terms of BIC but the AMP model provides better fit in terms of AIC and MDL, and in another experiment the SMT provides better fit in terms of BIC and MDL but the AMP model provides better fit in terms of AIC. Taken together, this pattern of results highlights the superior performance of the SMT versus other models in fitting these data, but at the same time demonstrates the extent of heterogeneity among these comparisons.

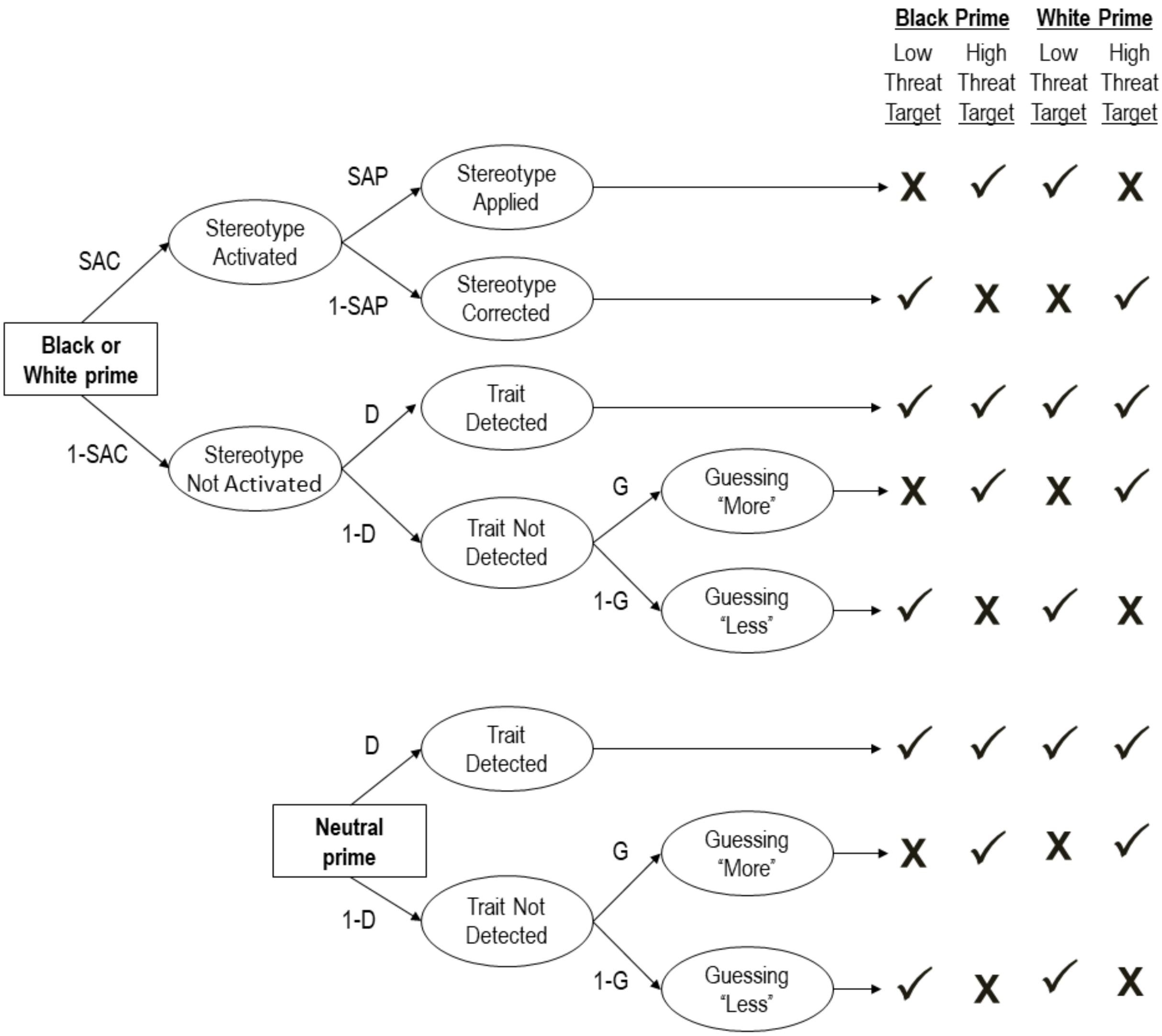

Figure 7. The SMT model (Krieglmeyer & Sherman, 2012) depicted for an SMT configured to assess racial threat stereotypes. The SMT model posits the influence of four qualitatively distinct processes. *Stereotype Activation* (SAC) reflects stereotypes activated by the prime stimuli. *Stereotype Application* (SAP) reflects the application of the activated stereotype to the response. *Detection* (D) reflects accurate responding to the target stimuli. *Guessing* (G) reflects any other processes that drive responses in the absence of influence from SAC, SAP, and D.

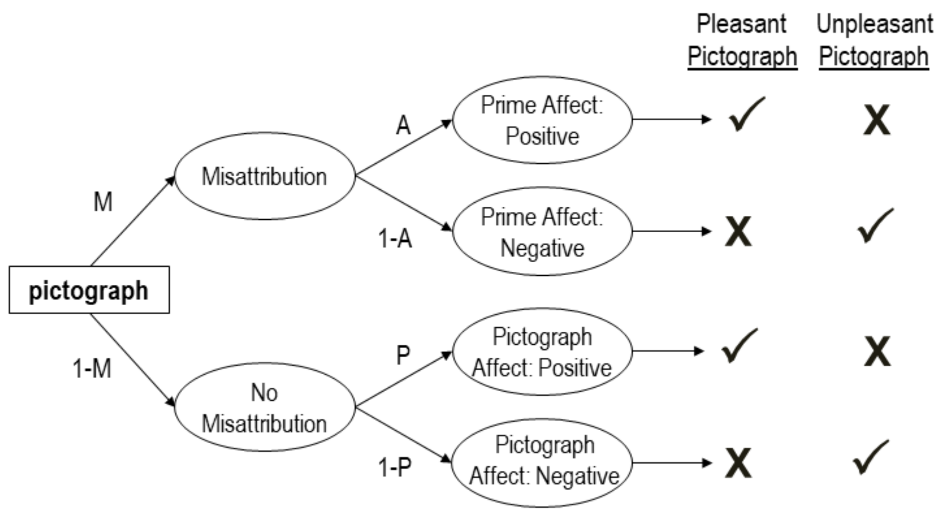

Figure 8. The AMP model (Payne et al., 2010) posits the influence of three qualitatively distinct processes. *Affect* (A) reflects an affective response to the prime stimuli. *Pictograph* (P) reflects an affective response to the target stimuli. *Misattribution* (M) reflects confusing an affective response to the prime stimuli for an affective response to the target stimuli.

Why not both?

As the above examples illustrate, when multiple MPTs have been applied in previous research, the goal is usually to determine which model provides better fit to data from a specific implicit measure. However, an alternate approach is to use multiple implicit measures within the same line of research, along with the MPT that is "right" for each measure (in terms of previous validation, etc.) in search of converging evidence across methods and models. For example, implicit evaluations can be reversed with exposure to a single impression-inconsistent behavior, and my colleague and I (Cone & Calanchini, 2020) were interested in the cognitive mechanism(s) underlying this rapid implicit revision. From a traditional dual-process perspective, rapid implicit revision should be expected to reflect control processes because associations are assumed to form slowly and change slowly (e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007; Smith & DeCoster, 2000; Wilson et al., 2000). Over a series of experiments, participants completed a learning paradigm in which they were exposed to multiple pieces of positive information about a target person, followed by either an additional neutral piece of information or a highly negative piece of information. Then, participants completed either an IAT or an AMP, and we applied the Quad or AMP model, respectively, to these data. Across measures, and in contrast to the traditional dual-process perspective, we reliably observed effects of the single piece of negative versus neutral information on the attitude-related parameters of each model (i.e., Association Activation and Affect, respectively), suggesting that associations can be changed quickly. This approach demonstrates that MPTs are not natural enemies, competing for parsimony or hegemony. Instead, different MPTs can be used in tandem to provide insight into implicit measures that generalizes beyond one specific model or measure.

Reinterpreted Non-Mysteries

Of Course Older People are Racist

Age differences in implicit racial bias are well documented: older people demonstrate more negative attitudes towards racial outgroups than do younger people (e.g., Nosek, Banaji, & Greenwald, 2002), in terms of latency-based summary statistics of implicit bias. One widely-accepted explanation for this pattern of results is that older people grew up when social norms were less egalitarian, and negative depictions of Black people were more common and more accepted than they are today (e.g., Wilson, 1996). However, MPT analysis tells a different story.

When the Quad model was applied to the responses of participants who completed a race IAT, racial Association Activation did not increase with age (Gonsalkorale, Sherman, & Klauer, 2009); instead, Association Activation remained relatively constant across age groups, even decreasing slightly with age. In contrast, Overcoming Bias decreased significantly starting around age 41-50. This pattern of results indicates that the relatively stronger racial bias demonstrated by older adults does not reflect more biased attitudes but, instead, reflects decreased ability to constrain the expression of those (more weakly) biased attitudes. Stewart, von Hippel, and Radvansky (2009) report very similar findings using the PD model, which suggests that the observed age-related decline is not an artifact of a specific MPT or operationalization of cognitive control. Thus, MPT modeling challenges a perfectly face-valid explanation (i.e., attitudes reflect the society in which we were raised) in favor of a more nuanced one (i.e., attitudes depend upon cognitive functioning).

Of Course Everybody is Ageist

Whereas implicit racial bias is positively correlated with age, implicit age bias does not vary by age: younger and older people alike demonstrate strong pro-young implicit attitudes in terms of latency-based summary statistics of implicit bias (Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002; Jost, Banaji, & Nosek, 2004; Nosek et al., 2002, 2007). This pattern of results has been interpreted to support system justification perspectives, such that an evaluative preference for youth legitimizes older people's disadvantaged status in society (Jost et al., 2004). But, once again, MPT analysis tells a different story.

The Quad model was applied to the responses of participants who completed an age IAT, which revealed lower age-related Association Activation among old versus young people, as well as lower Overcoming Bias (Gonsalkorale, Sherman, & Klauer, 2014). This pattern of results indicates that apparent age invariance in implicit age bias reflects older adults' weaker biased attitudes and decreased ability to constrain the expression of those (more weakly) biased attitudes. Here, MPT modeling highlights two countervailing effects underlying an apparent null effect.

When considered separately, Gonsalkorale and colleagues' (2009) / Stewart and colleagues' (2009) evidence of age-related differences in control in the context of racial bias, and

Gonsalkorale and colleagues' (2014) evidence of age-related differences in control in the context of age bias, could each be interpreted to reflect differences in motivation rather than differences in control, *per se*. That is, perhaps older people grew up during times that they were never taught to (or never needed to) inhibit their biased associations towards either Black or older people. However, two points speak against this interpretation. First, the motivational account is not parsimonious because it posits two domain-specific processes when one domain-general process can account for the observed outcome. Second, domain-specific (i.e., racial) egalitarian motivations are related to two parameters of the Quad model (i.e., Association Activation and Detection), but are unrelated to the parameter that (i.e., Overcoming Bias) varies across age groups (Gonsalkorale, Sherman, Allen, Klauer, & Amodio, 2011). Thus, converging evidence from multiple MPTs and multiple content domains speaks against accepted, but countervailing, motivational accounts for age differences in implicit racial bias but age invariance in implicit age bias, and in favor of a more parsimonious account of age-related deficits in general cognitive function (e.g., Connelly, Hasher, & Zacks, 1991; Hasher & Zacks, 1988).

Mysteries to Come

The future is necessarily more abstract than the past. Whereas the past is filled with concrete examples, the future consists of multiple alternative possibilities. So far in this article I have provided specific examples of how MPTs have helped to advance research using implicit measures. In this final section, I speculate more broadly about ways in which MPTs can make further contributions.

Other MPTs and measures. I have aimed to at least say a few words about all of the MPTs that have been applied to implicit measures of attitudes and stereotypes. However, I have said more words about some MPTs than others, and this largely reflects the extent to which they have been used. According to Google Scholar, at the time of this writing the PD model (Payne, 2001) has 1193 citations and the Quad model (Conrey et al., 2005) has 557 citations, but the AMP model (Payne et al., 2010) has 110 citations; the ReAL model (Meissner & Rothermund, 2013) has 74 citations; the SMT model (Krieglmeyer & Sherman, 2012) has 67 citations; the ABC model (Stahl & Degner, 2007) has 53 citations; and the Trip model (Nadarevic & Erdfelder, 2011) has 13 citations. Naturally, citation count is correlated with year of publication but, from my perspective, this pattern of citations suggests that there are questions waiting to be answered by MPTs that have not yet been utilized to their fullest potential.

Similarly, this article focuses heavily on the IAT which, no doubt, reflects the dominant position the IAT enjoys among implicit measures. However, task conditions determine which cognitive processes influence responses on an implicit measure (Payne et al., 2008). Thus, other implicit measures – such as the GNAT and the EAST, and other sequential priming-type tasks (e.g., Fazio et al., 1995; Krieglmeyer & Sherman, 2012; Payne, 2001; Payne et al., 2005) – would seem well-positioned for MPT analyses to reveal insight into social cognitive processes that are not reflected in IAT responses. Moreover, MPTs have not yet been developed for some implicit measures, such as the implicit relational assessment procedure (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010) and the simple implicit procedure (O’Shea, 2017; O’Shea, Watson, & Brown, 2016). I see this as a growth market.

RT-MPTs. All of the MPTs reviewed here rely on categorical data, e.g., frequency of correct and incorrect responses, which stands in contrast to the latency-based metrics by which implicit bias is usually operationalized (e.g., Greenwald et al., 2003). While some may argue about the superiority of response frequency versus latency to reveal insight into mental contents, I contend that the better metric is the one that matches your research question. A researcher interested in how quickly police officers can respond to an armed suspect would naturally rely on response latency, whereas a researcher interested in whether officers can distinguish armed from unarmed suspects would rely on response frequency. That said, psychological scientists can now have their cake and eat it too: a new class of MPTs have been recently developed that incorporate both response frequency and latency (RT-MPTs: Heck & Erdfelder, 2016; Klauer & Kellen, 2018). Because they include both frequency and latency data, RT-MTs are more psychologically comprehensive than traditional MPTs – with this richer database producing more precise parameter estimates as an added bonus.

To date, RT-MPTs have not been applied to implicit measures, but they seem set to advance the field. For example, associations are often conceptualized to activate more quickly than control processes (e.g., Fazio, 1990), which implies that responses made quickly should primarily reflect the influence of associations rather than control. However, as MPTs reveal, control processes can influence responses on implicit measures that encourage or require fast responses, which complicates the traditional view of the time course of control processes. By incorporating both latency and frequency of responses on implicit measures, RT-MPTs may be able to resolve this, and other, mysteries about implicit measures.

Conclusion

In this article, I have briefly provided examples of some of the contributions that MPTs have made to research using implicit measures. MPTs have helped to resolve mysteries that have arisen in the literature over the years, re-interpreted findings that did not seem mysterious in the first place, and may help to solve mysteries in the future. I hope that this article contributes to the latter, and inspires researchers to include MPTs in their work to continue to solve mysteries and advance research using implicit measures.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Proceedings from the Second Symposium on Information Theory* (pp. 267-281). Budapest: Akadémiai Kiadó.

Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Social context and the self-regulation of implicit bias. *Group Processes and Intergroup Relations*, *13*, 137-150.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger III & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder*. Science conference series (pp. 117–150). Washington, DC: American Psychological Association.

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods*, *46*(3), 668-688.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. *Handbook of Social Cognition*, *1*, 1-40.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY, US: Guilford Press.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, *60*(3), 527-542.

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.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, *6*(3), 242-261.

Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*(4), 631-643.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology*, *81*(5), 760–773.

Cai, H., Sriram, N., Greenwald, A. G., & McFarland, S. G. (2004). The Implicit Association Test's D measure can minimize a cognitive skill confound: Comment on McFarland and Crouch (2002). *Social Cognition*, *22*(6), 673-684.

Calanchini, J., Gonsalkorale, K., Sherman, J. & Klauer, C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, 43(5), 321-325.

Calanchini, J., Lai, C. K., & Klauer, K. C. (2020), Reducing implicit racial preferences: III. A process-level examination of changes in implicit preferences. *Unpublished manuscript.*

Calanchini, J., Sherman, J., Klauer, C., & Lai, C. K. (2014). Attitudinal and non-attitudinal components of IAT performance. *Personality and Social Psychology Bulletin*, 40(10), 1285-1296.

Cone, J.* & Calanchini, J.* (2020). A process dissociation model of rapid implicit revision in response to diagnostic revelations. *Personality and Social Psychology Bulletin.* *co-first authors.

Connelly, S. L., Hasher, L., & Zacks, R. T. (1991). Age and reading: The impact of distraction. *Psychology and Aging, 6*, 533-541.

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.

Corneille, O., & Hütter, M. (2020). Implicit? What Do You Mean? A Comprehensive Review of the Delusive Implicitness Construct in Attitude Research. *Personality and Social Psychology Review*.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition*, *25*(5), 736-760.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*(5), 800-814.

DeHouwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50*(2), 77-85.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5-18.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*(3), 316-326.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23*(75-109), 60318-4.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline?. *Journal of Personality and Social Psychology*, *69*(6), 1013-1027.

Fazio, R. H., & Towles-Schwen, T. (1999). The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 97-116). New York, NY, US: Guilford Press.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. *The Handbook of Social Psychology, 2*(4), 357-411.

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences*, *32*(2), 207-208.

Gonsalkorale, K., Allen, T. J., Sherman, J.W., & Klauer, K. C. (2010). Mechanisms of exemplar exposure effects on implicit attitudes. *Social Psychology*, *41*, 158-168.

Gonsalkorale, K., Sherman, J. W., Allen, T. J., Klauer, K. C., & Amodio, D. M. (2011). Accounting for successful control of implicit racial bias: The roles of association activation, response monitoring, and overcoming bias. *Personality and Social Psychology Bulletin*, *37*(11), 1534-1545.

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.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2014). Measures of implicit attitudes may conceal differences in implicit associations: The case of antiaging bias. *Social Psychological and Personality Science*, *5*(3), 271-278.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480.

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.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

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.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology and Aging*, *17*(3), 482-495.

Hütter, M., & Klauer, K. C. (2016). Applying processing trees in social psychology. *European Review of Social Psychology*, *27*(1), 116-159.

Hütter, M., Sweldens, S., Stahl, C., Unkelbach, C., & Klauer, K. C. (2012). Dissociating contingency awareness and conditioned attitudes: Evidence of contingency-unaware evaluative conditioning. *Journal of Experimental Psychology: General*, *141*(3), 539-557.

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881-919.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*(5), 871–888.

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.

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.

Lai, C. K., Hoffman, K. M., & Nosek, B. A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass*, *7*(5), 315-330.

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 219–234.

McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*(6), 483-510.

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.

Meissner, F., & Rothermund, K. (2015). The Insect-Nonword IAT Revisited. *Social Psychology*, *46*(1), 46-54.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control porcesses. *Zeitschrift für Experimentelle Psychologie*, *48*(2), 107-122.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, *85*(6), 1180-1192.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology*, *65*(3), 469-485.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*(1), 167-184.

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.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*(6), 625-666.

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.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., ... & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36-88.

O'Shea, B. (2017). *Advancing the applicability of absolute implicit measures: using the Simple Implicit Procedure (SIP) to measure responses to pathogen threats* (Doctoral dissertation, University of Warwick).

O’Shea, B., Watson, D. G., & Brown, G. D. A. (2016). Measuring implicit attitudes: A positive framing bias flaw in the Implicit Relational Assessment Procedure (IRAP). *Psychological Assessment, 28*(2), 158–170.

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.

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.

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, *20*(1), 272-314.

Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, *94*(1), 16-31.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*(3), 277-293.

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.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Erlbaum: Potomac, MD.

Rees, H. R., Rivers, A. M., & Sherman, J. W. (2019). Implementation intentions reduce implicit stereotype activation and application. *Personality and Social Psychology Bulletin*, *45*(1), 37-53.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95*(3), 318-339.

Rivers, A. M., Sherman, J. W., Rees, H. R., Reichardt, R., & Klauer, K. C. (2020). On the roles of stereotype activation and application in diminishing implicit bias. *Personality and Social Psychology Bulletin*, *46*(3), 349-364.

Rothermund, K., & Wentura, D. (2004). Underlying processes in the implicit association test: dissociating salience from associations. *Journal of Experimental Psychology: General, 133(2), 139-165*.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review*, *115*(2), 314.

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.

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of EAST performance. *Experimental Psychology*, *54*(2), 99-112.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108-131.

Smith, E. R., & Zárate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*(1), 3-21.

Stewart, B. D., von Hippel, W., & Radvansky, G. A. (2009). Age, race, and implicit prejudice: Using process dissociation to separate the underlying components. *Psychological Science*, *20*(2), 164-168.

Wilson, T. C. (1996). Cohort and prejudice: whites' attitudes toward blacks, hispanics, jews, and asians. Public Opinion Quarterly, 60(2), 253-274.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*(1), 101-126.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*(5), 815-827.

Wu, H., Myung, J. I., & Batchelder, W. H. (2010). Minimum description length model selection of multinomial processing tree models. *Psychonomic Bulletin & Review*, *17*(3), 275-286.