TPM®

A POISSON RACE MODEL FOR THE ANALYSIS OF THE IMPLICIT ASSOCIATION TEST

PASQUALE ANSELMI
MICHELANGELO VIANELLO
LUCA STEFANUTTI
EGIDIO ROBUSTO
UNIVERSITY OF PADOVA

The article presents a Poisson race model for the analysis of the Implicit Association Test (IAT). Four independent and parallel Poisson processes are assumed, one for each category of the IAT. Information about specific characteristics of the stimuli accumulates on the counter of each process until a termination criterion is reached and a response is given. Model parameters are the rates of information accumulation, and the amount of information that is needed before a response is given. The model accounts for both reaction time and response accuracy. The former is determined by the time at which a process wins. The latter depends on the winning process, the category of the presented stimulus, and the kind of block. An empirical application on a Conscientiousness-IAT shows the potential usefulness of the model for understanding the meaning of the implicit measure.

Key words: Implicit association test; Implicit measures; Poisson race model; Poisson model; Response time.

Correspondence concerning this article should be addressed to Pasquale Anselmi, Dipartimento FISPPA – Sezione di Psicologia Applicata, Università di Padova, Via Venezia 8, 35131 PADOVA (PD), Italy. Email: pasquale.anselmi@unipd.it

INTRODUCTION

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is the most popular procedure for measuring automatic associations. At the heart of the procedure is a pair of object categories (e.g., *Flowers* and *Insects*) and a pair of attribute categories (e.g., *Good* and *Bad*), displayed in the top-left and top-right corners of a computer screen. Stimuli representing each of the categories appear, one at a time, in the center of the computer screen, and participants have to categorize them into one of the categories by pressing, as quickly and accurately as possible, one of two response keys. The IAT consists of practice and critical blocks, that are exemplified in Figure 1 (for typographical reasons, the colours in the figure do not reflect the real colours used in the experimental procedure. In the experiment, the background was black, the object labels were white, and the attribute labels were green). The practice blocks involve the categorization of stimuli that represent either the object categories (Figure 1a) or the attribute categories (Figure 1b). The critical blocks involve the categorization of stimuli representing the object categories and stimuli representing the attribute categories, presented in alternating trials. There are two response mappings. In one mapping (Figure 1c), the categories Flowers and Good share a response key (the left key), and the categories Insects and Bad share the other (the right key). In the other mapping (Figure 1d), the categories Flowers and Bad share a response key, and

the categories Insects and Good share the other. The logic behind the IAT is that the categorization task should be easier when the two categories sharing the response key are strongly associated than when they are weakly associated. The mapping that leads to faster and more accurate responses is called compatible mapping, whereas the other one is called incompatible mapping. Compatible and incompatible mapping might differ across individuals. For the sake of simplicity, in the sequel the condition Flowers-Good/Insects-Bad will be referred to as compatible mapping because it is usually found to be easier than the condition Insects-Good/Flowers-Bad (e.g., Govan & Williams, 2004; Greenwald et al., 1998; Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009). The difference in performance between the compatible and incompatible mappings is known as IAT effect.

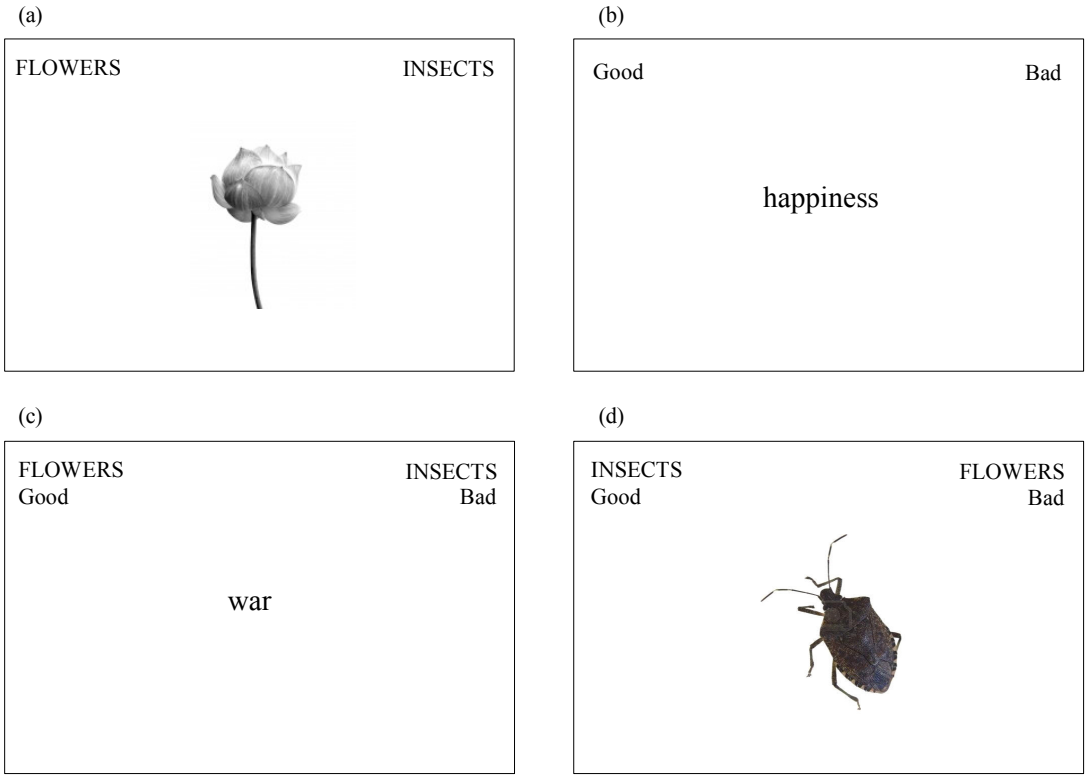


FIGURE 1
Examples of trials concerning the practice (a, b), compatible (c) and incompatible (d) blocks of a Flowers-Insects IAT.

Some authors have tried to investigate the process component underlying the IAT. Conrey, Sherman, Gawronski, Hugenberg, and Groom (2005) proposed a model (called Quad model) that disentangles the influence on the responses to the IAT of four qualitatively distinct processes: the automatic activation of an association because of the presented stimulus, the ability to determine a correct response, the success at overcoming automatically activated associations, and the influence of response biases that may guide responses in the absence of available information. An important limit of this model relies on the fact that it only accounts for response accuracy and, therefore, it only uses a very small part of the information provided by an IAT.

Klauer, Voss, Schmitz, and Teige-Mocigemba (2007) proposed a diffusion model (DM) analysis of the IAT that disentangles three distinct components of the response process: the rate at which information about the presented stimulus is accumulated, the amount of information that must be accumulated before a response is given, and the effect of nondecision components (e.g., encoding stimuli, response execution). The analysis accounts for both accuracy and reaction time, and is based on a hypothesis of serial processing of the stimuli. It follows that accumulation of information toward the two responses is perfectly negatively correlated. Different DMs have to be estimated separately for the blocks of the IAT, so that comparability of parameter estimates across blocks is not straightforward. Moreover, a DM analysis separated for object and attribute stimuli requires a number of trials that is much larger than usual.

A formal model for analyzing the IAT is presented, that has been derived to fit the characteristics of the IAT and the needs of researchers using it. The model is simultaneously estimated on all blocks of the IAT, accounts for both response accuracy and reaction time, and allows a distinct analysis of object and attribute stimuli by using the typical number of trials. In addition, the model is based on a hypothesis of parallel processing of the stimuli. Such a hypothesis is consistent with empirical findings that support most parallel — rather than serial — processing of visual stimuli (see, e.g., Evans, Horowitz, & Wolfe, 2011; Thornton & Gilden, 2007; Townsend, 1990).

OVERVIEW OF THE MODEL

Let us consider a collection O of *object* stimuli, and a collection A of *attribute* stimuli. The stimuli in O have to be categorized into the two object categories [a] and [b], whereas the stimuli in A have to be categorized into the two attribute categories [+] and [−]. In a Flowers-Insects IAT, for example, the object stimuli are pictures of flowers and insects, the attribute stimuli are words with positive (e.g., happiness, love, peace) and negative (e.g., hate, pain, war) valence, the object categories are Flowers and Insects, and the attribute categories are Good and Bad.

The existence of four parallel and independent processes is assumed, one for each category of the IAT. These processes are denoted $X_a(t)$, $X_b(t)$, $X_+(t)$, and $X_-(t)$ [in the aforementioned example, $X_{Flowers}(t)$, $X_{Insects}(t)$, $X_{Good}(t)$, and $X_{Bad}(t)$].

Model assumptions are:

1. every stimulus in $S = O \cup A$ potentially contains — albeit in a variable quantity — evidence for each of the four categories [a], [b], [+], [−];
2. every stimulus in S is simultaneously and independently processed by each of the four processes $X_a(t)$, $X_b(t)$, $X_+(t)$, $X_-(t)$.

Once a stimulus is presented on the screen, each process starts accumulating, on its own counter, selective information about a specific characteristic of it. For example, process $X_a(t)$ accumulates information about the membership of the stimulus to category [a]. The process which accrues the required amount of information (called *termination criterion*) in the shortest time produces the observable response. Thus, our model represents an extension of the standard two-process Poisson race model (Townsend & Ashby, 1983). The four processes are assumed to behave as Poisson processes. This implies that, in each process, interarrival times (i.e., time

intervals between consecutive units of information) are independent and exponentially distributed with rate λ.

Model parameters are the rates at which information accumulates on the counter of each process, and the termination criteria.

Discrimination Rates and Association Rates

The rate at which information is accumulated on the counter of each process depends only on the process and the category of the presented stimulus. Therefore, there is a different rate for each pair that can be formed by taking one of the four processes $X_a(t)$, $X_b(t)$, $X_+(t)$, $X_-(t)$ and one of the four categories [a], [b], [+], [−]. According to this formulation, for $i, j \in \{a, b, +, -\}$, the parameter λ_{ij} is the average amount of information that process $X_i(t)$ accumulates, in the time unit, when a stimulus of category [j] is presented. For example, λ_{aa} and λ_{+a} represent the amount of information respectively accumulated by processes $X_a(t)$ and $X_+(t)$, when the presented stimulus belongs to category [a].

Table 1 displays the 16 rates of the model. They can be grouped into discrimination rates and association rates. The discrimination rates regard the amount of information that object (respectively attribute) categories accumulate when object (respectively attribute) stimuli are presented. In particular, the four rates λ_{aa}, λ_{ab}, λ_{ba}, λ_{bb} (upper left 2 × 2 submatrix of the table) are involved in the discrimination between categories [a] and [b], whereas the four rates λ_{++}, λ_{+-}, λ_{-+}, λ_{--} (lower right 2 × 2 submatrix) are involved in the discrimination between [+] and [−].The discrimination rates provide information about *stimuli discrimination*, that is, whether the stimuli adequately represent their own category. The rates λ_{aa}, λ_{bb}, λ_{++}, λ_{--} are involved in the correct discrimination of the stimuli, whereas the rates λ_{ab}, λ_{ba}, λ_{+-}, λ_{-+} are involved in the incorrect discrimination. The better the discrimination, the smaller the latter rates.

TABLE 1
Discrimination and association rates

	Stimulus categories			
Processes	[a]	[b]	[+]	[−]
$X_a(t)$	$\lambda\lambda_{aa}$	λ_{ab}	λ_{a+}	λ_{a-}
$X_b(t)$	λ_{ba}	λ_{bb}	λ_{b+}	λ_{b-}
$X_+(t)$	λ_{+a}	λ_{+b}	λ_{++}	λ_{+-}
$X_-(t)$	λ_{-a}	λ_{-b}	λ_{-+}	λ_{--}

The association rates regard the amount of information that object (respectively attribute) categories accumulate when attribute (respectively object) stimuli are presented. The four rates λ_{+a}, λ_{+b}, λ_{-a}, λ_{-b} (lower left 2 × 2 submatrix) are the rates at which information concerning membership to attribute categories is accumulated when an object stimulus is presented (*object-driven* associations). The four rates λ_{a+}, λ_{a-}, λ_{b+}, λ_{b-} (upper right 2 × 2 submatrix) are the rates at which information concerning membership to object categories is accumulated when an attribute stimulus is presented (*attribute-driven* associations).

It is stressed here that the association rates are very informative in practical applications of the model. The particular values taken by these parameters might allow the identification of patterns of *automatic association* between objects and attributes that differ from one individual to another in both nature and meaning.

Termination Criteria

The termination criteria concerns the amount of information that has to be accumulated before a response is given. They vary across types of block (practice, compatible, incompatible), response categories (correct, incorrect) and stimulus types (object, attribute), giving rise to a total number of 12 termination criteria. Every single termination criterion is denoted by K_{tbr}, where t specifies the stimulus type (O = object, A = attribute), b specifies the block type (P = practice, C = compatible, I = incompatible), and r specifies the response category (0 = *incorrect*, 1 = *correct*). For instance, K_{OC1} is the termination criterion of the process which provides the correct response in the compatible blocks, when an object stimulus is presented on the screen.

The termination criteria may be the combined result of individual cautiousness, task difficulty, and their interaction. In practical applications of the model, it is reasonable to expect that the practice blocks are the easiest and the incompatible blocks are the hardest. Formally, this would imply that the following inequalities hold true:

$$K_{tPr} < K_{tCr} < K_{tIr} .$$

Moreover, an incorrect response is expected to be the effect of carelessness or inattention. Formally, this would imply that the following inequality holds true:

$$K_{tb0} < K_{tb1}.$$

Compound Processes and Components of the IAT Effect

Figure 2 depicts how the four processes are related to both the presented stimulus and the observable response in a single trial of the compatible (left-hand diagram) and incompatible (right-hand diagram) blocks of an IAT.

The left-hand diagram refers to the condition in which categories [*a*] and [+] are mapped onto the left key, whereas [*b*] and [−] are mapped onto the right key. Arrows departing from the stimulus and arriving at each of the four processes represent the rate of information accumulation of each process (parameters λ). Concerning the relationship between each of the processes and each of the two responses, it is seen that both processes $X_a(t)$ and $X_+(t)$ produce the response *left*, whereas both processes $X_b(t)$ and $X_-(t)$ produce the response *right*.

It is assumed that information units that contribute to the same observable response are accumulated by the same compound process. According to a property of Poisson processes, the superposition $X_1(t) + X_2(t)$ of two independent Poisson processes, with rates λ_1 and λ_2 respectively, is a Poisson process with rate $\lambda_1 + \lambda_2$.

In the compatible blocks, a race takes place between the two compound processes $X_{a+}(t) = X_a(t) + X_+(t)$ and $X_{b-}(t) = X_b(t) + X_-(t)$, each of which has its own termination criterion K (Figure 3). The process which first meets its own criterion will produce the observable response.

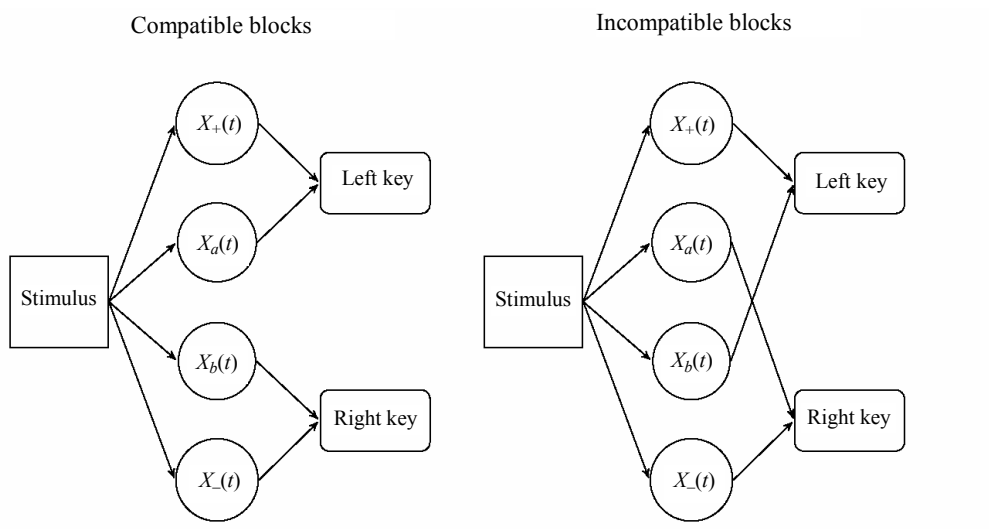


FIGURE 2
Connection of the four processes with the presented stimulus and the observable response in the compatible (left-hand diagram) and incompatible (right-hand diagram) blocks.

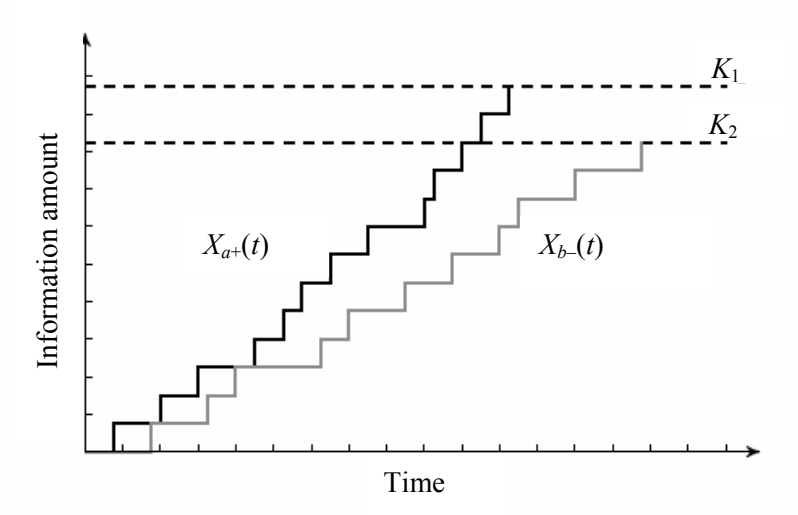


FIGURE 3
A race between two compound Poisson processes. Process $X_{a+}(t)$ has both a higher rate and a higher termination criterion (K_1) than process $X_{b-}(t)$. In this case the race is won by process $X_{a+}(t)$, which accumulates the required information in the shortest time.

In the right-hand diagram of Figure 2, the positions of the labels [*a*] and [*b*] are interchanged on the screen so that now categories [*a*] and [–] are mapped onto the right key, whereas [*b*] and [+] are mapped onto the left key. In this condition, the race is therefore between the two compound processes $X_{a-}(t) = X_a(t) + X_-(t)$ and $X_{b+}(t) = X_b(t) + X_+(t)$.

The inversion of the two labels between the compatible and incompatible blocks represents a constituent element of the IAT effect. Let us suppose that a stimulus of category [*a*]

is presented on the screen. Then $X_{a+}(t)$ is the correct process (i.e., the process that provides the correct response) in the compatible blocks, whereas $X_{a-}(t)$ is the correct process in the incompatible blocks. Assuming, for the moment, identical termination criteria in the two blocks, if the condition $\lambda_{+a} > \lambda_{-a}$ holds true, then the expected response time of the correct response will be longer in the incompatible blocks, compared to the compatible blocks. This occurs because the rate of the correct process in the incompatible blocks ($\lambda_{aa} + \lambda_{-a}$) is smaller than that of the correct process in the compatible blocks ($\lambda_{aa} + \lambda_{+a}$).

Assuming now identical association rates for the stimulus of category [a] (i.e., $\lambda_{+a} = \lambda_{-a}$), then the two correct processes $X_{a+}(t)$ (compatible blocks) and $X_{a-}(t)$ (incompatible blocks) would have identical rates. Nonetheless, an IAT effect would still be observed whenever the condition $K_{OC1} < K_{OI1}$ holds true.

An Empirical Application

Participants

One hundred and one psychology students at the University of Padua participated in the study with no financial reward. Their mean age was 23.22 (*SD* = 1.39), and 53 were female.

Material and Procedure

Participants were presented with a Conscientiousness-IAT according to the structure represented in Table 2.

Table 2
Structure of the Conscientiousness-IAT

Block	N. of trials	Left labels	Right labels
1 (Practice)	20	Others	Self
2 (Practice)	20	Not conscientious	Conscientious
3 (Compatible)	20	Others-Not conscientious	Self-Conscientious
4 (Compatible)	36	Others-Not conscientious	Self-Conscientious
5 (Practice)	30	Self	Others
6 (Incompatible)	20	Self-Not conscientious	Others-Conscientious
7 (Incompatible)	36	Self-Not conscientious	Others-Conscientious

Note. Blocks 3 and 4 were counterbalanced across participants with blocks 6 and 7.

The Conscientiousness-IAT used the category labels Self, Others, Conscientious and Not conscientious. Eight pronouns were used to represent the object categories Self (I, me, mine, myself) and Others (others, them, they, you). The eight bipolar pairs of adjectives of the conscientiousness scale of the Big Five Observer (Caprara, Barbaranelli, & Borgogni, 1994) were used to represent the attribute categories Conscientious (careful, conscientious, dependable, efficient, hard-working,

orderly, organized, persevering) and Not conscientious (careless, inefficient, lazy, messy, negligent, disorganized, sloppy, superficial).

Participants were tested individually in a laboratory. The stimuli were presented in the center of the computer screen in an alternating fashion, and participants were asked to categorize them by pressing, as quickly and accurately as possible, the response key "A" or "L," respectively. A red cross appeared in case of mistake and it disappeared after the response was corrected. The order of compatible (Others-Not conscientious/Self-Conscientious) and incompatible (Self-Not conscientious/Others-Conscientious) blocks was counterbalanced across the participants.

Analysis Procedure

Four vectors were obtained for each participant. The vectors specify the kind of block (practice, compatible, incompatible), the category a stimulus belongs to (S-Self, O-Others, C-Conscientious, N-Not conscientious), the accuracy of a response (1, 0), and its latency in milliseconds. Usually the latency of incorrect responses is increased by the time that is required to correct them (see, e.g., Greenwald, Nosek, & Banaji, 2003). We used the latency of incorrect responses and discarded the time needed to correct them. The length of the vectors was equal to the number of trials (n = 182). The four vectors represent the input of a Matlab function (implemented by the authors) for computing maximum likelihood estimates of the parameters of the model. Since maximum likelihood estimation is sensitive to outliers and contaminants in the distributions of response times, trials whose latencies were outliers in the response time distribution were discarded according to Tukey's criterion (see, e.g., Hoaglin, Mosteller, & Tukey, 1983). This led to the deletion of 6.14% of all the responses. The estimate convergence criterion was that the maximum adjustment among all parameters was less than 10^{-7}.

Model identifiability was tested for each participant by computing the Hessian matrix of the likelihood function (Gradshteyn & Ryzhik, 2000) and checking its positive definiteness. The goodness-of-fit of the model was tested for each participant using Pearson's chi-square statistic according to the procedure described in Klauer et al. (2007). The probability of the chi-square was computed both according to the theoretical chi-square distribution and using a parametric bootstrap procedure. For each participant, parametric bootstrap was performed in the following way. Using the parameters estimated on the observed data, 750 pairs of vectors specifying accuracy and latency of responses to the IAT were simulated, and on these vectors the model parameters were estimated. Then, the proportion of simulated data whose chi-square was greater than that obtained on the observed data was computed.

RESULTS

In the following section, the analysis is restricted to the 73 out of 101 participants whose model was accepted ($\chi^2 p > .05$). For each of these, the Hessian matrix was positive definite.

To deal with aberrant estimates, we decided to use Tukey's criterion which led to the deletion of 7.31% of the K parameters and 7.11% of the λ parameters. Our expectation that blocks are of increasing difficulty, with the practice blocks being the easiest and the incompatible blocks being the most difficult, was confirmed. Mean termination criteria were computed for

practice, compatible, and incompatible blocks. Friedman's test for dependent samples showed that blocks were of different difficulty (Mean Rank = 1.18, 2.23, 2.59 for practice, compatible and incompatible block, respectively; $\chi^2_{(2)} = 78.60, p < .001$). Wilcoxon's test highlighted that the incompatible blocks were more difficult than the compatible blocks ($Z = -2.993$, $p < .01$), which were more difficult than the practice blocks ($Z = -6.754$, $p < .001$). Mean termination criteria were also computed for correct and incorrect responses. Wilcoxon's test showed that the mean criterion of incorrect responses was lower than that of correct responses ($Z = -5.47$, $p < .001$), confirming our expectation.

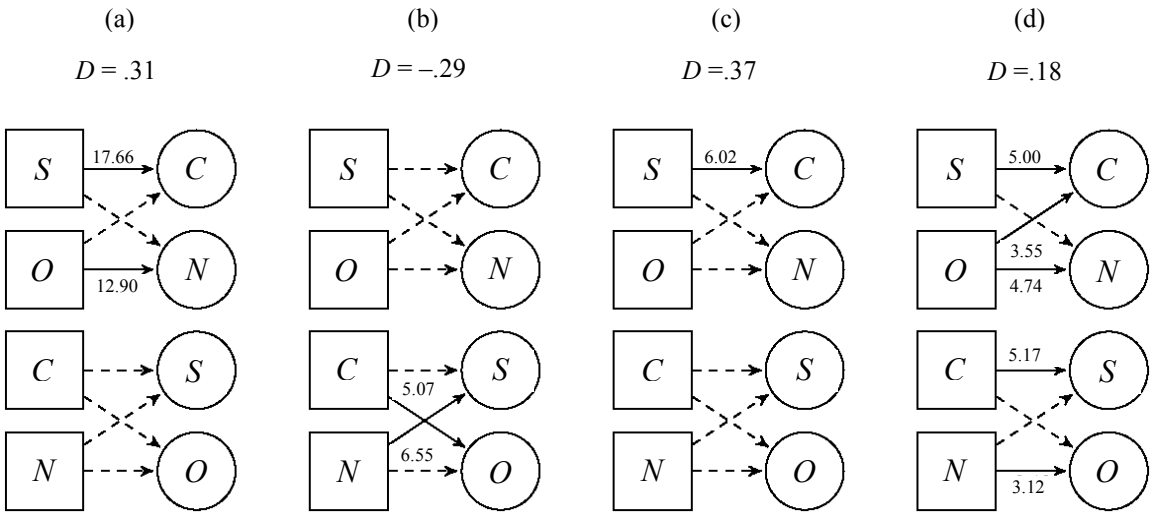
In order to highlight the discriminability of stimuli belonging to the four categories, the ratio between the rates concerning correct and incorrect discrimination was computed for each stimulus category (e.g., $R_S = \lambda_{SS} / \lambda_{OS}$). In all participants, the stimuli — and in particular those belonging to the two object categories — provided, in the time unit, more information toward the correct response than toward the incorrect one ($R_S = 22.22$; $R_O = 21.36$; $R_C = 3.74$; $R_N = 4.63$; the values are the median across participants).

We now show how the association rates might highlight automatic associations that differ in nature and meaning. In the following, the term "object-driven positive association" is used when $\lambda_{CS} > \lambda_{NS}$ and $\lambda_{NO} > \lambda_{CO}$. It means that, on average, the Self stimuli provide more information, in the time unit, about the category Conscientious than Not conscientious, as well as the Others stimuli provide more information about Not conscientious than Conscientious. In other words, the Self-Conscientious association is stronger than the Self-Not conscientious association, and the Others-Not conscientious association is stronger than the Others-Conscientious association. The term "object-driven negative association" is used when $\lambda_{NS} > \lambda_{CS}$ and $\lambda_{CO} > \lambda_{NO}$. The terms "attribute-driven positive association" and "attribute-driven negative association" are used when $\lambda_{SC} > \lambda_{OC}$ and $\lambda_{ON} > \lambda_{SN}$, and when $\lambda_{OC} > \lambda_{SC}$ and $\lambda_{SN} > \lambda_{ON}$, respectively.

Figure 4 depicts the standardized association rates of four participants. Unbroken arrows represent the standardized rates significantly greater than 0. The *D* scores of these participants are also provided. The *D* is the most common measure of the IAT effect size. It basically involves dividing the difference in average response latency between the compatible and incompatible blocks by the standard deviation of latencies for all the critical blocks (for details, see Greenwald et al., 2003). A positive *D* score means that the stimuli are categorized faster in the compatible blocks than in the incompatible blocks. A negative *D* score means the opposite.

An object-driven positive association is observed in participant *a* (Figure 4a), whereas an attribute-driven negative association is observed in participant *b* (Figure 4b). Exemplars of object-driven negative association and attribute-driven positive association are not observed in the present sample of 73 participants. An association is observed in participant *c* which can be regarded as a partial object-driven positive association because the Self stimuli only contribute to the measure (Figure 4c). The *D* scores of participants *a* and *c* are similar in size but different in meaning. Participant *c* only associates the Self stimuli with Conscientious, whereas participant *a* also associates the Others stimuli with Not conscientious. The *D* score of participant *d* is rather small compared to that of the other participants. Nevertheless, implicit associations exist in this participant. Behind the *D* score, an object-driven positive association and an attribute-driven positive association are constrasted by an object-driven negative association (Figure 4d). In fact, the Others stimuli provide a sizeable amount of information not only about Not conscientious but

also about Conscientious. An unpolarized evaluation of the Others stimuli is present here, that would not have been noticed without the decomposition of the association rates.



Note. Squares represent stimuli, circles represent processes, arrows represent association rates. *S* = Self, *O* = Others, *C* = Conscientious, *N* = Not conscientious. Unbroken arrows represent significant association rates ($p < .05$; Bonferroni correction).

FIGURE 4
Standardized association rates and *D* scores of four participants.

How Model Parameters Predict the IAT Score *D*

A stepwise regression analysis was run in which the model parameters were the independent variables and the *D* was the dependent variable. In order to avoid multicollinearity, the termination criteria were replaced by difficulty indexes (*DI*) computed by taking the difference between termination criteria of compatible and incompatible blocks. Four *DI*s were computed, one for each stimulus category (objects and attributes) and each response category (correct and incorrect; e.g., $DI_{A1} = K_{AI1} - K_{AC1}$). The *DI*s and the accumulation rates were entered into the model if the probability value of their *F* was smaller than .05 and were removed if it was greater than .10. The following independent variables entered in the regression model: the association rates concerning Conscientious stimuli (λ_{SC}, λ_{OC}), the rate concerning correct discrimination of Not conscientious stimuli (λ_{NN}), the rate concerning incorrect discrimination of Self stimuli (λ_{OS}), the *DI*s concerning correct and incorrect responses to attribute stimuli (DI_{A1}, DI_{A0}). None of the parameters showed multicollinearity (Tolerance ≥ .39). The final regression model (Table 3) accounted for approximately 80% of variance of *D* ($R^2 = .82$; $AdjR^2 = .78$). The indexes DI_{A1} and DI_{A0} account altogether for 48% of variance. The rates concerning the association of Conscientious stimuli to categories Self and Others accounts for 18% of variance. Therefore, in our Conscientiousness-IAT the score *D* mostly expresses the different difficulty in categorizing the attribute stimuli in the compatible and incompatible blocks, and the associations driven by Conscientious stimuli.

TABLE 3
Variance of the *D* score accounted for by the independent variables that entered in the regression model

Independent variable added	R^2 change	β	*t*	*p*
DI_{A1}	.31	1.06	9.44	< .01
DI_{A0}	.17	.34	3.42	< .01
λ_{OC}	.11	–.48	–5.22	< .01
λ_{OS}	.10	–.22	–2.59	< .05
λ_{NN}	.06	–.39	–4.08	< .01
λ_{SC}	.07	.41	3.18	< .01

DISCUSSION

A formal model has been presented, that disentangles the influences on the responses to the IAT of three qualitatively distinct process components: stimuli discrimination, automatic association and termination criterion. These components provide information about the response process which is relevant at different levels.

Stimuli discrimination informs us whether the stimuli that have been selected for representing a certain category are prototypical exemplars of that category and can be easily recognized and categorized. The relevance of this aspect for the validity of the implicit measure is well documented in the literature (see, e.g., Lane, Banaji, Nosek, & Greenwald, 2007; Steffens & Plewe, 2001). In our application, we found that the adjectives used to represent the categories Conscientious and Not conscientious, and even more so the pronouns used to represent the categories Self and Others, provided more information toward the correct discrimination than toward the incorrect one.

Automatic association informs us about the strength of the associations between objects and attributes. By identifying object-driven and attribute-driven associations, the model enables the researcher to observe differences between individuals in the nature and meaning of the implicit measure. In our application, we were able to identify participants in which the implicit measure of conscientiousness was mostly driven by the concepts (Self and Others), the attributes (Conscientious and Not conscientious), or both. The possibility of distinguishing between implicit measures that hold the same direction (i.e., the same *D* score) but different meanings represents a strong potential in some research fields, such as the investigation of race attitude and consumer behavior. In the former, the model would allow the distinction between individuals with an actual implicit prejudice and individuals who merely hold an implicit preference for their own group. In the latter, the model would allow the distinction between consumers who choose a product because they like it and consumers who choose the same product because they do not like the alternatives that are available (see, e.g., Stefanutti, Robusto, Vianello, & Anselmi, 2013).

Termination criteria determine the amount of information that must be accumulated before a decision toward a response can be made, and the associated response key is pressed. Thus, these parameters reflect individual cautiousness, task difficulty, and their interaction.

The model enables a fine-grained analysis of the IAT effect. It is interesting to note that, in our Conscientiousness-IAT, the classic metric *D* mostly reflected the different difficulty in categorizing the attribute stimuli in the compatible and incompatible blocks, and that the most

important associative component was represented by the conscientious stimuli. The *D* also reflects, in a smaller part, discrimination components that might confound the resulting measure of association.

The proposed model attempts to take into account the complexity of the IAT procedure through parameters that are specific for correct and incorrect responses to stimuli representing different categories and presented in different blocks. This results in a total of 28 parameters for each individual. The model is identifiable, despite the number of parameters, and the parameter estimates are reasonable (e.g., λ of incorrect responses are smaller than those of correct responses; *K* of practice blocks are smaller than those of critical blocks). Moreover, the model is much more parsimonious than the DM which, for a separate analysis of object and attribute stimuli in the different blocks of the IAT, estimates 49 parameters for each individual (see Klauer et al., 2007). The number of parameters might be reduced by introducing equality constraints among parameters. Specific assumptions about the constraints must be based on peculiar features of the IAT at hand, and need to be tested on empirical data.

Usually there are only a few incorrect responses in the blocks of the IAT and, sometimes, none. When this happens, there is not sufficient information in the data to compute stable and nondegenerate estimates of all parameters. For example, if there are no incorrect responses to the attribute stimuli in the incompatible blocks, the parameter K_{AI0} can not be estimated. No or few incorrect responses represent a limit also for the Quad model and the DM. We point out here that the presented model considers separate processes for correct and incorrect responses. As a consequence, even if the estimates of the parameters concerning the incorrect responses were not reliable, this would not be the case for those of the parameters concerning the correct responses. A possible solution to reduce the impact of no or few incorrect responses on parameter estimates might be to impose equality constraints to termination criteria of the correct and incorrect processes. Otherwise, it may be possible to try to increase the number of incorrect responses by adding trials or by introducing a response window in the IAT procedure.

The empirical application presented in this article aimed to show the potential usefulness of the model for a deeper understanding of the IAT measure. However, it is worth noting that further empirical studies are required in order to validate the proposed interpretation of parameters. Research on how much the model parameters reflect construct-specific variance is ongoing at the present time. Current research is devoted to analyzing test-retest reliability of parameters. We expect that, in repeated administrations of the same IAT to the same sample of participants, the termination criteria and the discrimination rates would vary as an effect of training with the task and the stimuli, respectively. On the contrary, the association rates are expected to remain essentially unchanged, if no intervention is made in order to modify automatic associations. Current research is also aimed at improving the model fit. Response latencies of objects and attributes in the different blocks, as well as the amount of information concerning incorrect responses are being considered. Moreover, work is being done to incorporate nondecision components within the model.

REFERENCES

Caprara, G. V., Barbaranelli, C., & Borgogni, L. (1994). *BFO, Big Five Observer. Manuale* [FO, Big Five Observer. Manual]. Firenze, Italy: Organizzazioni Speciali.

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. doi:0.1037/0022-3514.89.4.469

Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). When categories collide: Accumulation of information about multiple categories in rapid scene perception. *Psychological Science*, *22*, 739-746. doi:10.1177/0956797611407930

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology*, *40*(3), 357-365. doi:10.1016/j.jesp.2003.07.002

Gradshteyn, I. S., & Ryzhik, I. M. (2000). *Table of integrals, series and products* (6th ed.). New York, NY: Academic Press.

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

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

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (Eds.). (1983). *Understanding robust and exploratory data analysis*. New York, NY: John Wiley & Sons.

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*(3), 353-368. doi:10.1037/0022-3514.93.3.353

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies* (pp. 59-102). New York, NY: Guilford Press.

Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *The Quarterly Journal of Experimental Psychology*, *62*(1), 84-98. doi:10.1080/17470210701822975

Stefanutti, L., Robusto, E., Vianello, M., & Anselmi, P. (2013). A Discrimination-Association model for decomposing component processes of the Implicit Association Test. *Behavior Research Methods*, *45*, 393-404. doi:10.3758/s13428-012-0272-3

Steffens, M. C., & Plewe, I. (2001). Items' cross category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie*, *48*, 123-134. doi:10.1026//0949-3946.48.2.123

Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search. *Psychological Review*, *114*(1), 71-103. doi:0.1037/0033-295X.114.1.71

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46-54. doi: 10.1111/j.1467-9280.1990.tb00067.x

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York, NY: Cambridge University Press.

TPM®