



# Forgiving, fast and slow: validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm

Ramzi Fatfouta<sup>1,2\*</sup>, Michela Schröder-Abé<sup>3</sup> and Angela Merkl<sup>1,4</sup>

<sup>1</sup> Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Berlin, Germany

<sup>2</sup> Affective Neuroscience and Psychology of Emotion, Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

<sup>3</sup> Department of Psychology, Technische Universität Darmstadt, Darmstadt, Germany

<sup>4</sup> Department of Neurology with Experimental Neurology, Charité–Universitätsmedizin, Berlin, Germany

## Edited by:

John M. Zelenski, Carleton University, Canada

## Reviewed by:

Monika Fleischhauer, Technische Universität Dresden, Germany  
Michael Dufner, University of Leipzig, Germany

## \*Correspondence:

Ramzi Fatfouta, Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany  
e-mail: ramzi.fatfouta@fu-berlin.de

This study examined the role of automaticity in forgiving a real-life offense. As an alternative to self-report, an Implicit Association Test (IAT) of forgiveness was developed. Implicit (IAT-measured) and explicit (self-reported) forgiveness predicted shorter response times of state forgiveness ratings. The forgiveness IAT was highly reliable, moderately stable over time, and demonstrated incremental validity. Results suggest that the newly introduced forgiveness IAT could advance personality research beyond what is known from self-report measures, further corroborating the notion of implicit forgiveness. Implications for personality assessment are discussed.

**Keywords:** forgiveness, implicit self-concept, automaticity, Implicit Association Test (IAT), response latency measures

## INTRODUCTION

Individual differences in forgiveness have traditionally been assessed via self-report. Contemporary measures, however, can be biased in two fundamental ways. First, they are susceptible to social desirability (Hoyt and McCullough, 2005), and second, they are insensitive to processes outside of awareness (Bornstein, 2002; Moors and De Houwer, 2006). This latter point is problematic as one might explicitly state one has forgiven an offense, yet implicitly continue harboring a grudge (hollow forgiveness; Fincham, 2010). To remedy these limitations, indirect measures may prove useful.

Indirect measures tap into different psychological processes than direct measures (i.e., self-reports). In accordance with dual-process models (for an overview, see Strack and Deutsch, 2004; Back et al., 2009), human behavior is a joint function of both reflective and impulsive processes. Reflective processes operate slowly and refer to propositional or explicit representations of the self that result from reasoning. Impulsive processes operate fast and refer to associative or implicit representations of the self that are activated automatically (i.e., non-deliberatively) when encountering situational cues. While reflective processes can be assessed via direct measures, impulsive processes can be assessed indirectly. Notably, indirect measures have been found to provide increments in predictive validity beyond direct measures—which highlights their value for a more complete assessment of trait factors and for the prediction of trait-relevant behavior (e.g., Back et al., 2009; Rudolph et al., 2010; Fleischhauer et al., 2013).

The Implicit Association Test (IAT; Greenwald et al., 1998) is the most popular indirect measure of automatic associations

and has been shown to be relatively resistant to faking (Steffens, 2004; Röhner et al., 2011). Using a speeded categorization task, the IAT measures the (relative) association strength between a target concept (e.g., me–others) and an attribute dimension (e.g., forgiving–vengeful). The rationale is that when individuals strongly associate their self with congruent (e.g., me–forgiving) vs. incongruent (e.g., me–vengeful) attributes, categorization will be easier (i.e., faster). The IAT effect, defined as response-latency difference between congruent and incongruent pairings, thus reflects associative links between the self and a trait concept (in our case, forgiveness) and may be referred to as implicit self-concept of personality (Schnabel and Asendorpf, 2010). To date, however, no study has attempted to look at the implicit self-concept of forgiveness.

Existing research has conceptualized forgiveness as a deliberative process with a presumed endpoint, that is, one's decision to forgive a transgressor (Fincham, 2000; Fincham et al., 2005). Specifically, to forgive involves a prosocial change in individuals' thoughts, feelings, and motivations, whereby they become more positively disposed toward their transgressor (McCullough et al., 1998; Worthington and Wade, 1999). However, as Fincham et al. (2006, p. 422) noted, the focus on deliberative processes "may overlook aspects of forgiveness that occur outside of conscious awareness." Indeed, recent work has provided empirical support for this contention (for an overview, see Karremans and Lange, 2008). In close relationships, for instance, the inclination to forgive occurs automatically and without the individuals' intention. The reasoning is that individuals automatically associate forgiveness with beneficial outcomes (e.g., relationship

satisfaction), thereby reinforcing impulsive/automatic responding in a pro-relational manner (Karremans and Aarts, 2007, Studies 1 and 2). This is consistent with associative-learning processes in dual-process models (Strack and Deutsch, 2004; Back et al., 2009) and calls for a multimethod approach including indirect measures like the IAT. Consequently, it is necessary to examine both explicit *and* implicit aspects of forgiveness.

Based on these findings, the current study aimed to develop an IAT that taps individual differences in *implicit forgiveness* and to test its predictive validity over and above a corresponding self-report questionnaire. As forgiveness-related behavioral criteria, we used the response times (RTs) individuals need to rate their state forgiveness (i.e., their current thoughts and feelings toward an offender; McCullough et al., 1998). In past research, the measurement of response latency has been employed as an operative indicator of attitude strength (Fazio, 2001). The logic here is that RTs reflect the individual's mental effort to arrive at a response and therefore indicate the accessibility of information in memory. The more accessible the information is, the faster the individual responds (Fazio, 1990a,b). A personality-oriented adaptation of this approach has been provided by Lischetzke et al. (2005), who demonstrated that when individuals can easily access internal cues about their affective state they are faster to rate it.

With respect to forgiveness, a similar account has been proposed. Specifically, individuals who are consistently inclined to forgive are assumed to “have both a higher rate of forgiveness and a *shorter* latency of response” (Sutton and Thomas, 2006, p. 33; emphasis added). Hence, it appears plausible that when prompted (1) to recall an offense and then (2) to rate their state forgiveness, individuals who have already worked through the pain should more easily access internal cues about their forgiveness and, hence, exhibit faster responses. By contrast, individuals who have not yet achieved forgiveness (i.e., who find internal forgiveness cues to be rather difficult to access), should need comparatively more time to contemplate and report their responses. For example, responding to an item such as “I am finding it difficult to act warmly toward him/her” (McCullough et al., 2006) should be less effortful (i.e., RT should be faster) for a person who has come to terms with the hurt<sup>1</sup>.

Response latencies reflect both deliberative and automatic processing (Fazio, 1990a). Therefore, we hypothesized that both the newly developed forgiveness IAT and a forgiveness self-report would predict RT of forgiveness ratings. Specifically, high-forgiving individuals—measured by self-report and indirectly—should be faster to rate their state forgiveness (*Hypothesis 1*). Given that incremental validity is an important psychometric property when evaluating indirect measures (Perugini and Banse,

2007), we tested whether the novel forgiveness IAT predicted RT of state forgiveness ratings beyond explicit (self-reported) forgiveness (*Hypothesis 2*).

## MATERIALS AND METHODS

We report all data exclusions (if any), manipulations, and measures, and how we determined our sample size. The latter was determined a priori with the goal to obtain around 100 observations.

### PARTICIPANTS

One hundred and four students (70 women;  $M_{\text{age}} = 25.10$ ,  $SD = 4.39$ ) of Freie Universität Berlin participated in exchange for money or course credit.

### PROCEDURE

Data were collected individually or in sessions of two individuals. After providing written informed consent, participants completed the forgiveness IAT and a set of questionnaires on a computer screen described below. Each questionnaire item appeared separately on the screen. Participants answered by mouse click on the respective response category. They had to click a *next* button to proceed with the next item.

At the beginning of the computer-based questionnaire, participants answered questions to obtain a measure of their baseline speed (see Baseline speed). Participants then had (1) to think of a close person; (2) type in his/her name; (3) indicate the type of relationship involved, and (4) rate their perception of closeness to the self-selected other (Karremans and Aarts, 2007). Next, participants were instructed to bring to mind a real-life situation in which the other had hurt them and to rate their state forgiveness. RT (i.e., the time between item presentation and click on the *next* button) was unobtrusively recorded (Lischetzke et al., 2005). Subsequently, participants rated the severity of the transgression and completed measures of explicit forgiveness as well as social desirability. Finally, participants were thanked and debriefed.

### MEASURES

#### *Implicit forgiveness*

Implicit forgiveness was measured with an IAT (Greenwald et al., 1998). The forgiveness IAT was specifically designed for this study and comprised five blocks (for procedural details, see **Table 1**). In each trial, a stimulus word was presented in the center of the screen. Participants then pressed either a left or a right key to categorize the stimulus as quickly and accurately as possible into one of the categories. The category labels appeared in the upper left and right hand corners of the screen.

The first two blocks consisted of a simple discrimination task, in which participants practiced correctly categorizing stimuli from the target category (me–others) and attribute category (forgiving–vengeful). The third block combined both discrimination tasks (i.e., me–forgiving for the left key; others–vengeful for the right key). In block 4, the labels of the attribute categories were reversed (vengeful–forgiving). Block 5 consisted of the reversed combined discrimination task (i.e., me–vengeful for the left key; others–forgiving for the right key).

<sup>1</sup>To demonstrate that higher forgiveness is indeed related to faster responses, we performed a pilot study ( $N = 120$ ,  $M_{\text{age}} = 22.47$ ,  $SD = 4.33$ ), instructing participants to recall an experienced hurt and to rate their motivation to seek revenge against the transgressor (while RT was unobtrusively measured). Participants then indicated their forgiveness (i.e., “I have forgiven what he/she did to me”). As predicted, RT was significantly negatively correlated with forgiveness ( $r = -0.20$ ,  $p = 0.031$ ).

**Table 1 | Forgiveness IAT: Task Sequence and Stimuli.**

Block	Task	Target concepts	Attribute concepts <sup>a</sup>	Key assignment <sup>b</sup>	
				Left (A)	Right (5)
1 (20)	Target discrimination	Me–others		Me, my, own, I, self	They, your, them, you, others
2 (20)	Attribute discrimination		Forgiving–Vengeful	Forgiving, conciliatory, merciful, gracious, lenient	Vengeful, punitive, hostile, merciless, unforgiving
<b>3 (60)</b>	Combined discrimination	Me–others	Forgiving–Vengeful	Me, my, own, I, self	They, your, them, you, others
				Forgiving, conciliatory, merciful, gracious, lenient	Vengeful, punitive, hostile, merciless, unforgiving
4 (40)	Reversed attribute discrimination		Vengeful–Forgiving	Vengeful, punitive, hostile, merciless, unforgiving	Forgiving, conciliatory, merciful, gracious, lenient
<b>5 (60)</b>	Combined reversed discrimination	Me–others	Vengeful–Forgiving	Me, my, own, I, self	They, your, them, you, others
				Vengeful, punitive, hostile, merciless, unforgiving	Forgiving, conciliatory, merciful, gracious, lenient

IAT, Implicit Association Test. The number of trials per block is in parentheses. Critical blocks are in bold. <sup>a</sup>Stimuli were piloted ( $N = 50$ ,  $M_{age} = 24.96$ ,  $SD = 4.43$ ) and matched according to familiarity, valence, similarity, and potency. <sup>b</sup>Number 5 from the numeric keyboard.

The order of critical blocks (i.e., blocks 3 and 5) was held constant to maximize the reliability of person effects (Banse, 2001). IAT-effects were computed using the improved scoring algorithm ( $D_1$ ; Greenwald et al., 2003). Higher scores reflect stronger automatic associations between me–forgiving (vs. me–vengeful), and hence, a more forgiving implicit self-concept.

To calculate the forgiveness IAT's reliability, we applied the  $D_1$ -algorithm to two mutually exclusive subsets of the critical trials (Schmukle and Egloff, 2006). The Spearman-Brown corrected split-half correlation of this two-part measure was close to excellent ( $r_{tt} = 0.89$ ). To provide an estimate of temporal stability, the last 30 participants (19 women;  $M_{age} = 26.60$ ,  $SD = 4.35$ ) were re-contacted after 1 month ( $M_{time\ interval} = 32.70$  days,  $SD = 3.70$ ) to complete the forgiveness IAT in a follow-up assessment. The temporal stability was moderate (0.50; 95% confidence interval = 0.34; 0.63) and similar to that previously reported for IATs (Egloff et al., 2005).

### Baseline speed

Because we used RT as an individual difference measure, we controlled for differential baseline response speed (Fazio, 1990b). Baseline response speed was assessed by means of seven easy-knowledge questions [e.g., “What type of celestial body is the earth? (1) White giant; (2) Planet; (3) Asteroid; (4) Moon”; (Lisetzke et al., 2005)] to which the correct answer was evident, so that RT ( $\alpha = 0.71$ ) tapped only differential reading speed and motor behavior (e.g., muscle speed). Mean item accuracy was high (99.36%,  $SD = 0.03$ ), thus corroborating this idea. Consistent with Lisetzke and colleagues, we omitted the first (training) item and calculated the median of the remaining six RT

scores. Further validity evidence for this baseline speed measure can be found in Lisetzke et al. (2011).

### Closeness

Closeness to the transgressor was assessed with the Inclusion of the Other in the Self scale (IOS; Aron et al., 1992), a pictorial measure comprising seven pairs of two increasingly overlapping circles, labeled “Self” and “Other.” Each pair was numbered and participants had to click on the respective response option to indicate their perceived closeness (1 = *not at all close*, 7 = *very close*). The IOS has demonstrated high alternate-form and test–retest stability as well as strong convergence with multi-item measures of relational closeness (Aron et al., 1992).

### State forgiveness

State forgiveness was measured with the Transgression-Related Interpersonal Motivations Inventory (TRIM; McCullough et al., 1998, 2006). It assesses participants' motivational changes toward a specific transgressor and is divided into three subscales, with five items measuring revenge (e.g., “I'm going to get even”), seven items measuring avoidance (e.g., “I keep as much distance between us as possible”), and six items measuring benevolence (“I have given up my hurt and resentment”). Items were rated on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). Past research has evidenced that the subscales have high internal consistency, moderate temporal stability, and evidence of construct validity (McCullough et al., 1998, 2006; McCullough and Hoyt, 2002). In this sample, internal consistency ( $\alpha$ ) was good for avoidance and benevolence (0.87 and 0.76, respectively), but somewhat modest for revenge (0.53). While this alpha is less than ideal, similar estimates have been reported with the revenge scale ( $\alpha = 0.52$ ;

Allemand et al., 2013). Unlike previous research, however, our primary analyses focused on participants' RTs, and hence, this value is deemed reasonable.

### RT of state forgiveness ratings

As an objective measure of state forgiveness, RT to each TRIM item was recorded. Again, the first item was regarded as a training item and its RT was excluded from the analyses. We then log-transformed ( $\log_e[x]$ ) RTs for each person to correct for positive skewness (Fazio, 1990b) and calculated mean RT scores for each of the three TRIM subscales:  $RT_{\text{Revenge}}$  ( $\alpha = 0.65$ ),  $RT_{\text{Avoidance}}$  ( $\alpha = 0.78$ ), and  $RT_{\text{Benevolence}}$  ( $\alpha = 0.70$ ). Using three regression equations, we computed—separately for each TRIM subscale—residual RT scores by partialling out baseline speed. Values greater (smaller) than zero indicate that an individual is slower (faster) than predicted by his or her baseline speed. Specifically, the residual (i.e., baseline-corrected) RT scores reflect the speed with which individuals rate their revenge, avoidance, and benevolence, respectively.

### Transgression severity

On completion of the TRIM, participants indicated how severe they perceived their transgression to be on a face-valid, single-item scale (Ghaemmaghami et al., 2011). Specifically, they were asked: "How painful was the transgression to you at the time it occurred?" Responses were made on a 7-point scale (1 = *not painful at all*, 7 = *very painful*).

### Explicit forgiveness

Explicit forgiveness was measured with the four-item Tendency to Forgive Scale (TTF; Brown, 2003; e.g., "I tend to get over it quickly when someone hurts my feelings,"  $\alpha = 0.76$ ). Participants responded to each item on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). The TTF has consistently demonstrated good psychometric properties, including strong predictive validity across several studies (Brown, 2003; Brown and Phillips, 2005).

### Social desirability

The Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1994; German version by Musch et al., 2002) was used to assess both factors of social desirability, impression management (10 items; e.g., "I never swear,"  $\alpha = 0.76$ ) and self-deceptive enhancement (10 items; e.g., "My first impressions of people usually turn out to be right,"  $\alpha = 0.62$ ). Participants rated each item on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). We computed continuous (vs. dichotomized) subscale scores due to their superior psychometric properties (Stöber et al., 2002). The BIDR is widely used and both subscales have provided evidence of reliability and validity in past research (Paulhus, 1984, 1994; Musch et al., 2002).

### STATISTICAL ANALYSES

We analyzed bivariate correlations and conducted hierarchical multiple regression analyses using implicit and explicit forgiveness to predict residual (i.e., baseline-corrected) RT scores (i.e.,  $RT_{\text{Revenge}}$ ,  $RT_{\text{Avoidance}}$ , and  $RT_{\text{Benevolence}}$ ). All statistical tests adopted a significance level of  $\alpha = 0.05$  (two-tailed).

## RESULTS

### DESCRIPTIVE STATISTICS AND CORRELATIONS AMONG MEASURES

Participants reported transgressions pertaining to romantic partners (38.5%), close friends (37.5%), parents (10.6%), other family members or relatives (10.6%), coworker, acquaintance or neighbor (1.0%), or other person (1.9%). Table 2 details descriptive statistics for all measures.

Intercorrelations are presented in Table 3. Implicit and explicit forgiveness were not significantly correlated, a result that parallels findings on implicit-explicit relations in other research domains such as self-esteem (Bosson et al., 2000). Importantly, implicit forgiveness was unrelated to both social- desirability components. Explicit forgiveness, by contrast, correlated significantly positively with self-deceptive enhancement.

### PREDICTING RT FROM IMPLICIT AND EXPLICIT FORGIVENESS

Our main predictions were that both implicit and explicit forgiveness would predict shorter RT of forgiveness ratings (*Hypothesis 1*) and that implicit forgiveness would demonstrate increments in predictive validity (*Hypothesis 2*). To test these hypotheses, we computed three hierarchical multiple regression analyses. Each residual RT score (i.e.,  $RT_{\text{Revenge}}$ ,  $RT_{\text{Avoidance}}$ ,  $RT_{\text{Benevolence}}$ ) was treated as a separate criterion and the trait forgiveness measures were entered in two steps (Step 1: explicit forgiveness; Step 2: implicit forgiveness). As such, we were able to estimate the

**Table 2 | Means, Standard Deviations, and Actual Ranges of all Study Variables.**

Variable	<i>M</i>	<i>SD</i>	Min	Max
<b>BASELINE SPEED (S)<sup>a</sup></b>				
RT to knowledge items	5.87	1.40	3.55	10.02
<b>TRAIT FORGIVENESS</b>				
Forgiveness IAT (D <sub>1</sub> )	0.47	0.25	-0.20	1.03
TTF	3.79	1.24	1.25	6.75
<b>STATE FORGIVENESS (TRIM)</b>				
Revenge	1.66	0.66	1.00	3.80
Avoidance	2.36	1.12	1.00	5.57
Benevolence	5.89	0.76	3.50	7.00
<b>RT OF STATE FORGIVENESS RATINGS (S)<sup>a</sup></b>				
$RT_{\text{Revenge}}$	4.89	1.37	3.00	8.75
$RT_{\text{Avoidance}}$	4.57	1.11	2.81	8.06
$RT_{\text{Benevolence}}$	5.68	1.32	3.48	9.49
<b>TRANSGRESSION-RELATED CHARACTERISTICS<sup>b</sup></b>				
Closeness (IOS)	5.23	1.17	3.00	7.00
Transgression severity	4.93	1.65	1.00	7.00
<b>SOCIAL DESIRABILITY (BIDR)</b>				
Self-deceptive enhancement	4.05	0.72	2.50	6.00
Impression management	3.75	0.88	1.60	6.20

IAT, Implicit Association Test; TTF, Tendency to Forgive Scale; TRIM, Transgression-Related Interpersonal Motivations Inventory; RT, response time; IOS, Inclusion of the Other in the Self Scale; BIDR, Balanced Inventory of Desirable Responding; Possible range of self-report scores: 1–7. <sup>a</sup>For clarity of presentation, these values represent raw (mean) latencies. <sup>b</sup>Single-item measures.

**Table 3 | Intercorrelations Among Study Variables.**

Variable	1	2	3	4	5	6	7	8	9	10	11	12
<b>TRAIT FORGIVENESS</b>												
1. Forgiveness IAT (D <sub>1</sub> )	–											
2. TTF	0.00	–										
<b>STATE FORGIVENESS (TRIM)</b>												
3. Revenge	–0.04	–0.29**	–									
4. Avoidance	–0.05	–0.26**	0.37***	–								
5. Benevolence	0.01	0.28**	–0.37***	–0.51***	–							
<b>RESIDUAL RT OF STATE FORGIVENESS RATINGS<sup>a</sup></b>												
6. Residual RT <sub>Revenge</sub>	–0.22*	–0.24*	0.36***	0.22*	–0.07	–						
7. Residual RT <sub>Avoidance</sub>	0.01	–0.17 <sup>#</sup>	0.12	0.40***	–0.15	0.58***	–					
8. Residual RT <sub>Benevolence</sub>	–0.18 <sup>#</sup>	–0.27**	–0.02	0.10	–0.02	0.43***	0.54***	–				
<b>TRANSGRESSION-RELATED CHARACTERISTICS<sup>b</sup></b>												
9. Closeness (IOS)	–0.11	–0.02	–0.17 <sup>#</sup>	–0.20*	0.25*	–0.16	–0.15	–0.10	–			
10. Transgression severity	–0.18 <sup>#</sup>	–0.16	–0.09	–0.05	–0.12	0.14	0.02	0.02	0.04	–		
<b>SOCIAL DESIRABILITY (BIDR)</b>												
11. Self-deceptive enhancement	–0.10	0.23*	0.03	–0.01	0.06	–0.04	–0.05	–0.07	0.14	–0.07	–	
12. Impression management	–0.04	0.02	–0.08	–0.19 <sup>#</sup>	0.09	0.00	–0.01	0.00	0.14	0.01	0.15	–

IAT, Implicit Association Test; TTF, Tendency to Forgive Scale; TRIM, Transgression-Related Interpersonal Motivations Inventory; RT, response time; IOS, Inclusion of the Other in the Self Scale; BIDR, Balanced Inventory of Desirable Responding; Possible range of self-report scores: 1–7. <sup>a</sup>Baseline speed partialled out. <sup>b</sup>Single-item measures. <sup>#</sup>*p* < 0.10. \**p* < 0.05. \*\**p* < 0.01. \*\*\**p* < 0.001.

incremental proportion of variance in residual RT accounted for by implicit forgiveness, after controlling for the effect of explicit forgiveness. We obtained qualitatively identical results when controlling for closeness and transgression severity in the first step of the models. Also, the interaction between implicit and explicit forgiveness yielded no significant effects. In the interest of parsimony, these variables were trimmed from the final analyses. **Table 4** details the results.

The first regression model revealed that both implicit and explicit forgiveness significantly predicted residual RT<sub>Revenge</sub>. As expected, individuals high in implicit and explicit forgiveness were faster to rate their revenge motivations. Introducing implicit forgiveness in Step 2 added significantly to the prediction of residual RT<sub>Revenge</sub>, over and above explicit forgiveness. The full model explained 10.7% of variance in residual RT<sub>Revenge</sub>.

The second regression model with residual RT<sub>Avoidance</sub> as criterion revealed that explicit forgiveness predicted the time individuals needed to rate their avoidance motivations, but implicit forgiveness did not. Consequently, the addition of implicit forgiveness in Step 2 explained no variation in residual RT<sub>Avoidance</sub>. The restricted model (i.e., the model containing only explicit forgiveness as predictor) explained a small (2.9%) and marginally significant proportion of variance in the RT measure.

The model predicting residual RT<sub>Benevolence</sub> was examined last. As expected, high levels of implicit and explicit forgiveness were associated with shorter RT for items tapping benevolence. The addition of implicit forgiveness in Step 2, however, contributed only a marginally significant amount of variance to the prediction of residual RT<sub>Benevolence</sub>, beyond explicit forgiveness. The full model explained 10.2% of variance in residual RT<sub>Benevolence</sub>.

In summary, *Hypothesis 1* was generally supported for explicit forgiveness: among individuals high in explicit forgiveness, RT

**Table 4 | Hierarchical Multiple Regression Analyses Predicting Response Time by Implicit and Explicit Forgiveness.**

Predictor	Residual RT of State Forgiveness Ratings <sup>a</sup>					
	RT <sub>Revenge</sub> <sup>b</sup>		RT <sub>Avoidance</sub> <sup>c</sup>		RT <sub>Benevolence</sub> <sup>d</sup>	
	ΔR <sup>2</sup>	β	ΔR <sup>2</sup>	β	ΔR <sup>2</sup>	β
Step 1	0.060*		0.029 <sup>#</sup>		0.071**	
TTF		–0.24*		–0.17 <sup>#</sup>		–0.27**
Step 2	0.047*		0.000		0.031 <sup>#</sup>	
TTF		–0.24*		–0.17 <sup>#</sup>		–0.27**
Forgiveness IAT		–0.22*		0.01		–0.18 <sup>#</sup>

RT, response time; TTF, Tendency to Forgive Scale; IAT, Implicit Association Test; β, standardized beta coefficient; ΔR<sup>2</sup>, change in explained variance from one step to the next. <sup>a</sup>Baseline speed partialled out. <sup>b</sup>Step 1: *F*<sub>(1, 102)</sub> = 6.46, *p* = 0.013; Step 2: *F*<sub>(2, 101)</sub> = 6.05, *p* = 0.003; <sup>c</sup>Step 1: *F*<sub>(1, 102)</sub> = 3.03, *p* = 0.085; Step 2: *F*<sub>(2, 101)</sub> = 1.50, *p* = 0.227; <sup>d</sup>Step 1: *F*<sub>(1, 102)</sub> = 7.78, *p* = 0.006; Step 2: *F*<sub>(2, 101)</sub> = 5.71, *p* = 0.004. <sup>#</sup>*p* < 0.10. \**p* < 0.05. \*\**p* < 0.01.

was shorter for all three TRIM dimensions (i.e., revenge, avoidance, benevolence) and significantly so for revenge and benevolence. As regards implicit forgiveness, *Hypotheses 1* and *2* were supported for one of the three RT scores (i.e., residual RT<sub>Revenge</sub>), providing suggestive evidence for the indirect forgiveness measure in terms of incremental predictive validity.

**DISCUSSION**

In the present study, we analyzed the predictive and incremental validity of a non-self-report (indirect) measure of forgiveness, the forgiveness IAT. As forgiveness-related behavioral criteria,

we examined differential response latencies of state forgiveness ratings.

The new forgiveness IAT provides an indirect assessment of forgiveness by measuring chronically accessible self-associations pertaining to forgiving vs. vengeful. Noteworthy, we found that the IAT and the TTF scale were virtually unrelated. This may indicate that both measures tap different modes of representation (i.e., automatic vs. controlled; Moors and De Houwer, 2006) and different content (i.e., the stimuli included in both instruments). Notably, such factors (i.e., measurement correspondence and processing mode) have been shown to reduce the correlation between direct and indirect measures (Hofmann et al., 2005).

As desired for personality assessment, the forgiveness IAT demonstrated high split-half reliability, comparable with coefficients in related research on the implicit personality self-concept (Egloff and Schmukle, 2002; Back et al., 2009) and similar to those of explicit forgiveness measures (Brown, 2003; Brown and Phillips, 2005). Concerning temporal stability (i.e., test–retest reliability), IAT scores demonstrated a moderate convergence across measurement occasions. The 1-month retest correlation in the current study ( $r = 0.50$ ) was in the range typically found for IATs ( $r = 0.25–0.69$ , with mean and median estimates of 0.50; Lane et al., 2007). Given the small test–retest sample size, however, this estimate should be viewed with caution.

As to the prediction of behavior, results demonstrated that among individuals high in explicit forgiveness, RT was shorter for two of the three TRIM dimensions (i.e., revenge and benevolence). For implicit forgiveness, hypotheses were supported for one of the three RT scores (i.e., residual  $RT_{\text{Revenge}}$ ). Lack of conceptual correspondence (Ajzen and Fishbein, 1977) might explain why implicit forgiveness did not predict residual  $RT_{\text{Avoidance}}$ , given that the concept of avoidance is not directly tapped with our IAT.

Overall, our findings are consistent with well-established evidence that indirect personality measures allow prediction of variance in behavior beyond what is already predicted by self-reports (Back et al., 2009; Greenwald et al., 2009). Although not significantly related to TTF scores, the forgiveness IAT provided incremental insights into trait-relevant criteria that would have been overlooked if our focus had been restricted to self-report. This parallels recent results from the fields of need for cognition or self-esteem (Rudolph et al., 2010; Fleischhauer et al., 2013), which demonstrated that indirect measures (although not significantly associated with direct measures) incrementally explained variance in relevant behavior.

Apart from those results, some methodological considerations need to be discussed. For example, our criterion measures were based on response latencies and not on *actual* forgiving behavior. Although important to mention, prior personality research using the IAT similarly operationalized behavior with RTs. In one study on implicit self-esteem, for example, the time individuals spent reading social feedback has been taken as an objective measure of defensiveness (Schröder-Abé et al., 2007; study 2). Additionally, unlike other personality traits (e.g., neuroticism) that are comparatively easy to observe (e.g., non-verbal nervousness; Back et al., 2009), forgiveness is an intrinsically processual phenomenon that unfolds over time (McCullough et al., 2003), thus not manifested

in a specific act (Fincham, 2000). Therefore, actual behavior might be less suited for further research on implicit forgiveness.

It might be valuable, however, to examine other indices, such as physiological reactivity. Johnston et al. (2013), for instance, have recently demonstrated that implicit—but not explicit—moral identity predicted increases in blood pressure and heart rate in response to moral transgressions. Given evidence that the same physiological indicants are associated with unforgiving responses (Witvliet et al., 2001), one testable prediction is that the forgiveness IAT should predict individual variation in physiological responses to *personal* transgressions.

Together, our research has important implications for personality theory and assessment. First, it suggests that an individual's implicit cognition may influence the way he or she responds to transgression. This resonates with recent work on the role of implicit processes in forgiveness (Karremans and Aarts, 2007). Second, our findings provide an example of how direct and indirect measures complement each other in targeting implicit *and* explicit personality features, thereby yielding a more comprehensive understanding of behavior (Bornstein, 2002). Future research needs to place the construct of *implicit forgiveness* within a broader nomological network, further establishing its construct validity with respect to related constructs (e.g., empathy for the offender). Third, the IAT's resistance to social desirability underscores its utility for assessing even constructs as desirable as forgiveness. Last, the results of this study demonstrate that implicit associations may contribute a valuable, albeit understudied, source of information in understanding the implicit self-concept of forgiveness. Therefore, we emphasize the importance of administering multiple assessment formats, especially in applied (e.g., clinical) contexts. Promising work has already begun on the differential treatment sensitivity of implicit (vs. explicit) personality measures (Gamer et al., 2008) and should be further explored.

## CONCLUSION

The present study developed an indirect measure of forgiveness—the forgiveness IAT—and examined its incremental validity for the prediction of differential response latencies in a transgression-recall paradigm. We thereby complement previous work on the role of automatic processes in forgiveness and, importantly, provide an addition to existing self-reports. We found preliminary evidence that implicit forgiveness explained additional variance above and beyond a corresponding explicit measure. Together, our findings provide novel insights that will hopefully stimulate both research and practice of implicit personality functioning.

## ACKNOWLEDGMENTS

We wish to thank our reviewers whose constructive comments helped to improve this article. The research was financially supported by the Cluster of Excellence “Languages of Emotion” at Freie Universität Berlin, funded by the DFG (German Research Foundation).

## REFERENCES

- Ajzen, I., and Fishbein, M. (1977). Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychol. Bull.* 84, 888–918. doi: 10.1037/0033-2909.84.5.888

- Allemand, M., Steiner, M., and Hill, P. L. (2013). Effects of a forgiveness intervention for older adults. *J. Couns. Psychol.* 60, 279–286. doi: 10.1037/a0031839
- Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Pers. Soc. Psychol.* 63, 596–612. doi: 10.1037/0022-3514.63.4.596
- Back, M. D., Schmukle, S. C., and Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *J. Pers. Soc. Psychol.* 97, 533–548. doi: 10.1037/a0016229
- Banise, R. (2001). Affective priming with liked and disliked persons: prime visibility determines congruency and incongruency effects. *Cogn. Emot.* 15, 501–520. doi: 10.1080/02699930126251
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *J. Pers. Assess.* 78, 47–68. doi: 10.1207/s15327752JPA7801\_04
- Bosson, J. K., Swann, W. B., and Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *J. Pers. Soc. Psychol.* 79, 631–643. doi: 10.1037/0022-3514.79.4.631
- Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: construct validity and links with depression. *Pers. Soc. Psychol. Bull.* 29, 759–771. doi: 10.1177/0146167203029006008
- Brown, R. P., and Phillips, A. (2005). Letting bygones be bygones: further evidence for the validity of the tendency to forgive scale. *Pers. Individ. Diff.* 38, 627–638. doi: 10.1016/j.paid.2004.05.017
- Egloff, B., and Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *J. Pers. Soc. Psychol.* 83, 1441–1455. doi: 10.1037/0022-3514.83.6.1441
- Egloff, B., Schwerdtfeger, A., and Schmukle, S. C. (2005). Temporal stability of the implicit association test-anxiety. *J. Pers. Assess.* 84, 82–88. doi: 10.1207/s15327752jpa8401\_14
- Fazio, R. (1990a). Multiple processes by which attitudes guide behavior: the MODE Model as an integrative framework. *Adv. Exp. Soc. Psychol.* 23, 75–109.
- Fazio, R. H. (1990b). “A practical guide to the use of response latency in social psychological research,” in *Research Methods in Personality and Social Psychology*, eds C. Hendrick and M. S. Clark (Thousand Oaks, CA: Sage Publications, Inc.), 74–97.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cogn. Emot.* 15, 115–141. doi: 10.1080/02699930125908
- Fincham, F. D. (2000). The kiss of the porcupines: from attributing responsibility to forgiving. *Pers. Relat.* 7, 1–23. doi: 10.1111/j.1475-6811.2000.tb00001.x
- Fincham, F. D. (2010). “Forgiveness: integral to a science of close relationships?” in *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature*, eds M. Mikulincer and P. R. Shaver (Washington, DC: American Psychological Association), 347–365. doi: 10.1037/12061-018
- Fincham, F. D., Hall, J., and Beach, S. R. H. (2006). Forgiveness in marriage: current status and future directions. *Family Relat.* 55, 415–427. doi: 10.1111/j.1741-3729.2005.callf.x-11
- Fincham, F. D., Hall, J. H., and Beach, S. R. H. (2005). “Til lack of forgiveness doth us part: forgiveness in marriage,” in *Handbook of Forgiveness* ed E. L. Worthington (New York, NY: Routledge), 207–226.
- Fleischhauer, M., Strobel, A., Enge, S., and Strobel, A. (2013). Assessing implicit cognitive motivation: developing and testing an implicit association test to measure need for cognition. *Eur. J. Pers.* 27, 15–29. doi: 10.1002/per.1841
- Gamer, J., Schmukle, S. C., Luka-Krausgrill, U., and Egloff, B. (2008). Examining the dynamics of the implicit and the explicit self-concept in social anxiety: changes in the implicit association test-anxiety and the social phobia anxiety inventory following treatment. *J. Pers. Assess.* 90, 476–480. doi: 10.1080/00223890802248786
- Ghaemmaghami, P., Allemand, M., and Martin, M. (2011). Forgiveness in younger, middle-aged and older adults: age and gender matters. *J. Adult Dev.* 18, 192–203. doi: 10.1007/s10804-011-9127-x
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216. doi: 10.1037/0022-3514.85.2.197
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., and Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* 97, 17–41. doi: 10.1037/a0015575
- Hofmann, W., Gschwendner, T., Nosek, B. A., and Schmitt, M. (2005). What moderates implicit-explicit consistency? *Eur. Rev. Soc. Psychol.* 16, 335–390. doi: 10.1080/10463280500443228
- Hoyt, W. T., and McCullough, M. E. (2005). “Issues in the multimodal measurement of forgiveness,” in *Handbook of Forgiveness*, ed E. L. Worthington (New York, NY: Routledge), 109–123.
- Johnston, M. E., Sherman, A., and Grusec, J. E. (2013). Predicting moral outrage and religiosity with an implicit measure of moral identity. *J. Res. Pers.* 47, 209–217. doi: 10.1016/j.jrp.2013.01.006
- Karremans, J. C., and Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *J. Exp. Soc. Psychol.* 43, 902–917. doi: 10.1016/j.jesp.2006.10.012
- Karremans, J. C., and Lange, P. A. M. V. (2008). Forgiveness in personal relationships: its malleability and powerful consequences. *E. Rev. Soc. Psychol.* 19, 202–241. doi: 10.1080/10463280802402609
- Lane, K. A., Banaji, M. R., Nosek, B. A., and Greenwald, A. G. (2007). “Understanding and using the Implicit Association Test: IV: What We Know (So Far) about the Method,” in *Implicit Measures of Attitudes*, eds B. Wittenbrink, N. Schwarz, B. Wittenbrink, and N. Schwarz (New York, NY: Guilford Press), 59–102.
- Lischetzke, T., Angelova, R., and Eid, M. (2011). Validating an indirect measure of clarity of feelings: evidence from laboratory and naturalistic settings. *Psychol. Assess.* 23, 447–455. doi: 10.1037/a0022211
- Lischetzke, T., Cuccodoro, G., Gauger, A., Todeschini, L., and Eid, M. (2005). Measuring affective clarity indirectly: individual differences in response latencies of state affect ratings. *Emotion* 5, 431–445. doi: 10.1037/1528-3542.5.4.431
- McCullough, M. E., Fincham, F. D., and Tsang, J.-A. (2003). Forgiveness, forbearance, and time: the temporal unfolding of transgression-related interpersonal motivations. *J. Pers. Soc. Psychol.* 84, 540–557. doi: 10.1037/0022-3514.84.3.540
- McCullough, M. E., and Hoyt, W. T. (2002). Transgression-related motivational dispositions: personality substrates of forgiveness and their links to the big five. *Pers. Soc. Psychol. Bull.* 28, 1556–1573. doi: 10.1177/014616702237583
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L. Jr., Brown, S. W., and Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *J. Pers. Soc. Psychol.* 75, 1586–1603. doi: 10.1037/0022-3514.75.6.1586
- McCullough, M. E., Root, L. M., and Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *J. Consult. Clin. Psychol.* 74, 887–897. doi: 10.1037/0022-006X.74.5.887
- Moors, A., and De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychol. Bull.* 132, 297–326. doi: 10.1037/0033-2909.132.2.297
- Musch, J., Brockhaus, R., and Bröder, A. (2002). Ein inventar zur erfassung von zwei faktoren sozialer erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica* 48, 121–129. doi: 10.1026//0012-1924.48.3.121
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *J. Pers. Soc. Psychol.* 46, 598–609. doi: 10.1037/0022-3514.46.3.598
- Paulhus, D. L. (1994). *Reference Manual for BIDR Version 6*. Vancouver, BC: University of British Columbia.
- Perugini, M., and Banise, R. (2007). Personality, implicit self-concept and automaticity. *Eur. J. Pers.* 21, 257–261. doi: 10.1002/per.637
- Röhner, J., Schröder-Abé, M., and Schütz, A. (2011). Exaggeration is harder than understatement, but practice makes perfect! Faking success in the IAT. *Exp. Psychol.* 58, 464–472. doi: 10.1027/1618-3169/a000114
- Rudolph, A., Schröder-Abé, M., Riketta, M., and Schütz, A. (2010). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behavior, but not self-reported or controlled behavior. *J. Psychol.* 218, 12–19. doi: 10.1027/0044-3409/a000003
- Schmukle, S. C., and Egloff, B. (2006). Assessing anxiety with extrinsic Simon tasks. *Exp. Psychol.* 53, 149–160. doi: 10.1027/1618-3169.53.2.149
- Schnabel, K., and Asendorpf, J. B. (2010). “The self-concept: new insights from implicit measurement procedures,” in *Handbook of Implicit Social Cognition: Measurement, Theory, And Applications*, eds B. Gawronski and B. K. Payne (New York, NY: Guilford Press), 408–425.
- Schröder-Abé, M., Rudolph, A., Wiesner, A., and Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *Int. J. Psychol.* 42, 174–183. doi: 10.1080/00207590601068134

- Steffens, M. C. (2004). Is the implicit association test immune to faking? *Exp. Psychol.* 51, 165–179. doi: 10.1027/1618-3169.51.3.165
- Stöber, J., Dette, D. E., and Musch, J. (2002). Comparing continuous and dichotomous scoring of the balanced inventory of desirable responding. *J. Pers. Assess.* 78, 370–389. doi: 10.1207/S15327752JPA7802\_10
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803\_1
- Sutton, G. W., and Thomas, E. K. (2006). Restoring christian leaders: how conceptualizations of forgiveness and restoration used in empirical studies can influence practice and research. *Am. J. Pastor. Counsel.* 8, 27–42. doi: 10.1300/J062v08n02\_02
- Witvliet, C. V., Ludwig, T. E., and Vander Laan, K. L. (2001). Granting forgiveness or harboring grudges: implications for emotion, physiology, and health. *Psychol. Sci.* 12, 117–123. doi: 10.1111/1467-9280.00320
- Worthington, E. L. Jr., and Wade, N. G. (1999). The psychology of unforgiveness and forgiveness and implications for clinical practice. *J. Soc. Clin. Psychol.* 18, 385–418.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 April 2014; accepted: 23 June 2014; published online: 11 July 2014.

Citation: Fatfouta R, Schröder-Abé M and Merkl A (2014) Forgiving, fast and slow: validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Front. Psychol.* 5:728. doi: 10.3389/fpsyg.2014.00728

This article was submitted to *Personality and Social Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Fatfouta, Schröder-Abé and Merkl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.