European Journal of Social Psychology
Eur. J. Soc. Psychol. **38**, 977–997 (2008)
Published online 14 January 2008 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.487

Reliability and validity of the Single-Target IAT (ST-IAT): Assessing automatic affect towards multiple attitude objects

MATTHIAS BLUEMKE[1]* AND MALTE FRIESE[2]
[1]*Psychological Institute, University of Heidelberg, Heidelberg, Germany*
[2]*Institute of Psychology, University of Basel, Basel, Switzerland*

Abstract

In contrast to the original Implicit Association Test (IAT), the Single-Target Implicit Association Test (ST-IAT) measures the evaluation of a target object without the need to simultaneously evaluate a counter-category. The present research investigates (a) whether position within a series of several ST-IATs affects reliability and validity, and (b) whether the ST-IAT exhibits adequate construct validity if the target objects are closely interrelated. We address these questions by taking five interrelated yet distinct political parties in Germany as an exemplary domain. The ST-IAT reliably and validly assessed attitudes towards political parties (Study 1). Serial position effects did not affect the results. The ST-IATs mostly captured a specific party evaluation and exhibited discriminant validity. At the same time, discriminant validity was limited among parties within one wing of the political left–right spectrum that underlies implicit and explicit party evaluations (Study 2). If used with caution, the ST-IAT can be a valuable supplement to implicit measures in the case of multiple single-target assessments. Copyright © 2008 John Wiley & Sons, Ltd.

The aim of this paper is to illuminate the usefulness of a new implicit measure tailored to assess automatic affective reactions—the Single-Target Implicit Association Test (ST-IAT; Wigboldus, Holland, & van Knippenberg, 2004; cf. Karpinski & Steinman, 2006)—and to explore the methodological properties of this recent offspring of the well-known Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). In particular, we investigate whether the ST-IAT can achieve what has not been demonstrated for other implicit tools so far, namely, to reliably and validly assess the evaluation of several (five) interrelated target objects—while remaining resistant to fatigue or exercise effects in a line of repeated measurements.

Scientists have developed several indirect approaches to tapping automatic affective components that are important for guiding spontaneous behaviour (Fazio & Olson, 2003). The IAT has gained particular support as a tool for the assessment of such spontaneous evaluations. While acknowledging the IAT's valuable features (flexibility, reliability, validity; Nosek, Greenwald, & Banaji, 2006), it has been pointed out that (a) IATs provide only an ambiguous answer to the question of the absolute evaluation of target concepts (Blanton, Jaccard, Gonzales, & Christie, 2006; Fiedler, Messner, & Bluemke, 2006; Nosek, Greenwald, & Banaji, 2005), (b) the choice of a counter-category against which the target object of interest is contrasted may pose a natural complement (e.g. *men* vs. *women*), but in many cases the choice of category is highly subjective (compare a *liberal/conservative* IAT to a *liberal/socialist* IAT; Karpinski, 2004) and (c) the evaluation of multiple target concepts such as social groups within a multi-ethnic nation (e.g. White vs. Asian Americans, White vs. African Americans, African vs. Asian Americans; Devos & Banaji, 2005) requires numerous pairwise comparisons for a complete picture. Taking Germany's complex political landscape as another example, one would need to compare liberals to conservatives, social-democrats, socialists and environmentalists at the same time.

*Correspondence to: Matthias Bluemke, Psychological Institute, University of Heidelberg, Hauptstrasse 47-51, 69117 Heidelberg, Germany.
E-mail: matthias.bluemke@psychologie.uni-heidelberg.de

Received 16 November 2006
Accepted 21 October 2007

Several researchers have taken steps to overcome these limitations of the IAT. Among the efforts developed as non-relative measures are the Go/No-Go Association Task (GNAT; Nosek & Banaji, 2001), the Extrinsic Affective Simon Task (EAST; De Houwer, 2003), the newly developed identification-EAST (ID-EAST; De Houwer & De Bruycker, 2007a) and IATs using *neutral*, <blank>, non-words, a negated category (John—*not* John) or an unrelated concept (e.g. *tree*) as the counter-category (Brendl, Markman, & Messner, 2001; Czopp, Monteith, Zimmerman, & Lynam, 2004; De Jong, van den Hout, Rietbroek, & Huijding, 2003; Kim, 2004; Zayas & Shoda, 2005; see also section 'General Discussion'). Despite the advances, either methodological problems remain (interpretational ambiguities due to arbitrary counter-categories; insufficient reliability) or the convenience of common IAT applications is lost (e.g. signal detection outcomes that are difficult to compare with latency indices). Interestingly, a prominent approach to assessing single-target associations comes from a modification of the traditional IAT procedure.

THE SINGLE-TARGET IAT AND ITS ADVANTAGES

By abandoning one of the target categories and keeping two attribute categories, Wigboldus et al. (2004) derived a single-target IAT variant. For instance, they assessed the association of the category *Islamic* towards positive and negative valence without applying a counter-category (e.g. *Christian*). Generally, the ST-IAT proceeds with the following three steps (cf. Table 1): After a single discrimination block of the evaluative stimuli, a second block of trials follows, mapping target stimuli and positive items onto one response key and negative stimuli onto the other, before finally an inverted response key assignment maps target stimuli and negative items together. Whereas the IAT yields a preference index relative to a contrast concept, the ST-IAT reduces such an arbitrary influence in the evaluation of a target category and raises hopes of obtaining higher criterion correlations.

Unsurprisingly, researchers have shown growing interest in the ST-IAT as a measure that aims to tap into associative structures (e.g. Friese, Hofmann, & Wänke, in press; Hofmann, Rauch, & Gawronski, 2007; Penke, Eichstaedt, & Asendorpf, 2006; Richetin, Perugini, Adjali, & Hurling, 2007). For instance, Karpinski and Steinman (2006) applied a conceptually similar procedure, termed the Single-Category-IAT (SC-IAT). The SC-IAT is distinct from the ST-IAT in applying a response window of 1500 or 2000 milliseconds maximum latency that supports a sense of urgency for quick responding (feedback had to be provided in less than 1% of trials). Furthermore, whereas Wigboldus et al. (2004) included only the target trials for the analysis of ST-IAT latencies (and found lower reliability than is typical for IATs), the SC-IAT relied on the full power of both target and attribute stimuli for the calculation of the index and obtained sufficient reliability. Overall, there are more similarities than dissimilarities between the ST-IAT and the SC-IAT, the differences between them are not of crucial concern for the purposes of our research, and we reserve the name SC-IAT for the response window procedure.

Table 1. Category assignment and stimulus proportions across ST-IAT blocks for an exemplary participant

				Number of stimuli		
Block	Task description	Left key concepts	Right key concepts	Positive	Negative	Party
1	Evaluative training trials	Positive	Negative	10	10	—
2	Initial block	Positive + CDU	Negative	10	15	10
3	Reversed block	Positive	Negative + CDU	15	10	10
4	Initial block	Negative	Positive + SPD	10	15	10
5	Reversed block	Negative + SPD	Positive	15	10	10
6	Initial block	Positive + FDP	Negative	10	15	10
7	Reversed block	Positive	Negative + FDP	15	10	10
8	Initial block	Negative	Positive + PDS	10	15	10
9	Reversed block	Negative + PDS	Positive	15	10	10
10	Initial block	Positive + GREEN	Negative	10	15	10
11	Reversed block	Positive	Negative + GREEN	15	10	10

DOI: 10.1002/ejsp

OPEN RESEARCH QUESTIONS

With the ST-IAT procedure an efficient evaluation of several target objects in succession is possible. Several topics pertaining to the psychometric status of the ST-IAT have not been addressed so far, though they directly relate to the usefulness of ST-IAT applications and have implications for IAT research in general: (a) whether the psychometric properties of ST-IATs are affected by position effects (fatigue or exercise effects), particularly when more than two target objects are evaluated; and (b) whether the ST-IAT exhibits adequate construct validity in terms of convergent and discriminant validity if the target categories constitute a class of several interrelated target objects. We briefly discuss the importance of these methodological questions and their practical implications.

Psychometric Properties of Multiple ST-IAT Assessments

To date, we know little about the reliability and validity of ST-IATs, especially when several ST-IATs are assessed in a single session. First, Karpinski and Steinman (2006) explicitly questioned the ST-IAT's reliability despite its similarities with the SC-IAT, raising the question of whether response windows are a crucial feature of single-target assessments if they are to reach satisfying levels of reliability. Second, caution is even more warranted as the ST-IAT's simpler task structure (compared to the original IAT) might foster strategic processing. Effects of original IATs decrease across multiple measurement occasions (Greenwald, Nosek, & Banaji, 2003) and they sometimes depend on the serial position of the task within the course of multiple implicit measures (Steffens, 2005). Serial position could affect ST-IAT effects as well (Wigboldus et al., 2004). The ST-IAT's strength of efficiently assessing several single-target evaluations would then be constricted by severe disadvantages, either because of response strategies, fatigue effects or specific carry-over effects (Nosek et al., 2006; Steffens & Schulze-König, 2006). Following Schnabel, Banse, and Asendorpf's (2006) call for an analysis of serial position effects of implicit measures, we will test the hypothesis that serial position adversely affects means and validities of ST-IATs. As regards reliability, we will analyse internal consistency and—for the first time—provide estimates of test-retest reliability of ST-IATs.

Construct Validity

Convergent and discriminant validity of implicit measures have often been demonstrated with few target objects, yet mostly when those objects were quite distinct from each other (e.g. German–Turk and German–Asian IATs; Gawronski, 2002). Suppose a researcher wants to measure the associations towards four groups of immigrants. Two of the groups may be better liked than the other two because the groups may have come from two different continents. The researcher might choose the ST-IAT as a means to measure the evaluations of the groups in a non-relative way. Would the ST-IAT still capture affective tendencies specifically for each group? We consider it important to explore the specificity, or discriminant validity, of ST-IATs pertaining to several related, but still distinct target objects.

Overview of Studies

To answer the aforementioned questions, which are all related to the psychometric status of the ST-IAT, we illustrate the ST-IAT's properties within the domain of political preferences in Germany. Political parties served as target objects. Study 1 concerns the reliability, stability and construct validity of ST-IATs in the German parliamentary elections in 2002 and the impact of ST-IAT position on the psychometric properties. Study 2 further delves into the ST-IAT's construct validity by demonstrating relations to a political left-wing/right-wing continuum.

STUDY 1: RELIABILITY AND VALIDITY OF THE ST-IAT

During the 3 months leading up to the 2002 German parliamentary elections (first phase) we examined participants' political preferences by explicitly asking them for their evaluations of the five major parties in the German Parliament,

DOI: 10.1002/ejsp

including their intention to vote for a specific party, and we additionally measured participants' attitudes by means of five party-specific ST-IATs. Voting behaviour was assessed during the 2 weeks following the election (second phase). A sub-sample yielded estimates of retest reliabilities.[1]

Political Spectrum

Common to both studies is the spectrum of the political parties in Germany. Five parties gained seats in Parliament in the 1998 election and could be expected to do so again in 2002. Two of them are usually strong: the conservative Christian-Democratic and Christian-Social alliance (CDU/CSU) and the Social-Democratic Party (SPD). The Liberals (FDP) and the Green Party (Bündnis90/Die Grünen; GREEN) received fewer votes, but at least one of them is required by the big parties to form a coalition in order to hold the majority of Parliament seats. A fifth party entered the stage after German reunification—the successors of the former East German socialist party (PDS, now 'Die Linke'). Although it represents a crude ranking, it has been shown by using multi-dimensional scaling that these parties can be ordered according to their presumed position on a political left–right continuum (PDS, GREEN, SPD, FDP, CDU/CSU; Pappi, 1983; von Collani & Blank, 2003).

Hypotheses

In terms of construct validity we expected the intercorrelations of five ST-IATs to mirror the intercorrelations of explicit attitude measures. Convergent validity of a given ST-IAT would be reflected in its correlation with the respective explicit attitude measure. Particularly strong evidence for convergent validity would be evident if the strongest correlations emerged between the ST-IAT of a given party and the explicit measure of this party. Based on previous research in the political domain we expected medium to strong implicit–explicit correlations (Nosek et al., 2005; cf. Karpinski, Steinman, & Hilton, 2005 and Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005).

Discriminant validity is difficult to establish in the case of several interrelated political groups. It should be evident in lower correlations between a party-specific ST-IAT and explicit measures of attitudes towards other parties. Due to the ideological overlap of the parties, we anticipated positive intercorrelations among attitude measures of politically related parties, that is, among the left-wing (SPD, GREEN, PDS) and among the right-wing (CDU/CSU and FDP) parties, but negative correlations between parties across the left-wing/right-wing border (Conover & Feldman, 2004). Taken together, the ST-IAT should discriminate between parties from different wings and optimally between parties from the same political wing.

Finally, we tested whether position effects would weaken reliability and validity. It could well be that the internal consistencies and/or the correlation between the ST-IATs on the one hand and explicit attitudes as well as behavioural criteria on the other hand would continually decline across position. To test if either fatigue or exercise effects are responsible for any changes across the measurement procedure, we additionally analysed the mean response latencies.

METHOD

Participants

All data collection was carried out on the Internet because of a better chance of recruiting a larger and politically more diverse sample (e.g. finding supporters of parties receiving low percentages of votes) than a lab study would have allowed.

[1]Other data of this study were presented in Friese, Bluemke, and Wänke (2007) with a focus on the moderation of the implicit–explicit consistency by attitude importance and the incremental validity of the ST-IAT for voting behaviour. The analyses of those data are not repeated. Aspects that are irrelevant for the present purpose are not reported here. For further details on recruitment, drop-out, technical aspects and the respective analyses the reader is referred to this literature.

DOI: 10.1002/ejsp

We supplied a lottery of 15 vouchers for an Internet shop (worth 10€ each) for participants in the first phase and a chance to win a city tour (value 100€) for respondents in the second phase.

In the first phase, complete responses were obtained from 2556 participants, of whom 1753 (69%) returned for the second phase (67.7% male). A small number of participants (101) took part in the optional retest analysis. Control questions indicated diversity in socio-demographic variables (education and profession, $M_{age} = 31.48$ years, $SD = 11.47$). In terms of regional provenance, participants responding from each of the 16 German federal states represented the proportions in the population, including the proportion of East and West Germans.

Procedure and Materials

First Phase

The study was made available 3 months before the German parliamentary elections (22 September 2002). After some initial information about the study purpose, the second page requested comprehensive socio-demographic data. Then, we assessed the crucial variables: (a) *explicit measures of attitudes* towards CDU/CSU, FDP, SPD, GREEN and PDS ('How do you evaluate the following parties overall?'; 8-point rating scale ranging from 'very negative' to 'very positive') presented in a fixed order that was common in public opinion polls; (b) *voting intention* ('Whom will you vote for in the German parliamentary elections on 22 September 2002?'; response options: CDU/CSU, SPD, FDP, GREEN, PDS; other party; I don't know yet; I will not vote; I will cast an invalid ballot; I will not be allowed to vote); (c) *implicit measures of attitudes* (*ST-IATs*) with CDU/CSU, FDP, SPD, GREEN and PDS as categories applied in random order.

With regard to the order of implicit and explicit measures, a meta-analytical finding suggests that there is no systematic impact of order on implicit–explicit correlations (Hofmann et al., 2005). As regards the order of the implicit measures themselves, a fixed order of ST-IATs would be helpful for establishing correlations, but it would also prevent the analysis of position effects and hinder the comparison of the magnitude of ST-IAT effects across the five parties.

Succeeding questions checked for any disturbances, interruptions or lost motivation during the course of the assessment. We asked for informed consent and an e-mail address—stored separately from attitudinal data—in order to invite participants for the second phase. We debriefed participants and provided feedback on the sample's average results of the ST-IATs. Overall, participation in the first phase of the study took about 20 minutes.

Second Phase

After the election date, we invited participants via e-mail to answer a 1-minute online questionnaire. Participants logged in with an anonymous code used for matching the two data sets and supplied an answer to the criterion variable *voting behaviour*: 'Which party did you vote for in the 2002 German parliamentary elections?' with response options being: CDU/CSU, SPD, FDP, GREEN, PDS, other party; I did not vote; I cast an invalid ballot; I was not allowed to vote. After some control questions, participants had the option to voluntarily partake in a retest of explicit and implicit measures of attitudes, or to instantly receive individual results from Phase 1.

ST-IATs

A Java applet presented the stimuli and stored the latencies (accuracy dependent on a participant's local timing resolution) before sending the data back to the Web server after completion. The interstimulus interval after correct responses was set to 300 milliseconds, and a red cross in the lower part of the screen alerted participants to incorrect responses. Participants started with 20 training trials for the evaluative words prior to the first combined block. Each of the five consecutive ST-IATs, presented in individually randomized order, consisted of two combined blocks (first *party + positive*, then *party + negative*). The task was always explained to participants ahead of each block, and the category labels, which were visible at the top of the screen, served as a reminder. We balanced the side on which the first party was presented (right or

DOI: 10.1002/ejsp

left response key) *between* participants. In addition, for each party the assignment of the evaluative categories to the left and right side switched *within* participants to reduce the likelihood of strategic recoding of the tasks (cf. Table 1).

We used five different words for the positive and negative categories as well as five stimuli for each political party, represented by party emblems, images and names of well-known party members (cf. Appendix). Each stimulus was presented at least twice, adding up to 35 trials per combined block. Target stimuli, coupled and uncoupled evaluative stimuli occurred in a ratio of 10:10:15 trials, leading to a proportion of left-hand and right-hand responses of 4:3 in one and 3:4 in the other combined block (cf. Table 1).[2] Additionally, we applied fixed sequences of target and attribute stimuli so that task-switching costs across the combined blocks, target parties and participants were equal (cf. Mierke & Klauer, 2003). This enabled us to compare both ST-IAT means across parties and interindividual differences among participants.

RESULTS AND DISCUSSION

Explicit Measures of Attitudes and Voting Intentions

For a direct comparison of explicitly and implicitly measured attitudes, we rescaled the explicit ratings by subtracting the midpoint of the 8-point rating scales as a neutral reference point from the explicit likings and dividing the result by the average standard deviation of all likings. On average participants held neutral attitudes or disliked the parties, with the exceptions being SPD and GREEN, which finally won the elections by forming a coalition, $F(4, 7008) = 280.2$, $p < .001$ (see Figure 1). This tendency is also evident in the self-reported voting intention. The majority of participants were willing to vote for SPD (28.2%) and GREEN (22.5%), while CDU/CSU (18.0%), FDP (13.6%) and PDS (5.6%) clearly attracted fewer voters ($\chi^2_8 = 1305.4$, $p < .001$).

As predicted, participants had attitudes that indicated a left-wing or right-wing preference. People who liked a left-wing party (e.g. SPD) were also more likely to express support for another party of the political left (GREEN, PDS). They were

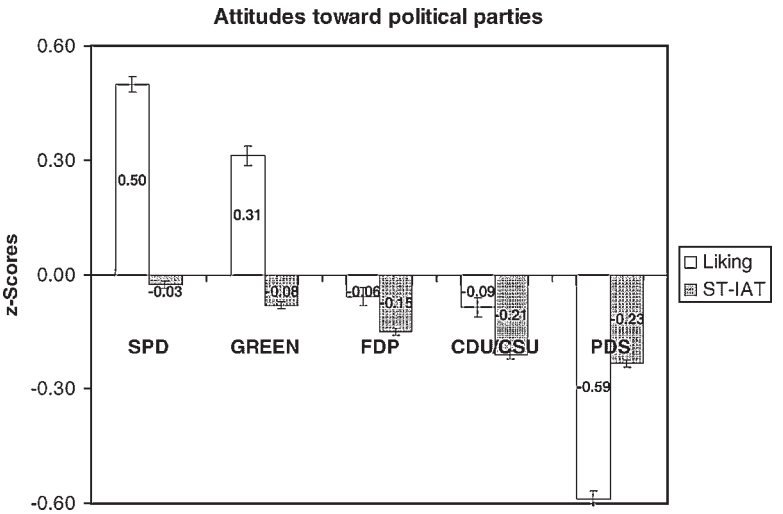

Figure 1. Means of measures of attitudes towards political parties before the German parliamentary elections in 2002 (Study 1). *Liking* refers to explicitly assessed attitudes with rating scales (*z*-like standardization on scale-midpoint); *ST-IAT* refers to implicitly assessed attitudes by means of Single-Target Implicit Association Tests (ST-IAT effects)

[2]Wigboldus et al. (2004) equalled the number of left-hand and right-hand responses, which led to a confound of positive and negative block with the number of valence trials in these blocks. By contrast, Richetin et al. (2007) equalled the number of valence stimuli, which confounds positive and negative block with the number of left-hand and right-hand responses. We used a stimulus set up in-between both of these options and similar to the one used by Karpinski and Steinman (2006).

DOI: 10.1002/ejsp

less likely to prefer parties from the right, resulting in negative correlations between left- and right-wing parties. We discovered the same pattern for participants supportive of right-wing parties. Taken together, these results show a clear-cut ideological distinction in the intercorrelations of explicit measures (see upper half of Table 2).

Single-Target IATs

Data Preparation of Latencies

As in previous IAT research, we skipped error trials and recoded latencies below 300 milliseconds and above 3000 milliseconds to the respective values. We controlled for irrelevant differences in response latency level and latency variability between participants by individually z-transforming participants' latencies, thereby reducing method variance substantially (McFarland & Crouch, 2002; Mierke & Klauer, 2003). We thus subtracted an individual's mean response latency (across all 10 ST-IAT blocks excluding the attribute training trials) from a given reaction time and divided by the individuals' standard deviation (across the response latencies of all 10 ST-IAT blocks except for the attribute training trials). This method is similar to Greenwald et al.'s (2003) *D*-algorithm save for the treatment of error penalties and training trials in the calculation of means and standard deviations (Bluemke & Friese, 2006). We dropped the first trial of each block as a training trial requiring orientation, because it typically deviated strongly from the grand mean (z values > 1). Furthermore, we omitted 10.6% of the participants who had committed 20% or more errors in at least one of the 10 combined blocks ($M = 3.9\%$), leaving 1568 data sets for ST-IAT analyses. Seventy-nine retest participants had less than 20% errors in each of the 20 ST-IAT blocks at both occasions.

In line with the calculation of IAT scores, subtracting the mean latency in the *party + positive* block from the *party + negative* block yielded the ST-IAT effects. A positive score indicates that a participant associated a political *party* and *positive* faster than the same political party paired with *negative*, which can be interpreted as a favourable spontaneous reaction to the respective party.

Analyses of Mean ST-IAT Effects and Mean Block Latencies

Although the ST-IATs varied less than explicit measures, the rank ordering of the parties was identical to those based on explicit ratings (cf. Figure 1). Assuming that the ST-IATs were well calibrated on a neutral point (none or equally positive and negative associations), then on average all of the parties were automatically evaluated negatively. We additionally checked if position effects attenuated means. One-factorial analyses of variance (ANOVAs) on the ST-IAT effects confirmed that, except for the GREEN ST-IAT, there was significant variability between positions, though only to a small extent (cf. Table 3). We tested if this variability would fit the hypothesis of linear attenuation across positions (curve-fitting regression analysis). With the exception of the GREEN Party, the ST-IAT effects were significantly less pronounced (i.e. less negative) towards the end of the session, as evident in the positive β-weights of the trend analysis.

We do not interpret this linear trend as a meaningful increase in positivity towards the parties across the procedure. Rather, we suspect that exercise gains diminished differences between the mean block latencies that existed in the beginning of the procedure. According to a linear contrast, $F(1, 1567) = 948.61$, $p < .001$, $\eta^2 = .38$, the attenuation seems to be due to a linear increase in response speed across positions 1–5 in both positive and negative blocks, Ms = 0.053, 0.096, −0.096, −0.084, −0.143. As evident in Figure 2, latencies in the faster (negative) block could not accelerate to the same extent as latencies in the positive block (slopes = −0.048 vs. −0.066), a floor effect led to the shifting IAT effects. In sum, these data cast doubt on any absolute interpretations of effect sizes when ST-IATs are repeatedly applied within an experimental session, and even more so when not counterbalancing the order of the measures.

Reliability

We calculated Cronbach's α based on a full 34-item scale of trial-wise latency-differences (with missing data being replaced by participants' mean latencies, i.e. a z-score of zero). These reliability estimates amounted to a mean α of .69 (1st

DOI: 10.1002/ejsp

Table 2. Intercorrelations of explicit and implicit measures of political attitudes before and after the German parliamentary elections in 2002 (Study 1), corrected for measurement error of implicit measures

		First phase										Second phase									
		Explicit liking					ST-IAT					Explicit liking					ST-IAT				
	α	CDU	FDP	SPD	GREEN	PDS	CDU	FDP	SPD	GREEN	PDS	CDU	FDP	SPD	GREEN	PDS	CDU	FDP	SPD	GREEN	PDS
First phase																					
Explicit liking ($N = 1753$)																					
CDU/CSU	—	—																			
FDP	—	.59	—																		
SPD	—	−.46	−.30	—																	
GREEN	—	−.58	−.43	.67	—																
PDS	—	−.49	−.37	.29	.40	—															
ST-IAT ($N = 1568$)																					
CDU/CSU	.75	**.56**	.36	−.38	−.45	−.37	—														
FDP	.66	.28	**.42**	−.19	−.25	−.21	.41	—													
SPD	.70	−.39	−.32	**.43**	.38	.20	−.32	−.09	—												
GREEN	.68	−.40	−.34	.30	**.51**	.26	−.27	−.11	.45	—											
PDS	.67	−.23	−.24	.13	.20	**.44**	−.20	−.06	.27	.34	—										
Second phase																					
Explicit liking ($N = 101$)																					
CDU/CSU	—	***.89***	.61	−.55	−.68	−.54	.65	.24	−.38	−.58	−.31	—									
FDP	—	.62	***.76***	−.41	−.49	−.28	.49	.47	−.38	−.46	−.14	.66	—								
SPD	—	−.48	−.35	***.82***	.68	.22	−.35	−.47	.39	.31	−.01	−.61	−.53	—							
GREEN	—	−.64	−.53	.69	***.87***	.36	−.60	−.42	.47	.62	.04	−.78	−.62	.78	—						
PDS	—	−.54	−.39	.25	.37	***.83***	−.41	−.20	.32	.40	.56	−.56	−.35	.21	.40	—					
ST-IAT ($N = 79$)																					
CDU/CSU	.74	.51	.24	−.43	−.45	−.31	***.48***	.13	−.17	−.25	−.07	**.55**	.31	−.40	−.47	−.26	—				
FDP	.67	.33	.48	−.18	−.26	−.17	.40	***.32***	−.40	−.40	−.06	.27	**.50**	−.17	−.37	−.26	.17	—			
SPD	.77	−.46	−.29	.59	.34	.28	−.36	−.26	***.62***	.39	−.10	−.47	−.43	**.55**	.46	.36	−.42	−.28	—		
GREEN	.70	−.53	−.46	.38	.65	.35	−.56	−.22	.56	***.48***	−.03	−.53	−.45	.40	**.67**	.43	−.31	−.47	.55	—	
PDS	.74	−.34	−.20	.15	.23	.58	−.16	−.17	.07	.31	***.50***	−.39	−.26	.17	.25	**.64**	−.20	−.18	.30	.37	—

Note: $N_1 = 1568$–1753; all $ps < .001$; $N_2 = 79$–101; $ps < .05$ for $rs > .22$.

DOI: 10.1002/ejsp

Table 3. Means and reliabilities of ST-IAT effect sizes across measurement position (Study 1)

	ST-IAT effect													
		By position					Effect of position on ST-IAT (overall *F*-test)			Linear trend analysis				
Political party	Overall	1	2	3	4	5	$F(4, 1563)$	p	η^2	B	SE_B	β	t	p
CDU/CSU														
M	−.21	$-.27_{A}$	$-.24_{AB}$	$-.16_{B}$	$-.19_{AB}$	$-.16_{B}$	4.10	.003	.010	0.03	0.01	.09	3.47	.001
SD	.44	.45	.45	.43	.45	.41								
Cohen's *d*	−.67	−.83	−.75	−.52	−.59	−.57								
α	.75	.76	.73	.74	.77	.73								
FDP														
M	−.15	$-.20_{AB}$	$-.23_{AC}$	$-.10_{D}$	$-.11_{D}$	$-.14_{BD}$	7.22	$9^{*}10^{-6}$	.018	0.03	0.01	.10	3.83	.0001
SD	.37	.37	.36	.37	.37	.36								
Cohen's *d*	−.57	−.77	−.89	−.38	−.42	−.54								
α	.66	.64	.59	.69	.68	.67								
SPD														
M	−.03	$-.12_{A}$	$-.08_{AE}$	$.04_{CD}$	$.02_{CD}$	$.00_{CDE}$	10.67	$2^{*}10^{-8}$	.027	0.04	0.01	.12	4.86	$1^{*}10^{-6}$
SD	.40	.46	.38	.37	.38	.35								
Cohen's *d*	−.09	−.36	−.29	.16	.09	−.01								
α	.70	.76	.65	.70	.68	.65								
GREEN														
M	−.08	$-.10_{A}$	$-.11_{A}$	$-.03_{A}$	$-.07_{A}$	$-.08_{A}$	1.84	.119	.005	0.01	0.01	.03	1.28	.201
SD	.38	.40	.40	.35	.35	.40								
Cohen's *d*	−.30	−.34	−.39	−.14	−.29	−.27								
α	.68	.70	.67	.66	.64	.74								
PDS														
M	−.23	$-.29_{AB}$	$-.29_{A}$	$-.18_{C}$	$-.23_{ABC}$	$-.20_{BC}$	4.32	.002	.011	0.02	0.01	.08	3.10	.002
SD	.38	.39	.39	.36	.39	.37								
Cohen's *d*	−.87	−1.03	−1.06	−.72	−.85	−.78								
α	.67	.66	.66	.67	.70	.67								

Note: $N = 1568$. Means within rows not sharing a common subscript differ significantly at $p < .05$ (two-sided) according to Tukey's HSD-Test.

phase) and .72 (2nd phase). Taking the overall number of trials into account, together with an interindividually randomized order of five successive ST-IATs and along with the low number of 35 trials per block, the reliability was satisfactory. We also inspected if changes of ST-IAT effects were attributable to a drop in reliability across positions, but we found no equivalent trend for reliability (cf. Table 3).

We calculated test-retest reliability as an indicator of stability (mean time interval = 17.77 days, $SD = 15.68$). Analyses confirmed that whereas the explicitly measured attitudes were stable ($r = .76$–$.89$; cf. italics in lower left part of Table 2), the ST-IATs exhibited only low to medium stability ($r = .21$–$.46$). Presumably, the high number of ST-IATs and the low number of trials per ST-IAT resulted in less than optimal estimates of internal consistencies, which consequently led to lower stability indices. Accordingly, correction for attenuation led to higher test-retest reliabilities (cf. Table 2).[3] Although unsatisfactory, retest reliabilities within this range are not uncommon for original IATs as well, even with immediate retests (Banse, Seise, & Zerbes, 2001, Exp. 2; Cunningham, Preacher, & Banaji, 2001).

Given the seemingly low stabilities, we also tested the stability of the rank orders of the ST-IATs of each participant and calculated the concordance of the rankings according to Kendall's *W* for each participant. Averaging these values across participants resulted in a non-trivial mean coefficient of concordance of .65 for the ST-IAT rank orders (median = .75; mode = .80). Thus, the outcomes of both measurement occasions were not arbitrary and the rank order of ST-IATs remained quite stable within participants.

[3]We thank an anonymous reviewer for pointing this out.

DOI: 10.1002/ejsp

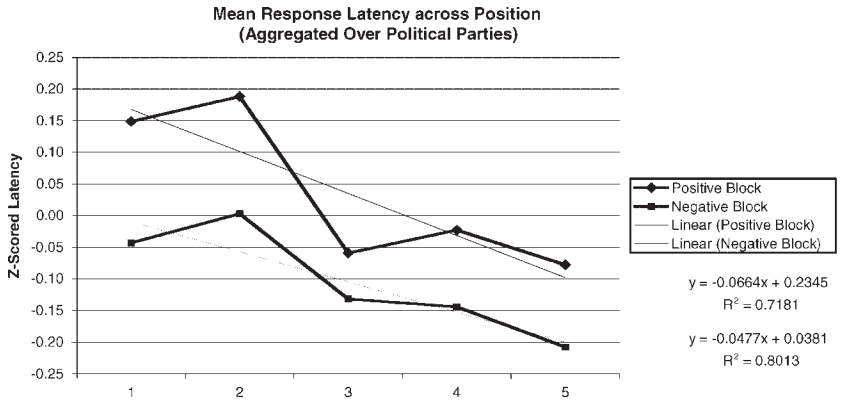

Figure 2. Decline of mean response latencies in positive and negative blocks across the course of five ST-IATs (Study 1)

Convergent and Discriminant Validity

For each party, the highest correlation emerged between the ST-IAT and the explicit measure of the respective party, as expected. Apparently, a specific ST-IAT mostly captured the attitude towards the specific party, $r = .34$–$.49$ (mean $r = .39$, uncorrected for attenuation). This pattern was confirmed in the retest sub-sample (mean $r = .50$). Corrected for measurement error, convergent validities reached mean *r*s of .47 and .58 in the 1st and 2nd phase, respectively. Simultaneously, a specific ST-IAT correlated with attitudes towards other parties to a significantly lesser extent (all $ts(1565) > 2.12$, $ps < .05$). Convergent validities (corrected for attenuation) between ST-IATs and explicit measures are depicted in Table 2 and printed in boldface.

The left-wing/right-wing distinction was clearly as evident in ST-IATs as in explicit measures. Looking at the first phase in Table 2, the ST-IAT reliably assessed the positive and negative intercorrelations among parties from the left or the right spectrum, regardless of whether explicit or implicit measures served as criteria, yielding evidence for discriminant validity of the ST-IAT. The correlations of the small retest sample in the second phase perfectly mirrored the pattern of the first phase, though small correlations remained non-significant more often due to a lack of power. Thus, despite the negative evaluation of all parties across all participants on average (see Figure 1), the ST-IATs did not simply measure a general negative evaluation of political parties or party members, irrespective of the specific target party. Rather, they captured the interindividually different evaluations of five interrelated target objects.

One shortcoming in terms of discriminant validity seems to be the correlations of similar magnitude between ST-IATs and explicitly measured attitudes of same-wing parties. For instance, the FDP ST-IAT correlated .34 with the explicit FDP liking in the first phase, but also .31 with the explicit CDU/CSU liking (uncorrected for attenuation). Also, the SPD ST-IAT captured both SPD and GREEN liking to a similar extent (.36 vs. .31). We presume that the left–right value orientation, which is known to underlie the explicitly measured attitudes, is partly responsible for the suboptimal discriminant validity of same-wing ST-IATs. We will pursue this topic in Study 2.

Position Effects

An important question is whether the attenuation of mean ST-IAT effects across positions (cf. Table 3) affected convergent validity. To this end, we conducted multiple regression analyses of each explicit measure on the respective implicit measure, the position of the ST-IAT and their interaction term (Baron & Kenny, 1986; all predictors centred on means; see Table 4). The only significant moderation occurred for the FDP, yet the positive regression weight suggests that the later the ST-IAT was applied, the more it was related to the explicitly measured attitude towards the FDP. Thus, in contrast to the decline of mean ST-IAT effects across positions, there was no linear decline of the implicit–explicit relationship.

We also checked whether position moderated the predictive value of the ST-IATs for voting behaviour. As some participants either did not supply voting intention or did not actually vote in the election, the following regression analyses

DOI: 10.1002/ejsp

Table 4. Correlations of explicit and implicit measures of political attitudes across measurement position (Study 1)

						Multiple regression model of explicit attitude including position effects														
	Implicit–explicit correlation by position					ST-IAT					Position of ST-IAT					ST-IAT × position				
Political party	1	2	3	4	5	*B*	SE_B	β	*t*	*p*	*B*	SE_B	β	*t*	*p*	*B*	SE_B	β	*t*	*p*
CDU/CSU	.46	.57	.47	.46	.45	2.27	0.11	.49	20.98	$2^{*}10^{-86}$	0.03	0.03	.02	0.96	.339	0.042	0.076	.01	0.56	.578
FDP	.30	.25	.33	.36	.48	1.86	0.13	.34	14.48	$1^{*}10^{-44}$	−0.07	0.03	−.05	−2.01	.045	0.240	0.093	.06	2.58	.010
SPD	.44	.32	.36	.43	.19	1.49	0.10	.35	14.36	$5^{*}10^{-44}$	−0.07	0.03	−.05	−2.26	.024	−0.104	0.072	−.04	−1.45	.147
GREEN	.43	.46	.38	.45	.40	2.30	0.13	.42	18.38	$2^{*}10^{-68}$	−0.05	0.03	−.03	−1.35	.176	−0.022	0.088	−.01	−0.25	.802
PDS	.35	.37	.30	.44	.33	1.78	0.12	.36	14.42	$2^{*}10^{-44}$	−0.02	0.03	−.02	−0.75	.455	0.032	0.083	.01	0.39	.700

Note: $N = 1568$.

rests on 1386 data sets. In binary logistic regression analyses, we regressed dummy-coded voting behaviour on the ST-IATs, serial position and the respective interaction term (see Table 5). Positive regression coefficients confirmed that higher ST-IAT scores raised the likelihood of voting for a specific party. The interaction terms in the multiple regression analyses showed that the variability of the predictive value of the ST-IAT did not linearly depend on position. The only significant interaction occurred for the FDP. Similar to the aforementioned increase of the FDP implicit–explicit correlation across position, the positive regression weight reflects an increase of predictive value of the FDP ST-IAT. Taken together the data disconfirm the hypothesis that the ST-IAT's usefulness declines with position in a series of multiple ST-IATs. (For analyses regarding the utility of ST-IATs for predicting voting behaviour beyond explicitly measured attitudes, the reader is referred to Friese, Bluemke, & Wänke, 2007.)

Summary and Discussion

We found notable evidence that the ST-IAT can be used to indirectly evaluate five interrelated target objects. Although our findings are limited to the political domain and our ST-IATs had slightly weaker reliabilities than typical IATs, we

Table 5. Criterion correlations of ST-IATs across measurement position (Study 1)

	Prediction of voting behaviour (binary logistic regression)						Binary logistic regression model of voting behaviour including position effects											
		By position					ST-IAT				Position of ST-IAT				ST-IAT × position			
Political party	Overall	1	2	3	4	5	*B*	SE_B	Wald	*p*	*B*	SE_B	Wald	*p*	*B*	SE_B	Wald	*p*
CDU/CSU																		
B	2.80	3.07	2.47	3.55	1.86	3.60	2.79	0.21	171.50	$3^{*}10^{-39}$	0.02	0.06	0.14	.706	−0.02	0.15	0.02	.893
Nagelkerke-R^2	.27	.32	.25	.33	.15	.32												
FDP																		
B	2.34	1.58	1.57	2.06	3.29	3.36	2.44	0.25	97.08	$7^{*}10^{-23}$	−0.19	0.07	7.32	.007	0.53	0.17	9.10	.003
Nagelkerke-R^2	.14	.08	.06	.11	.22	.24												
SPD																		
B	1.52	1.68	1.52	1.45	1.88	1.10	1.49	0.17	79.30	$5^{*}10^{-19}$	0.01	0.05	0.02	.900	−0.10	0.12	0.67	.414
Nagelkerke-R^2	.09	.14	.08	.08	.12	.05												
GREEN																		
B	1.83	1.87	1.80	1.99	1.81	1.93	1.84	0.18	104.75	$1^{*}10^{-24}$	0.03	0.05	0.41	.520	0.03	0.13	0.04	.842
Nagelkerke-R^2	.12	.13	.13	.10	.10	.14												
PDS																		
B	3.27	5.21	2.50	2.10	4.60	3.14	3.31	0.37	80.25	$3^{*}10^{-19}$	−0.09	0.12	0.63	.428	−0.03	0.25	0.02	.889
Nagelkerke-R^2	.20	.41	.13	.09	.34	.18												

DOI: 10.1002/ejsp

currently conclude that the ST-IAT displays sufficient reliability and stability to assess complex attitudinal patterns. In addition, we found support for the validity of the ST-IAT as apparent in explicit and behavioural criterion correlations and in the absence of linear position effects on reliability and validity. Nonetheless, repeated applications do weaken ST-IAT effects at least in absolute terms and put a question mark on the interpretation of absolute effect sizes. The implication of these findings is that changing means need not be detrimental to validity. Nevertheless, we hasten to add that the validity of a measure could still be seriously affected even if unaltered means indicated the absence of any group differences (cf. Perugini & Banse, 2007).

Having answered the question of serial position effects and, at least partly, the question of construct validity, more data on the validity of the ST-IAT seemed warranted. As a next step, we sought to extend convergent validity with respect to an implicit measure of a related yet distinct construct in order to substantiate our assumptions regarding a left–right value continuum underlying the evaluation of the parties.

STUDY 2: EXTENDING CONSTRUCT VALIDITY

Overview and Hypotheses

So far, we have examined construct validity of the political ST-IATs mainly in terms of convergence with explicit measures and other ST-IATs. Yet, another important aspect of construct validity is convergence with other latency-based measures assessing a related construct. This is a crucial topic because of the low intercorrelations often found among different implicit measures even when targeted at identical attitude objects (Brauer, Wasel, & Niedenthal, 2000; Cunningham et al., 2001). We sought supplementary evidence for the political left-wing/right-wing continuum presumably underlying the party appraisals and hypothesized that this ideological aspect can be assessed parsimoniously by applying a left–right IAT. Validating that automatic party preferences are related to the spontaneous evaluation of left- and right-wing concepts would strengthen the view that ST-IATs assess what the parties stand for in political terms, rather than only the liking of the respective politicians who were selected as stimuli (Bluemke & Friese, 2006). Our primary goal therefore was to extend the correlation matrix by the political left-wing/right-wing dimension, measured both explicitly and implicitly. We expected the explicit ideological self-identification to be significantly related to explicit and implicit measures of attitudes towards specific political parties. In the same manner, a left-right IAT should be related to ST-IAT effects of left- and right-wing parties. If we found such evidence for convergent validity, this could partly explain why discriminant validity of same-wing ST-IATs was limited in Study 1. We used the most prominent parties from the political left, the SPD, and from the political right, the CDU/CSU, for this examination.[4]

METHOD

Participants

One hundred and twenty-nine participants ($M_{age} = 25.69$ years, $SD = 9.09$) either took part in the study in partial fulfilment of course requirements ($N = 63$) or found their way to the experiment via hyperlinks at the Web lab at the University of Heidelberg ($N = 66$). We offered them deeper insight on their 'automatic preferences for leading politicians' on the starting page, but no monetary incentives.

[4]One might argue that a CDU/CSU–SPD IAT would be a more appropriate comparison standard to the CDU/CSU and SPD ST-IATs. Whereas the relation of single-target measures to an IAT evaluating identical target objects has already been explored (Karpinski & Steinman, 2006; Wigboldus et al., 2004), these interrelations could have resulted from overlapping target stimuli. By contrast, we investigated automatically activated ideological aspects underlying the evaluation of political parties and thus explored the convergent validity with another latency-based measure that rested on completely different target stimuli, that is, an associated, yet distinct construct. We did not explore if ST-IAT and IAT converge in general, but rather focused on the question whether an automatic evaluation of left-/right-wing concepts exists that underlies the evaluation of political parties. This would strengthen the view of the ST-IAT having content validity and limited discriminant validity for political parties of the same ideological wing.

DOI: 10.1002/ejsp

Although the Internet sample was perfectly balanced, overall females predominated (67%) due to a skewed lab sample. Internet participants differed somewhat from the lab sample in terms of age and varied more in terms of education and job level. Voters' party preferences during the election in 2002 differed somewhat from Study 1 (GREEN: 48.1%, SPD: 13.2%, CDU/CSU: 7.8%, FDP: 7.0%, PDS: 3.9%). Excluding data from participants with more than 20% errors in one of the ST-IAT blocks resulted in 107 remaining data arrays.

Procedure and Materials

The web pages were similar to Study 1. After explicitly measuring attitudes towards the parties, we first applied the CDU/CSU and SPD ST-IATs in counterbalanced order, then the left–right IAT. We did so to prevent carry-over effects of ideological associations to the evaluation of the parties, which might arbitrarily inflate any intercorrelations. After the examination we asked participants to locate themselves in terms of political ideology on a 9-point rating scale ('Do you consider yourself to be rather politically left or politically right?') with scale endings labelled 'left' and 'right' but without defining what it meant to be 'left' or 'right'.

ST-IATs

Each party concept was represented by five verbal stimuli (e.g. the name *SCHRÖDER*) plus five pictorial stimuli (e.g. a picture of former Chancellor Gerhard Schröder) matched in identity. Some exemplars were changed across the studies (cf. Appendix). The attribute dimension was represented by verbal stimuli exclusively. Target and attribute trials amounted to 35 trials in each combined block. We kept task switching constant across blocks and ST-IATs, and each participant received the same stimulus order.

IAT

The target concepts *left*/*right* and the attribute concepts *negative*/*positive* served as category labels. We randomly recruited 16 university students (about one half with psychology majors; $M_{age} = 22.14$ years, $SD = 2.69$) who were unaware of the specific hypotheses for a pretest of 55 potentially useful nouns. After inspection of the data, we chose seven unambiguously identifiable stimuli for the politically left and right categories, respectively, $M_{left} = 3.74$ ($SD = .24$) versus $M_{right} = 7.22$ ($SD = .61$) on 9-point rating scales (cf. Appendix), which were equivalent in valence on average, $M_{left} = 5.02$ ($SD = .94$) versus $M_{right} = 4.96$ ($SD = .67$) on 9-point rating scales. The stimulus sets were thus free of evaluative confounds. We held positive and negative attribute stimuli constant to the ST-IATs. Each combined block comprised 56 trials with target and attribute stimuli in alternating order. The IAT comprised 196 trials in total, including 28 training trials for target and attribute stimuli each prior to the first combined block and 28 training trials for the reversed target discrimination phase. All participants encountered the *left* + *positive* (*right* + *negative*) block first and always received the same fixed order of stimuli. However, we counterbalanced whether participants started with *left* + *positive* on the left response key or on the right response key, thus controlling for the Simon-like incongruence of the political concept and the geometric side of the reaction required (which did not bias ST-IAT effects).

Data preparation for the (ST-)IATs followed the procedure as described in Study 1.

RESULTS AND DISCUSSION

Sample Differences

A multivariate test of explicit and implicit measures revealed that there were differences between the Internet and lab participants in only a few cases (overall F-test: $F(9, 97) = 2.55$, $p = .01$). Participants in the lab sub-sample explicitly liked the GREEN party better than Internet participants, Ms = 6.16 versus 4.96, $F(1, 105) = 11.98$, $p = .0008$, $\eta^2 = .10$. They were also more left-wing oriented both explicitly, Ms = 3.24 versus 4.05, $F(1, 105) = 6.05$, $p = .02$, $\eta^2 = .06$ and implicitly,

DOI: 10.1002/ejsp

Ms = 0.44 versus 0.17, $F(1, 105) = 4.78$, $p = .03$, $\eta^2 = .04$. All other F-tests remained non-significant (Fs < 2.35, ps > .13, $\eta^2 \leq .02$).

Reliability

Cronbach's α amounted to .80, .80 and .88 for the CDU/CSU and SPD ST-IATs, and the left–right IAT, respectively. Unexpectedly, the reliability of the SPD ST-IAT declined slightly from the first ($\alpha = .85$) to the second position ($\alpha = .65$), although the overall reliabilities of the ST-IATs were slightly higher than in the previous study.

Construct Validity

We start with a look at the consistency of the correlations with the explicit measure. As evident in Table 6, participants' party likings were dependent on their personal stance on the left–right continuum. Moreover, the left–right bipolarity clearly underlay the explicit evaluation of the big parties, as a difference score between the explicit liking of SPD and CDU/CSU correlated highly with the ideological self-identification ($r = .71$, $p < .001$). We next tested whether this result would also hold for the implicit measures.

As the intercorrelations of the implicit measures show in Table 6, left–right IAT effects can be used to predict the automatic evaluation of the parties (ST-IAT effects): The more the category *left* is spontaneously preferred to the category *right*, the higher the ST-IAT score for the SPD and, correspondingly, the lower the ST-IAT score for the right-wing party, the CDU/CSU. Interestingly this held even though the SPD ST-IAT and the CDU/CSU ST-IAT did not show a negative correlation as they did in the first study. Nonetheless, the difference score between the two ST-IATs clearly revealed a positive correlation with the left–right IAT ($r = .40$, $p < .001$) as well as with the explicit left–right self-identification ($r = .38$, $p < .001$). These findings show that the concepts assessed by the ST-IATs partly reflected general political values. This is far from being trivial. Regarding the debate of stimulus versus label influence in IATs (Bluemke & Friese, 2006; Nosek et al., 2006), an ST-IAT effect might simply reflect the liking of the specific politicians. Although the evaluation of a party is likely to depend on its representatives, the ST-IAT also assesses what the parties stand for in ideological terms, yielding support for influence of the category labels.

Regarding the correlations between explicit and implicit measures of attitudes (mixed-method intercorrelations), we replicated the findings from Study 1. The ST-IATs correlated positively with explicit appraisals of parties of the same ideological pole, but negatively with opposite ones (cf. Table 6). In addition, the left–right IAT and the explicit ideological self-identification correlated positively. Also, the CDU/CSU and SPD ST-IATs reflected the explicit ideological self-identification as hypothesized. Overall the ST-IATs displayed the same, though attenuated, correlative pattern as in Study 1, with a rather low convergent validity of the SPD ST-IAT with explicit SPD liking.[5]

Taken together, in Study 2 we consolidated convergent and discriminant validity for the CDU/CSU ST-IAT, and to a lesser degree for the SPD ST-IAT. However, the results also show that discriminant validity in terms of correlations with explicit party likings of the *same* political wing was less than ideal. Yet, in the light of an underlying left–right concept, which seems to partly define the political camps also at an implicit level, convergence of ST-IATs *on the same side* of the spectrum may be driven by ideological overlap, but at the same time discrimination among parties *within this scope* may be limited. Both explicit and implicit criteria established a link between the ST-IATs and the ideological position of the parties.

GENERAL DISCUSSION

We intended to shed some light on open questions relating to the psychometric status of a relatively new measure: the ST-IAT. In two studies, we analysed psychometric properties as well as susceptibility to position effects.

[5]This might be due to the low variation of our sample in terms of political preferences and range restriction of the SPD liking, which had the smallest standard deviation of all explicit measures (SDs = 1.46 vs. 1.70, 1.70, 1.86, 1.90). The finding that the correlation was evident in the more heterogeneous Internet sub-sample, $r_{N=56} = .38$ ($p = .005$), but absent in the lab, $r_{N=51} = -.04$ ($p = .76$), is supportive of this assumption. Overall, the correlation patterns of the ST-IATs are quite similar for both sub-samples (cf. Table 6).

DOI: 10.1002/ejsp

Table 6. Intercorrelations of explicit and implicit measures of political attitudes and ideological self-identification (Study 2), corrected for measurement error of implicit measures

	Explicit measures						Implicit measures		
	CDU	FDP	SPD	GREEN	PDS	SI	CDU	SPD	IAT
Total sample ($N = 107$)									
Explicit measures									
CDU/CSU	1.00								
FDP	.67***	1.00							
SPD	−.09	.02	1.00						
GREEN	−.45***	−.27**	.43***	1.00					
PDS	−.25**	−.16	.04	.25*	1.00				
Self-identification (left-right)	−.68***	−.50***	.35***	.61***	.35***	1.00			
Implicit measures									
CDU/CSU ST-IAT	**.37*****	.35***	−.02	−.09	−.09	**−.22***	1.00		
SPD ST-IAT	−.20*	−.06	**.18**+	.10	.18+	**.33*****	.22*	1.00	
Left-right IAT	−.45***	−.30**	.24*	.53***	.31**	.60***	**−.24***	**.36*****	1.00
Lab sample ($N = 51$)									
Explicit measures									
CDU/CSU	1.00								
FDP	.72***	1.00							
SPD	−.15	−.08	1.00						
GREEN	−.40**	−.33*	.50***	1.00					
PDS	−.41**	−.41**	−.14	.16	1.00				
Self-identification (left-right)	−.75***	−.58***	.28*	.50***	.54***	1.00			
Implicit measures									
CDU/CSU ST-IAT	**.41****	.33*	−.14	−.29*	−.11	**−.45*****	1.00		
SPD ST-IAT	−.34*	−.24+	**−.04**	.06	.21	**.37*****	.16	1.00	
Left-right IAT	−.44**	−.22	.06	.33*	.18	.51***	**−.44****	**.36****	1.00
Internet sample ($N = 56$)									
Explicit measures									
CDU/CSU	1.00								
FDP	.65***	1.00							
SPD	−.06	.01	1.00						
GREEN	−.46***	−.27*	.42**	1.00					
PDS	−.07	−.13	.22	.26	1.00				
Self-identification (left-right)	−.63***	−.49***	.40**	.63***	.15	1.00			
Implicit measures									
CDU/CSU ST-IAT	**.38****	.37**	.05	−.07	−.14	**−.16**	1.00		
SPD ST-IAT	−.07	.19	**.38****	.11	.12	**.29***	.26+	1.00	
Left-right IAT	−.44***	−.40**	.38**	.61***	.41**	.62***	**−.17**	**.35****	1.00

Note: SI=Left-right self-identification (higher values indicate left wing orientation).; Convergent validities of ST-IATs are shown in boldface.
$^{+}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Psychometric Properties

Across 12 ST-IATs, internal consistencies regularly reached .70 or higher (mean $\alpha = .72$). It appears that the response window that Karpinski and Steinman (2006) applied in the SC-IAT is not a necessary precondition for acceptable reliability as long as the ST-IAT effects are calculated on the basis of target *and* attribute stimuli. In the present ST-IATs reliabilities were at the lower bound of what can be expected for traditional IATs (.70–.90, Nosek et al., 2006). Note that according to a recent meta-analysis (Hofmann et al., 2005) internal consistencies in IATs linearly depend on the amount of trials in the critical blocks ($r_{\alpha \sim \text{trials}} = .54$) and that we used a rather low number of 35 trials per block only. Higher reliabilities may be observed when more trials are used. Although reliability fluctuated across positions to some extent, there was no consistent linear decline across positions such that lower reliability emerged for ST-IATs that were assessed at a later stage of the sequence of five ST-IATs. We are not aware of any other studies examining position effects on the internal consistency of (ST-)IATs.

DOI: 10.1002/ejsp

The first report on test-retest reliabilities of ST-IATs hints at a range between .21 and .46 for five ST-IATs across different experimental sessions. Admittedly, these are fairly low numbers compared to the IAT ($r_{tt} \sim .50$; Nosek et al., 2005). Nevertheless, without measurement error, the stability of the ST-IAT effects would peak at .60. Even without correction for attenuation, rank ordering of participants' preferences in these ST-IATs demonstrated satisfactory stability (mean Kendall's $W = .65$). We are confident that higher stability for ST-IATs will be found in studies with more trials in combined blocks, less target objects assessed sequentially and with a shorter and less fluctuating retest interval.

Regarding convergent validity we obtained promising results. Correlations of ST-IAT scores with explicit measures of the same attitude constructs amounted to a mean $r = .34$ (uncorrected for attenuation) and $r = .43$ (corrected for attenuation) across the two studies.[6] This value exceeds the disattentuated mean population correlation of $\rho = .24$ between IATs and explicit measures of attitudes according to a recent meta-analysis in which even the research domain leading to the highest population correlation, consumer attitudes, only reached a disattenuated value of $\rho = .34$ (Hofmann et al., 2005). As a word of caution, these results could represent the upper bound of the ST-IAT's capabilities if we take the typically strong implicit–explicit correspondence in the political domain into account. Self-selected Internet samples might add to this correspondence if participants who held stronger political attitudes preponderated, because attitude importance is known to moderate the implicit–explicit relationship (Friese et al., 2007; Karpinski et al., 2005; Nosek et al., 2005).

Furthermore, the correlations between related implicit measures showed convergent validity, as well. In Study 2, two ST-IATs correlated meaningfully with an IAT that measured participants' automatic associations to the political left-wing/right-wing continuum. These results demonstrate that the ST-IATs captured not only evaluations of a specific party and its representatives but that these evaluations related in a more general sense to a comprehensible ideological pattern displayed even at a more automatic level. Given this ideological overlap of same-wing parties at the implicit level, discriminant validity—as evident in correlations with non-targeted implicit and explicit measures—can be said to be acceptable.

Serial Position Effects

The analysis of position effects showed that researchers need not fear systematic fatigue effects that would diminish reliability or validity across a lengthy ST-IAT measurement procedure. Correlations of the ST-IAT with several criteria were by and large unaffected by detrimental position effects. However, several single-target assessments in a series may be subject to exercise effects as evident in speed gains that can affect the magnitude of ST-IAT effects. Researchers are thus encouraged to be careful when comparing absolute scores, especially when relying on a fixed, rather than a counterbalanced order of the measures (cf. Karpinski & Steinman, 2006; Steffens & Schulze-König, 2006).

Even if fatigue and strategic responding need not pose a problem, the order of implicit measures can nevertheless create specific contexts, alter the activated associations and bias the measurement (cf. Bless & Schwarz, 1998). In Study 1, we tried to reduce the influence of such carry-over effects by randomly determining the ST-IAT order for each participant. Temporarily activating all the target parties before the implicit measures were taken should also have minimized carry-over effects across the following ST-IAT sequence. By contrast, not priming all the target objects beforehand would have maximized the dependency of each ST-IAT measurement on the specifically created context (Sudman, Bradburn, & Schwarz, 1996). Although we cannot preclude that an individual's results might still have been affected by the specific order he or she encountered, at the aggregate level serial position effects should not have been the consequence of prevailing context effects. Given that carry-over effects would have affected results interindividually and biased ST-IAT validity downward, the resulting correlation patterns were quite impressive. Caution is nevertheless in place since the chances that the serial positions will elicit detrimental context effects rise as the number of ST-IATs increases, as the sequence of ST-IATs gets fixed and as a full-scale context of the target objects is not made available beforehand.

[6]We computed the mean convergent implicit–explicit correlation separately for each study, drawing on 12 correlations (Study 1 split into two phases, Study 2), then the mean of the study correlations weighted by sample size (Schmidt, Hunter, & Raju, 1988), but limited the largest weight ($N_{\text{Study 1, First Phase}} = 1{,}568$) to the highest weight of the remaining samples ($N = 107$; cf. Hofmann et al., 2005).

DOI: 10.1002/ejsp

Limitations

Despite the encouraging support for reliability and validity of the ST-IAT, we would like to discuss some limitations of the present data. First, both studies were concerned with political attitudes in Germany. Conceptual replications are needed in other research domains with different contents and different ST-IAT specifications such as length of block and stimulus selection. In the present case, we do not know if the reliable assessment of evaluative associations was partly due to the global left–right dimension underlying the party evaluations. Also, a different content domain with lower implicit–explicit correlations could be better suited to demonstrate the value of ST-IATs for incrementally explaining behavioural variance, for instance when participants' control resources are limited or when positive mood fosters the reliance on associative information processing (cf. Friese et al., in press; Hermsen, Holland, & van Knippenberg, 2006; Hofmann & Friese, in press; Hofmann et al., 2007; Strack & Deutsch, 2004).

Second, in our ST-IATs the parties were represented by names and pictures of the most prominent party members. However, we did not collect explicit evaluations of these exemplars but confined ourselves to global party evaluations. Given the influence that stimuli can exert in traditional IATs (Bluemke & Friese, 2006), it might be wise to incorporate explicit ratings of the stimuli, as this may contribute to a more differentiated pattern of criterion correlations and thus add to the picture of discriminant validity.

Finally, given that an ideological IAT can also predict explicit party evaluations and given the high convergence of the ST-IATs with a left–right IAT, one could raise the question if we really need five evaluative ST-IATs. First of all, we think the same question could be asked with regard to the theoretical usefulness of explicit party evaluations. Though the preference of left-wing to right-wing concepts already conveys a lot of information on a participant's political preferences, researchers continue to assess explicit attitudes specifically because these measures carry additional evaluative aspects that are different from ideological orientation. Second, we must not conclude to lower validity of the ST-IAT only because we found higher relations towards explicit likings with the IAT rather than with the ST-IAT. Logically, these patterns could also reveal a stronger influence of the left–right dimension in *explicit* ratings rather than in the ST-IAT. Last but not least, our own empirical evidence shows that the evaluative ST-IATs did not capture the *ideological* aspect specifically, but that they rather assessed automatic *evaluations* that hinge on the liking of party representatives and appreciation of enacted politics (rather than programme statements). An analysis of partial correlations in Study 2 shows that the CDU/CSU ST-IAT predicted the explicit attitude even when controlling for the left–right IAT ($\Delta R^2 = .063$, $r_p = .28$, $p = .004$). The SPD ST-IAT explained additional variance at least for Internet participants ($\Delta R^2 = .063$, $r_p = .26$, $p = .048$). In sum, the global left–right distinction underlying the evaluation of political parties poses one ingredient of convergent validity, but it does not preclude the assessment of party specific evaluative associations. The value of evaluative ST-IATs, however, will increase whenever no such global dimension underlies the appraisal of multiple target objects.

Comparison to Other Single-Category Measures

Extending our validation by applying another indirect measure as in Study 2 is one step to support the usefulness of the ST-IAT, but even implicit–implicit correlations between the ST-IATs and a left–right IAT could be inflated due to the high degree of similarity of the measurement types and shared method variance. Future studies need to compare the usefulness of ST-IATs to alternative indirect measures of a different kind. It is particularly important to know whether the ST-IAT is doing a better job at predicting behaviour than other tools.

If we are to compare the ST-IAT to the SC-IAT in psychometric terms, we never observed a reliability drop as reported by Karpinski and Steinman (2006) who obtained reliabilities ranging from $r = .55$–$.85$ despite the fact that they applied almost twice as many critical trials as we did. Whether this difference is owed to the response-window technique of the SC-IAT or to the particular content domain remains speculative. Interestingly, Cunningham et al. (2001) also observed a decline of internal consistency from .78 to .63 and estimates of stability from .46 to .36 when they compared a typical IAT to a response-window IAT while keeping the target concepts constant. Purposely forcing participants to accelerate during responding may alter the nature of an (ST-)IAT so that the typically good psychometric properties of (ST-)IATs fall off.

In comparison to the original IAT, both ST-IAT and SC-IAT data show that single-category assessments can extract information beyond that gained by comparative IATs, such as a *Pepsi–Coke* IAT or a *Male–Female* IAT. There is accumulating evidence that these measures can, for instance, predict the amount of snacks or beverages consumed in a

DOI: 10.1002/ejsp

laboratory taste test (Friese et al., in press; Hofmann & Friese, in press; Richetin et al., 2007) or determine the evaluation of phobic stimuli (Huijding & de Jong, 2006). The value of a single-target measure is particularly evident when arbitrary counter-categories—such as the category 'other' in self-esteem and self-concept IATs—lessen criterion correlations (Karpinski & Steinman, 2006). Yet to our knowledge a direct comparison of the ST-IAT's and the IAT's predictive validity of future behaviour is still missing.

Of the remaining single-category measures, the EAST (De Houwer, 2003) resembles the original IAT most. It requires participants to evaluate the colour of target stimuli instead of their valence or semantic content. Despite some reports on the usefulness of the EAST to predict interindividual differences in behaviour, there have been doubts concerning the usefulness of the EAST in terms of reliability and validity (De Houwer & De Bruycker, 2007a, 2007b).

CONCLUSIONS

In sum, we think our results highlight two potential assets of the ST-IAT compared to the traditional IAT: its ability to capture the evaluation of a single target category and to reliably and efficiently do so. The ST-IAT aims at assessing automatic affective tendencies and removes a counter-category that would introduce nuisance variance in many traditional IAT applications. It seems even more beneficial when several attitude objects are to be evaluated simultaneously and naturally opposing categories are unavailable. The choice between the comparative and non-comparative measure depends on the research question. Applying an IAT might be an efficient way to analyse relative preferences towards two target categories. Whenever one is interested in non-relative evaluations, however, an ST-IAT may be a better choice. One might conceive of the ST-IAT as an absolute measure. However, neither does the outcome index yield an absolute scale value (Blanton & Jaccard, 2006), nor does an implicit association procedure circumvent the fundamentally relative nature of human judgement (Gawronski & Bodenhausen, 2005). The ST-IAT outcome is essentially relative in the sense that it is calculated as a difference between positive and negative associations towards the same target, yielding an effect size estimate (Fiedler et al., 2006). If anything, the ST-IAT can be said to discard a contrast category so that one can hope to get closer to a non-relative evaluation. It is doubtful whether any psychological measure will ever yield absolute evaluations. However, whenever one is interested in evaluations of multiple target objects, the ST-IAT could present an efficient way to get there. With its potential to detect differences that exist among several interrelated target objects—by and large unaffected by serial position effects—the ST-IAT promises to be a valuable research tool that expands the range of implicit measures of attitudes, stereotypes, self-concept and other domains.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of Klaus Fiedler. The research presented in this paper was supported in part by a Leibniz Award to Klaus Fiedler from the Deutsche Forschungsgemeinschaft. We thank Georg Johann for the development of the JAVA applet used for the single-target assessments as well as Klaus Fiedler, Henning Plessner, Sabine Sczesny, Anita Todd, Michaela Wänke and the anonymous reviewers for valuable comments on an earlier draft of this manuscript. We especially thank Stefanie Höhl, Daniel Konermann and Martin Stegmüller for their help in collecting the data of Study 2.

REFERENCES

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, *48*, 145–160.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, *61*, 27–41.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*, 192–212.

DOI: 10.1002/ejsp

Bless, H., & Schwarz, N. (1998). Context effects in political judgement: Assimilation and contrast as a function of categorization processes. *European Journal of Social Psychology*, *28*, 159–172.
Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology*, *42*, 163–176.
Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, *4*, 79–101.
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*, 760–773.
Conover, P. J., & Feldman, S. (2004). The origins and meaning of liberal/conservative self-identifications. In J. T. Jost, & J. Sidanius (Eds.), *Political Psychology* (pp. 200–216). New York: Psychology Press.
Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *121*, 163–170.
Czopp, A. M., Monteith, M. J., Zimmerman, R. S., & Lynam, D. R. (2004). Implicit attitudes as potential protection from risky sex: Predicting condom use with the IAT. *Basic and Applied Social Psychology*, *26*, 227–236.
De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.
De Houwer, J., & De Bruycker, E. (2007a). The identification-EAST as a valid measure of implicit attitudes toward alcohol-related stimuli. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 133–143.
De Houwer, J., & De Bruycker, E. (2007b). The Implicit Association Test outperforms the Extrinsic Affective Simon Task as a measure of interindividual differences in attitudes. *British Journal of Social Psychology*, *46*, 401–421.
De Jong, P. J., van den Hout, M. A., Rietbroek, H., & Huijding, J. (2003). Dissociations between implicit and explicit attitudes towards phobic stimuli. *Cognition and Emotion*, *17*, 521–545.
Devos, T., & Banaji, M. R. (2005). American = White? *Journal of Personality and Social Psychology*, *88*, 447–466.
Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.
Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the ''I'', the ''A'' and the ''T'': Logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*, 74–147.
Friese, M., Bluemke, M., & Wänke, M. (2007). Voting behavior and gut reaction: The 2002 German parliamentary elections. *Experimental Psychology*, *54*, 247–255.
Friese, M., Hofmann, W., & Wänke, M. (in press). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology.*
Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology*, *49*, 171–180.
Gawronski, B., & Bodenhausen, G. V. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation and retrieval experiences. *Journal of Personality and Social Psychology*, *89*, 672–685.
Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 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*, 197–216.
Hermsen, B., Holland, R. W., & van Knippenberg, A. (2006). The happy act on impulse, the sad think twice. Mood as moderator of the impact of implicit and explicit attitudes on behavior. Manuscript submitted for publication.
Hofmann, W., & Friese, M. (in press). Impulses got the better of me: Alcohol moderates the influence of immediate affective reactions to food cues on eating behavior. *Journal of Abnormal Psychology.*
Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–1385.
Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology*, *43*, 497–504.
Huijding, J., & de Jong, P. J. (2006). Automatic associations with the sensory aspects of smoking: Positive in habitual smokers but negative in non-smokers. *Addictive Behaviors*, *31*, 182–186.
Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin*, *30*, 22–34.
Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, *91*, 16–32.
Karpinski, A., Steinman, R. B., & Hilton, J. L. (2005). Attitude importance as a moderator of the relationship between implicit and explicit attitude measures. *Personality and Social Psychology Bulletin*, *31*, 949–962.
Kim, D.-Y. (2004). The Implicit Life Satisfaction Measure. *Asian Journal of Social Psychology*, *7*, 236–262.
McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition*, *20*, 483–510.
Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, *85*, 1180–1192.
Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*, 625–664.
Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*, 166–180.

DOI: 10.1002/ejsp

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

Pappi, F. U. (1983). Die Links-Rechts-Dimension des deutschen Parteienspektrums und die Parteipräferenz-Profile der Wählerschaft [The left-wing/right-wing dimension of the German party spectrum and the preference profiles of the electorate]. In M. Kaase, & H-. D. Klingemann (Eds.), *Wahlen und politisches System: Analysen aus Anlaß der Bundestagswahl 1980 [Elections and the political system: Analyses on the occasion of the parliamentary elections]* (pp. 422–441). Opladen: Westdeutscher Verlag.

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single Attribute Implicit Association Tests (SA-IAT) for the assessment of unipolar constructs: The case of sociosexuality. *Experimental Psychology*, *53*, 283–291.

Perugini, M., & Banse, R. (2007). Editorial: Personality, implicit self-concept, and automaticity. *European Journal of Personality*, *21*, 257–261.

Richetin, J., Perugini, M., Adjali, I., & Hurling, R. (2007). The moderator role of intuitive vs. deliberate decision making for the predictive validity of implicit and explicit measures. *European Journal of Personality*, *21*, 529–546.

Schmidt, F. L., Hunter, J. E., & Raju, N. S. (1988). Validity generalization and situational specificity: A second look at the 75% rule and Fisher's z transformation. *Journal of Applied Psychology*, *73*, 665–672.

Schnabel, K., Banse, R., & Asendorpf, J. B. (2006). Assessment of implicit personality self-concept using the Implicit Association Test (IAT): Concurrent assessment of anxiousness and angriness. *British Journal of Social Psychology*, *45*, 373–396.

Steffens, M. C. (2005). Implicit and explicit attitudes towards lesbians and gay men. *Journal of Homosexuality*, *49*, 39–66.

Steffens, M. C., & Schulze-König, S. (2006). Predicting spontaneous Big Five behavior with Implicit Association Tests. *European Journal of Psychological Assessment*, *22*, 13–20.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* San Francisco: Jossey-Bass.

von Collani, G., & Blank, H. (2003). Persönlichkeitsmerkmale, soziale Überzeugungen und politische Parteienpräferenzen: Eine Internetbefragung zur Bundestagswahl 2002 [Personality, social beliefs, and political party preferences: An Internet survey concerning the Parliamentary elections 2002]. *Zeitschrift für Politische Psychologie*, *11*, 207–324.

Wigboldus, D. H. J., Holland, R. W., & van Knippenberg, A. (2004). Single target implicit associations. Unpublished manuscript.

Zayas, V., & Shoda, Y. (2005). Do automatic reactions elicited by thoughts of romantic partner, mother, and self relate to adult romantic attachment? *Personality and Social Psychology Bulletin*, *31*, 1011–1025.

DOI: 10.1002/ejsp

APPENDIX: TARGET AND ATTRIBUTE STIMULI FOR IMPLICIT MEASUREMENT TOOLS (STUDIES 1–2)

Political party stimuli

CDU/CSU	FDP	SPD	GREEN	PDS
Merkel (picture)	Gerhardt (picture)	Müntefering (picture)	Schlauch (picture)	Zimmer (picture)
Schäuble (name)	Genscher (name)	Scharping (name)*	Trittin (name)	Bisky (name)
CDU-emblem*	FDP-emblem	SPD-emblem**	GREEN-emblem	PDS-emblem
Stoiber (name)	Westerwelle (name)	Schröder (name)	Künast (name)	Gysi (name)
Merz (picture)	Möllemann (picture)	Struck (picture)	Fischer (picture)	Wagenknecht (picture)
*Koch		*Däubler-Gmelin		
		**Eichel		

Evaluative stimuli

POSITIV (positive)	NEGATIV (negative)
Freude (joy)	Gestank (stink)
Geschenk (present)	Gift (poison)
Liebe (love)	Katastrophe (catastrophe)
Paradies (paradise)*	Krankheit (disease)
Urlaub (vacation)**	Tod (death)*
*Gesundheit (health)	*Schmerz (pain)
**Lachen (laughter)	

Ideological stimuli

LINKS (left)	RECHTS (right)
Anarchismus (anarchy)	Bürgertum (middle classes)
Chaos (chaos)	Kapitalismus (capitalism)
Gewerkschaft (union)	Nationalflagge (national flag)
Marx (Marx - philosopher)	Strenge (austerity)
Punk (punk)	Tradition (tradition)
Ökosteuer (ecological taxes)	Unternehmen (business venture)
Sozialismus (socialism)	Wirtschaft (economy)

Note: Stimuli marked with asterisks were exchanged in Study 2 with the stimuli below.

DOI: 10.1002/ejsp