

© 2018, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission. The final article will be available, upon publication, via its DOI: 10.1037/xge0000383

Bar-Anan, Y., & Vianello, M. (2018). A multi-method multi-trait test of the dual-attitude perspective. *Journal of Experimental Psychology: General*, 147, 1264-1272.

A Multi-Method Multi-Trait Test of the Dual-Attitude Perspective

Yoav Bar-Anan

Ben Gurion University of the Negev, Beer Sheva, Israel

Michelangelo Vianello

University of Padova, Padova, Italy

The dual-attitude perspective posits that it is useful for research and theory to assume two distinct constructs: explicit and implicit attitudes (or automatic and deliberate evaluation). Much evidence supports this perspective but some important tests are missing, casting doubts on studies that relied on the perspective for inference. We used a multi-method multi-trait design to extensively test the validity of the dual perspective. The dataset ($N = 24,015$) included measurements of attitudes in three domains (race, politics, the self) with seven indirect measures, and at least three self-report measures for each attitude domain. The dual-attitude model fit the data better than a single-attitude model. Six of the seven indirect measures were related to the implicit construct more than to the explicit construct. The evidence supports the dual-attitude perspective, bolsters the validation of six indirect measures, and clears doubts from countless previous studies that used only one indirect measure to draw conclusions about implicit attitudes.

Keywords: Attitudes; Implicit Social Cognition; Implicit Measures; Automatic Evaluation;

Multi-Trait Multi-Method

Correspondence should be addressed to: Yoav Bar-Anan, Department of Psychology, Ben-Gurion University in the Negev Be'er Sheva, Israel. E-mail: baranany@bgu.ac.il. This project was supported by grants from the Israel Science Foundation [779/16], and the United States – Israel Binational Science Foundation [2013214] to Y. B.-A. Study materials and data appear at <https://osf.io/66ycm/>. We thank **Brian Nosek** for his contribution to this work, and Marco Perugini for useful comments.

A Multi-method multi-trait test of the dual-attitude perspective

Dual models of attitudes and evaluation share the perspective that it is useful to assume two separate theoretical constructs reflecting distinct processes or mental representations (Gawronski & Bodenhausen, 2011; Greenwald & Banaji, 1995; Fazio, 2007; Petty & Briñol, 2006; Wilson, Lindsey, & Schooler, 2000). Automatic processes reflect the implicit construct (often named *implicit attitudes*) and non-automatic processes reflect the explicit construct (often named *explicit attitudes*). Self-report measures tap the explicit construct. To investigate the implicit construct, researchers use measures that infer the evaluation from a behavior that people initiate without intention to express an evaluation of the attitude object (Gawronski & De Houwer, 2014; Wittenbrink & Schwarz, 2007). These indirect measures help investigate evaluation and stereotypes (Gawronski & Payne, 2011; Nosek, Hawkins, & Frazier, 2012; Petty, Fazio, & Brinol, 2012) with many practical implications for various fields (e.g., Bluemke & Teige-Mocigemba, 2014; Krieger & Fiske, 2006; Roefs et al., 2011; Uhlmann et al., 2012). Nevertheless, the validation of the implicit construct and its measures is still an on-going effort, employing various research strategies, and targeting various open questions (for a review, see Gawronski & De Houwer, 2014).

Validation of the implicit construct and its measures has come mostly from *predictive validity* studies. Those studies test the relation between a supposed measure of implicit attitudes and a measure of another construct that—according to the dual perspective—is related to implicit attitudes. Evidence that *indirect attitude measures capture attitudes* comes mostly from research that found positive relations between the indirect measures and standard self-report attitude measures (e.g., Hofmann, Gawronski, Gschwendner, & Schmitt, 2005; Nosek, 2007). Evidence that *indirect measures capture a psychological construct that reflects automatic processes and is distinct from explicit attitudes* comes from studies that found that

indirect measures predicted unintentional evaluative behavior better than self-report measures, whereas self-report measures predicted intentional evaluative behavior better than indirect measures (e.g., Dasgupta & Rivera, 2006; Dovidio, Kawakami, & Gaertner, 2002; Fazio, Jackson, Dunton, & Williams, 1995; Friese, Hofmann, & Wänke, 2008).

Another research strategy for construct validation is the test of convergent and discriminant validity (Campbell & Fiske, 1959; John & Benet-Martinez, 2000). It is a concurrent test for the theory that suggested a construct and for the measures developed to tap that construct. If multiple instruments that have been developed to measure a theoretical construct are found to share variance (accounting for method variance), then the existence of the theoretical construct is empirically supported and the measures are construct validated.

According to the dual perspective, indirect measures are positively related to each other (convergent validity) and the relation between them is stronger than their relation to self-report measures (discriminant validity). Convergent and discriminant validity evidence comes from studies that collected data with multiple indirect and direct measures (Cunningham, Preacher, & Banaji, 2001; Ranganath, Smith, & Nosek, 2008). Those studies supported the dual perspective: The model with the best fit to the data loaded measures developed to tap implicit attitudes onto one factor and the supposed measures of explicit attitudes onto a second factor. However, because previous studies measured attitudes in a single domain, they could not distinguish attitude variance from variance related to the measurement methods. To overcome this weakness, the present study used a multi-trait-multi-method (MTMM) design (Campbell & Fiske, 1959).

In the MTMM design, participants complete measures that vary in their method (e.g., open-ended questions versus multiple-choice questions) and in the constructs they are designed to measure (e.g., verbal versus mathematical ability). This design tests convergent validity (the correlation between measures of the same construct) and discriminant validity

(the correlation between measures that share method but not construct). MTMM designs allow separating method-specific variance from construct-specific variance, for validating measures as sensitive to variance in a certain theoretical construct and not only to variance in the reaction to the measurement method.

We used more than eleven attitude measures in three domains. Some of the measures were developed to capture the implicit construct, whereas the others were developed to capture the explicit construct. Using Structural Equation Modelling (SEM) analysis, we examined the dual perspective assumption that the best model for our data would include six psychological constructs: implicit and explicit attitudes for each attitude domain. This is the first study to use multiple measures for each construct in multiple attitude domains. Therefore, our dataset is the first that allows examining whether the measures map to two different theoretical variables, and whether those variables are attitudinal (i.e., different for each attitude domain; and show stronger explicit-implicit correlation within each attitude domain than across domains). As a secondary investigation, we examined the validity of each individual measure by testing whether it was related to the construct it was developed to tap more than to the other construct.

Method

The study (approved by University of Virginia's IRB) was designed for the purpose of the present analysis. Full details about each measure procedure, stimuli materials, and scoring were reported in Bar-Anan and Nosek (2014; osf.io/qf9jx/) who used the dataset to compare the psychometric qualities of the indirect measures.

Participants. Participants volunteered at the Project Implicit website (Nosek, 2005) after registering for research. They were randomly assigned to the study from a pool of available studies. Stopping data collection did not depend on the analysis, conducted only

afterwards. We included in the analysis 24,015 participants (63% women, 36% men, 1% unknown; $M_{age} = 29.1$, $SD_{age} = 12.0$) who completed at least one measure.

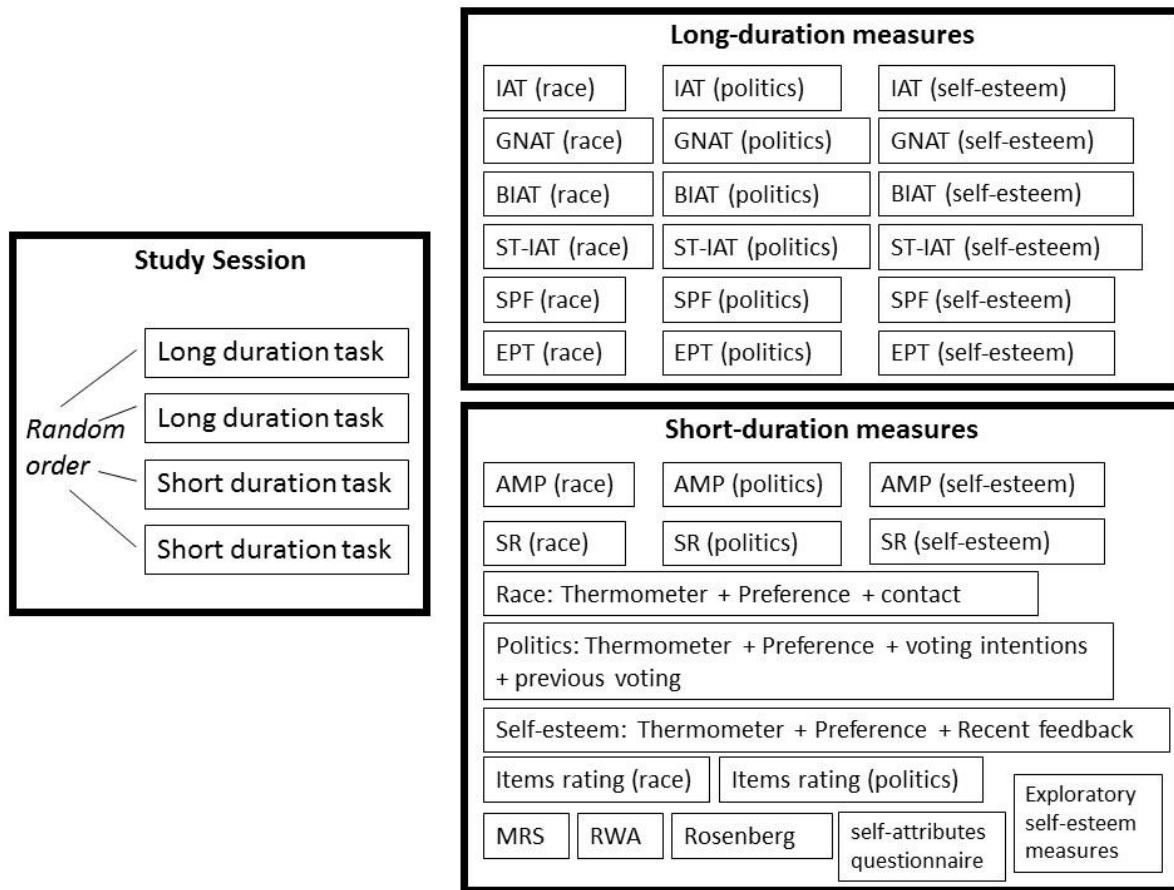


Figure 1. The study procedure. Each study session included two long and two short tasks, selected and ordered randomly. The tasks are listed on the right. Measures that share a rectangle appeared in the same questionnaire. IAT = Implicit Association Test; GNAT = Go/No-Go Association Test; BIAT = Brief IAT; ST-IAT = Single-target IAT. SPF = Sorting Paired Features; EPT = Evaluative Priming Task; AMP = Affect Misattribution Procedure; SR = Speeded rating; MRS = Modern Racism Scale; RWA = Right-wing Authoritarianism.

Procedure and Materials. Participants were assigned, in a random order, to two “long-duration” and two “short-duration” measures (Figure 1). Thus, the design for the factor analysis was *planned incomplete* (Graham, Taylor, Olchowski, & Cumsille, 2006). This design minimizes parameter estimation distortions that occur when data are not missing at random. We did not constrain the selection of tasks, other than preventing repetition of the same measure in the same session. Participants could initiate additional sessions up to 32 times and could receive identical measures from previous sessions. The present analysis did

not use scores from repeated measures. The procedure setup (Figure 2) resulted with variability in sample size for each measure (*mean N* = 3,709, *median N* = 3,085) and for each combination of two measures (*min N* = 259, *max N* = 3,869, *mean N* = 756, *median N* = 562), allowing good to high precision of estimates.

We measured attitudes in three domains: Politics (Republicans/Democrats), race (Black people/White people), and self-esteem (Self/Others). We used eight measures developed to tap implicit attitudes: the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), the Evaluative Priming task (EPT; Fazio, Sanbonmatsu, Powell, & Kardes, 1986), the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005), the Brief Implicit Association Test (BIAT; Sriram & Greenwald, 2009), the Go/No-go Association Task (GNAT; Nosek & Banaji, 2001), the Single-Target Implicit Association Test (ST-IAT; Karpinski & Steinman, 2006), the Sorting Paired-Features task (SPF; Bar-Anan, Nosek, & Vianello, 2009), and speeded rating (1200ms deadline) of the items and category labels from the indirect measures.

The supposed explicit attitude measures were self-reported preference between two categories (e.g., Black people versus White people), feeling thermometer of the two categories, rating of the images used as items in the indirect measures (politics and race only), the modern racism scale (MRS¹; McConahay, 1983), political identification (liberal/conservative), the Rosenberg self-esteem scale (Rosenberg, 1965), the self-attributes questionnaire for self-esteem (Pelham & Swann, 1989), and self-reported self-esteem rating (one item). Another direct measure of political attitudes—the Right-wing Authoritarianism scale (Altemeyer, 1981)—was administered but not included because it prevented identification of the models for lack of shared information across variables. In total, the

¹ Previous research suggests that MRS is not an implicit attitudes measure (Fazio et al., 1995; Greenwald et al., 1998).

analysis included 24 measures developed to tap implicit attitudes (eight techniques by three attitude domains), and 14 measures of, supposedly, explicit attitudes (four race measures, four politics measures, and six self-esteem measures).

Results

The number of observations, mean, standard-deviation, and internal consistency of each of the 37 measures appear in Table 1. A correlation matrix of all the variables is available online, at osf.io/n6bsp/.

Validity of the Implicit/Explicit Distinction

For the main analysis, we employed the general CFA approach to the covariance structure analysis of MTMM correlation matrices drawing from Widaman (1985). We specified and estimated a baseline model with six attitudes (implicit and explicit attitudes on each of the three topics) and two methods (for the supposed measures of implicit attitudes and the supposed measures of explicit attitudes). When speeded rating was assumed to measure implicit attitudes, the model estimation consistently failed, indicating that it was wrong. Hence, we specified the speeded ratings to load on explicit attitudes only, resulting in the baseline model depicted in Figure 2. We compared the baseline model against more specific nested models. Parameters have been estimated by Full Information Maximum Likelihood (Arbuckle, 1996).

The Correlated Trait-Correlated Method model with six trait factors and two method factors (model 6CT-2CM; C stands for correlated, T for trait, and M for method; illustrated in Figure 2; Table 2 shows the correlations between the latent constructs) was a good fit to the data ($\chi^2_{(576)}=1390.566$; $RMSEA=0.008$, 90% CI [0.007 0.008]; $CFI=0.973$; $TLI=0.967$), providing initial evidence of validity. Focal trait loadings were all significantly different from zero, higher on average than focal method loadings, and often at least moderate in size.

Table 1

All the measures: Descriptive Statistics

<i>Measure</i>	<i>N</i>	<i>Mean</i>	<i>STD</i>	<i>Min</i>	<i>Max</i>	<i>IC</i>
IAT (politics)	2,949	0.271	0.550	-1.399	1.570	.93
BIAT (politics)	2,630	0.273	0.434	-1.192	1.454	.89
GNAT (politics)	2,771	-0.005	0.888	-3.508	4.083	.84
ST-IAT (politics)	2,771	0.249	0.588	-1.682	2.231	.84
SPF (politics)	2,736	0.123	0.559	-1.855	1.964	.59
AMP (politics)	3,187	0.085	0.292	-1.000	1.000	.81
EPT (politics)	2,825	0.144	0.516	-2.264	2.139	.63
Preference (politics)	3,511	1.141	1.822	-3	3	-
Thermometer (politics)	3,568	2.407	4.000	-10	10	-
Items (politics)	3,717	1.298	2.721	-8.000	8.000	.85
Speeded (politics)	3,391	0.545	1.143	-3.000	3.000	.89
Political Identity	22,175	0.821	1.635	-3	3	-
IAT (race)	3,041	0.296	0.400	-1.108	1.433	.86
BIAT (race)	2,780	0.238	0.333	-0.959	1.272	.81
GNAT (race)	2,982	0.002	0.806	-4.047	2.871	.70
ST-IAT (race)	2,953	0.101	0.498	-2.058	1.737	.74
SPF (race)	2,944	0.119	0.491	-1.762	1.754	.52
AMP (race)	3,085	-0.051	0.223	-1.000	1.000	.66
EPT (race)	2,943	0.034	0.462	-2.114	1.566	.54
Preference (race)	3,659	0.398	1.028	-3	3	-
Thermometer (race)	3,778	0.546	1.781	-10	10	-
Item (race)	3,834	-0.834	1.044	-8.000	8.000	.63
Speeded (race)	3,274	-0.090	0.637	-3.000	3.000	.80
MRS	3,795	2.020	0.907	1.000	6.000	.81
IAT (self)	2,941	0.464	0.352	-0.904	1.527	.82
BIAT (self)	2,633	0.308	0.299	-1.038	1.303	.76
GNAT (self)	2,203	-0.001	0.818	-4.789	4.597	.65
ST-IAT (self)	2,858	0.255	0.442	-1.593	1.874	.70
SPF (self)	2,900	0.464	0.476	-1.405	1.803	.48
AMP (self)	3,172	0.032	0.193	-0.917	0.958	.55
EPT (self)	3,067	0.202	0.460	-1.690	1.730	.54
Preference (self)	3,856	0.588	1.351	-3	3	-
Thermometer (self)	3,869	0.759	2.079	-9	10	-
Direct self esteem	3,655	0.586	1.361	-3	3	-
Speeded (self)	3,323	0.305	0.706	-2.900	3.000	.69
Rosenberg	3,808	4.678	0.939	1.000	6.000	.90
Self-attributes questionnaire	3,665	4.568	0.794	1.000	7.000	.58

Notes. IC = Internal consistency; IC was computed using parcels for the indirect measures and the speeded rating, and the item scores for the scales and the item rating measures; IAT = Implicit Association Test; GNAT = Go/No-Go Association Test; BIAT = Brief IAT; ST-IAT = Single-target IAT. SPF = Sorting Paired Features; EPT = Evaluative Priming Task; AMP = Affect Misattribution Procedure; SR = Speeded rating; MRS = Modern Racism Scale; RWA = Right-wing Authoritarianism.

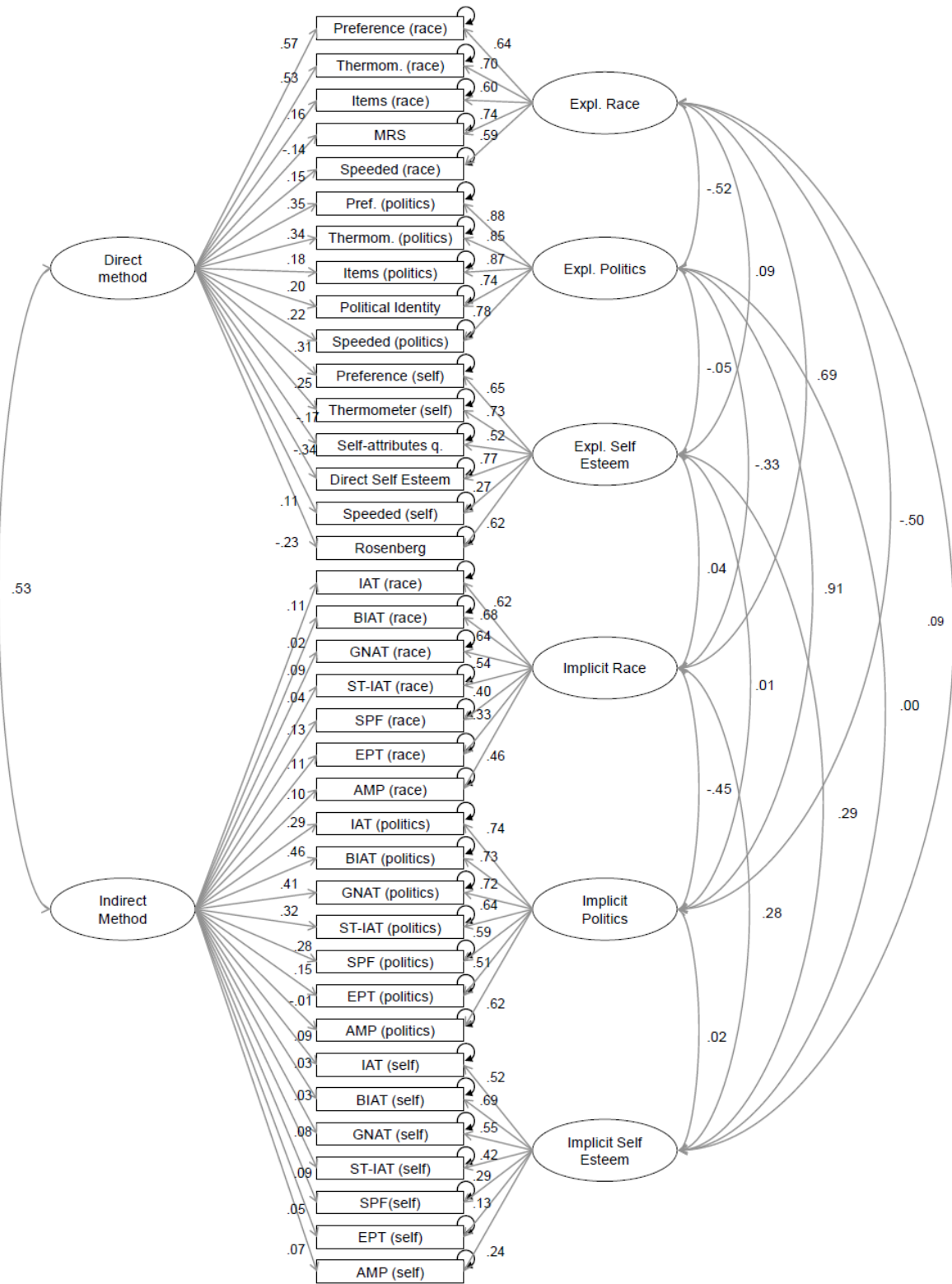


Figure 2. The Correlated Trait-Correlated Method (CT-CM) model with multiple indicators for each Trait-Method Unit, one factor for each trait and one factor for each method that has been used as a baseline against which more constrained nested models are compared ($\chi^2_{(576)}=1390.566$; RMSEA=0.008, 90% CI [0.007 0.008]; CFI=0.973; TLI=0.967). Correlations are disattenuated; Correlations higher than .06 and standardized loadings higher than .38 are significantly different from zero; IAT = Implicit Association Test; GNAT = Go/No-Go Association Test; BIAT = Brief IAT; ST-IAT = Single-target IAT. SPF = Sorting Paired Features; EPT = Evaluative Priming Task; AMP = Affect Misattribution Procedure; SR = Speeded rating; MRS = Modern Racism Scale; RWA = Right-wing Authoritarianism.

Table 2: Correlations between the latent attitude factors in the dual-attitude model

	Implicit Politics	Implicit Race	Implicit Self	Explicit Politics	Explicit Race	Explicit Self
Implicit Politics		-.45	.02	.91	-.50	.01
Implicit Race			.28	-.33	.69	.04
Implicit Self				.00	.09	.29
Explicit Politics					-.52	-.05
Explicit Race						.09

Note. Correlations higher than .06 are significantly different from zero.

Table 3 presents statistics for convergent and discriminant validity. The classic test of *convergent validity* in MTMM models is to examine the difference in fit between the 6CT-2CM model and the 2CM model (no trait factors). That test strongly supported the former model ($\Delta\chi^2_{(52)}=7185.99$, $\Delta CFI=.26$), suggesting that the indicators are effective in capturing inter-individual differences on the latent traits they are supposed to measure. The mirror-image test on method factors compared the baseline to a model without method factors (Model 6CT, $\Delta\chi^2_{(38)}=654.55$, $\Delta CFI=.02$). This test verifies that the fit loss when eliminating method factors is smaller than the fit loss when eliminating trait factors. This was indeed the case, suggesting that the trait variance shared among measures is much higher than shared method variance, a property of every good measure.

We conducted tests of discriminant validity among both traits and methods. First, we compared the baseline (6CT-2CM) with a more specific model in which all trait factors correlate perfectly (1T-2CM). The difference in fit between the models was huge ($\Delta\chi^2_{(20)}=5451.52$, $\Delta CFI=.177$), strongly supporting discriminant validity among traits (indeed, attitudes toward the different topics are certainly different constructs). The baseline model was better ($\Delta\chi^2_{(1)}=35.27$) than a model forcing method factors to be orthogonal (6CT-2M), although the increase in fit was very small ($\Delta CFI=.001$). In the 6CT-2CM model, the correlation between the two method factors was higher than we expected ($\phi=.53$), which might suggest that a non-attitudinal component (e.g., cognitive capabilities or motivation) contributes to correlations observed between indirect and self-report measures. On the other

hand, this finding might be spurious because correlation between method factors is sometimes inflated with trait variance when some trait factors (e.g., implicit and explicit political attitudes in the present dataset) are highly correlated (Marsh, 1989).

For answering the main research question, we estimated and compared to the baseline a model with one attitude factor per trait (3CT-2CM). That model fit the data very poorly and was inferior to the baseline (Table 3). For a more stringent test, we estimated three more specific models in which implicit and explicit attitude measures of one of the topics load on the same factor (5CT-2CM). These three models showed a significant misfit compared to the baseline ($\Delta\chi^2_{(6)} > 57.7$, $\Delta CFI > .003$; Table 3). These results confirm the viability of the distinction between implicit and explicit constructs.

Validity of Individual Measures

To test the validity of individual measures, we examined their multidimensionality. We estimated 37 models in which, for every indicator (e.g., race IAT), we released the constraint that forced the cross-loading (e.g., race IAT on explicit race attitude) to zero. We compared the fit of the constrained and unconstrained models and examined whether focal loadings are higher than cross-loadings (Table 4). All indirect methods showed cross-loadings on the explicit construct that are either zero or much smaller than focal loadings on the implicit construct, with the exception of the AMP that showed loadings of similar size on the explicit and the implicit constructs. The three speeded rating measures were slightly influenced by implicit constructs for race and politics and moderately so in the case of implicit self-esteem. Taken together, these tests provide evidence of good convergent and discriminant validity across all the indirect measures, except for the AMP, which seems to be moderately influenced by explicit processes. Speeded rating seems sensitive to the explicit constructs more than to the implicit constructs.

Table 3

Analysis of convergent and discriminant validity

Model description	χ^2	<i>Df</i>	<i>RMSEA</i> [90% <i>CI</i>]	<i>CFI</i>	<i>TLI</i>	$\Delta\chi^2$	$p(\Delta\chi^2)$
Baseline							
6 correlated traits, 2 correlated methods (6CT-2CM)	1390.57	576	.008 [.007 .008]	.973	.967		
Convergent Validity							
2 correlated methods, no traits (2CM)	9231.12	628	.023 [.023 .024]	.710	.68	7185.99	<.001
6 correlated traits, no methods (6CT)	2045.21	614	.010 [.009 .010]	.952	.946	654.55	<.001
six orthogonal traits – 2 correlated methods (6T-2CM)	2230.427	591	0.011 [.010 .011]	.946	.935	839.86	<.001
Discriminant Validity							
6 perfectly correlated traits, 2 correlated methods (1T-2CM)	6842.088	596	.021 [.020 .021]	.796	.759	5451.52	<.001
3 correlated traits, 2 correlated methods (3CT-2CM)	4818.95	149	.017 [.017 .018]	.860	.833	3428.38	<.001
6 correlated traits, 2 orthogonal methods (6CT-2M)	1426.29	577	.008 [.007 .008]	.972	.966	35.72	<.001
6 correlated traits, 2 perfectly correlated methods (6CT-1M)	1580.91	577	.009 [.008 .009]	.967	.959	190.34	<.001
direct and indirect measures of race attitude load onto the same trait (5CT-2CM)	1578.61	582	.008 [.008 .009]	.967	.96	188.044	<.001
direct and indirect measures of politics attitude load onto the same trait (5CT-2CM)	1448.233	582	.008 [.007 .008]	.970	.965	57.667	<.001
direct and indirect measures self esteem load onto the same trait (5CT-2CM)	1618.062	582	.009 [.008 .009]	.966	.958	227.496	<.001

Notes: tests of the difference in chi square (last two columns) always compare the model against the baseline; A likelihood ratio test of the difference in fit is given by $\Delta\chi^2$; Differences in practical fit that are less or not influenced by sample size are given by *RMSEA*, *CFI* and *TLI*.

Table 4

Test of cross-loadings: single-indicator tests of convergent and discriminant validity

Cross-loading added	χ^2	RMSEA [90% CI]	CFI	TLI	$\Delta\chi^2$	$p(\Delta\chi^2)$	Loading on implicit construct	Loading on explicit construct
Baseline	1390.57	.008 [.007 .008]	.973	.967				
IAT on explicit race	1388.13	.008 [.007 .008]	.972	.966	2.43	.118	.74	-.03
IAT on explicit politics	1383.56	.008 [.007 .008]	.973	.967	7.01	.008	.68	-.004
IAT on explicit self-esteem	1390.54	.008 [.007 .008]	.973	.967	.02	.880	.74	.01
BIAT on explicit race	1383.17	.008 [.007 .008]	.973	.967	6.85	.009	.87	-.20
BIAT on explicit politics	1382.71	.008 [.007 .008]	.973	.967	7.85	.005	.52	.16
BIAT on explicit self-esteem	1389.93	.008 [.007 .008]	.973	.967	.63	.427	.66	.04
AMP on explicit race	1341.44	.007 [.007 .008]	.973	.967	49.12	$2.4*10^{-12}$.12	.37
AMP on explicit politics	1375.35	.008 [.007 .008]	.973	.967	15.21	$9.6*10^{-5}$.31	.29
AMP on explicit self esteem	1383.12	.008 [.007 .008]	.973	.967	7.44	.006	.19	.11
GNAT on explicit race	1383.31	.008 [.007 .008]	.973	.967	7.26	.007	.83	-.19
GNAT on explicit politics	1383.20	.008 [.007 .008]	.973	.967	7.36	.007	.56	.10
GNAT on explicit self esteem	1386.34	.008 [.007 .008]	.973	.967	4.23	.040	.61	-.20
EPT on explicit race	1389.44	.008 [.007 .008]	.973	.967	1.13	.290	.27	.06
EPT on explicit politics	1376.83	.008 [.007 .008]	.973	.967	13.74	$2.1*10^{-4}$.94	-.49
EPT on explicit self esteem	1388.81	.008 [.007 .008]	.973	.967	1.76	.180	.10	.05
ST-IAT on explicit race	1390.03	.008 [.007 .008]	.973	.967	.54	.460	.59	-.05
ST-IAT on explicit politics	1382.15	.008 [.007 .008]	.973	.967	8.42	.003	.79	-.20
ST-IAT on explicit self esteem	1390.56	.008 [.007 .008]	.973	.967	.02	.880	.42	-.01
SPF on explicit race	1385.84	.008 [.007 .008]	.973	.967	.54	.460	.27	-.05
SPF on explicit politics	1390.12	.008 [.007 .008]	.973	.967	.44	.510	.49	-.20
SPF on imp self esteem	1389.51	.008 [.007 .008]	.973	.967	1.06	.300	.31	-.04
Speeded on implicit race	1366.46	.008 [.007 .008]	.974	.968	24.11	$9.11*10^{-7}$.21	.43
Speeded on implicit politics	1379.95	.008 [.007 .008]	.973	.967	10.62	.001	.20	.86
Speeded on implicit self esteem	1352.60	.008 [.007 .008]	.974	.968	37.96	$7.21*10^{-10}$.25	.17

Self-reported preference on implicit race	1390.18	.008 [.007 .008]	.973	.967	0.39	.530	.03	<i>.61</i>
Self-reported preference on implicit politics	1367.42	.008 [.007 .008]	.974	.968	23.14	2.0×10^{-6}	-.25	<i>1.00</i>
Self-reported preference on implicit self esteem	1390.01	.008 [.007 .008]	.973	.967	0.39	.53	-.03	<i>.66</i>
Thermometer on implicit race	1387.10	.008 [.007 .008]	.973	.967	3.47	.06	-.08	<i>.77</i>
Thermometer on implicit politics	1372.92	.008 [.007 .008]	.973	.968	17.64	2.7×10^{-5}	-.33	<i>.84</i>
Thermometer on exp implicit self esteem	1390.15	.008 [.007 .008]	.973	.967	0.41	.52	.02	<i>.72</i>
Item ratings on implicit race	1383.50	.008 [.007 .008]	.973	.967	7.06	.008	-.01	<i>.63</i>
Item ratings on implicit politics	1345.22	.008 [.007 .008]	.973	.968	45.34	1.65×10^{-11}	.39	<i>.49</i>
Item ratings on implicit self esteem	1390.15	.008 [.007 .008]	.973	.967	0.41	.52	-.02	<i>.53</i>
Rosenberg on imp self esteem	1388.61	.008 [.007 .008]	.973	.967	1.96	.16	.06	<i>.60</i>
Direct rating on impl self esteem	1382.99	.008 [.007 .008]	.973	.967	7.58	.006	-.11	<i>.81</i>

Note: All reported loadings are standardized. $df(\text{baseline}) = 576$; $df(\text{all other models}) = 575$; Focal loadings are reported in italics. The cross-loadings for Modern Racism Scale and reported political identity prevented the minimization algorithm to find an acceptable solution and are not reported. IAT = Implicit Association Test; GNAT = Go/No-Go Association Test; BIAT = Brief IAT; ST-IAT = Single-target IAT. SPF = Sorting Paired Features; EPT = Evaluative Priming Task; AMP = Affect Misattribution Procedure; SR = Speeded rating; MRS = Modern Racism Scale; RWA = Right-wing Authoritarianism.

General Discussion

One of the best methods to test whether a theoretical construct is useful for empirical research is to examine whether different measures developed to reflect that construct share variance that is not shared with measures of other constructs. Failure to find such evidence casts doubt on the theory, and prevents considering results found with the different measures as pertaining to the same construct. The present research is a comprehensive multi-trait-multi-method (MTMM) test of the theory that *implicit attitude* is a theoretical construct related but distinct from the construct captured by standard self-reported attitude measures. This is the first research to use multiple indirect and multiple direct measures to test whether there is convergent evidence that the theoretical construct exists, and multiple attitude domains to test whether the construct is attitudinal (i.e., varies by attitude domain) rather than methodological (e.g., captures cognitive ability to perform the tasks).

We found strong support for the distinction of implicit and explicit attitudes, which can be considered two distinct but related constructs, rather than facets of the same attitudinal construct. Although that distinction is widespread (Gawronski & Bodenhausen, 2011; Fazio, 2007; Petty, Briñol, & DeMarree, 2007; Wilson et al., 2000; but see Kruglanski & Gigerenzer, 2011), the present research is the first to confirm it with a factor analysis on multiple attitude measures developed to tap those constructs. Another important contribution of the present research is that it provides strong retrospective support for countless previous studies that used only one of the indirect measures to draw conclusions about the construct *implicit attitude*.

Further, our research provides convergent validation evidence for the individual measures. Five of the indirect measures showed acceptable convergent and discriminant validity – high loadings on the implicit constructs and low loadings on the explicit constructs. A sixth measure, the EPT, showed smaller loadings on the implicit constructs, but the

loadings were still always considerably higher than the EPT's loadings on the explicit constructs. This provides support that the construct captured by EPT is related to the implicit construct more than to the explicit construct.

Only one indirect measure, the AMP, did not show clear evidence for belonging to the implicit construct more than to the explicit construct. Therefore, research that relies on the AMP as a measure of implicit attitudes might suffer from questionable validity. The evidence that the AMP predicts judgment biases and behaviors that people might prefer to inhibit (Cameron, Brown-Iannuzzi, & Payne, 2012; Payne & Lundberg, 2014) is compatible with the possibility that the AMP is equally related to implicit and explicit constructs. Alternatively, the AMP captures a third construct, distinct from the constructs measured by other indirect measures and by self-report measures. The high rate of missing data prevented testing models with more than one implicit construct. Therefore, we hold off pessimistic conclusions about the AMP, pending more evidence.

The present research also questions previous suggestions that speeded rating tap implicit attitudes (Wilson et al., 2000; Ranganath et al., 2008). The higher loading of speeded ratings on the explicit than on the implicit construct are incompatible with previous findings of better fit for models that grouped this measure with implicit measures (Ranganath et al., 2008). Because the previous research did not include separate method variables in the tested models, it is possible that the previous findings reflected a non-attitudinal construct shared by indirect measures and the speeded-rating, such as inter-individual differences in reaction speed, task-switching abilities, and other confounding cognitive processes.

As a study of convergent and discriminant validity, the present research did not directly test whether the implicit construct has anything to do with automatic processes. That test requires an outcome measure of a construct that is not implicit attitudes but is—according to the theory—related to implicit attitudes more than to explicit attitudes. Currently, there is

no such measure that is highly replicable. Without such a measure, the validation of indirect measures as capturing a shared construct that reflects automatic evaluative processes remains incomplete. Here, we used an informative alternative approach that validates each supposed measure of implicit attitudes by testing whether it is related to other supposed measures of implicit attitudes more than to supposed measures of explicit attitudes. Another informative validation approach is to examine whether interventions that are supposed to change implicit attitudes have a similar impact on multiple (supposed) implicit attitude measures (Borsboom, Mellenbergh, & Heerden, 2004; e.g., Moran, Bar-Anan, & Nosek, in press).

Context of the Research

Due to the large number of measures and participants, the present research has been a challenging undertaking in the pursuit of more knowledge about the measures that have dominated implicit social cognition research for the past few decades. We continue to pursue further knowledge about indirect measures of social cognition with many projects, including an MTMM study of indirect measures of gender stereotypes, comparative research on the sensitivity of different indirect measures to evaluative information, a search for replicable outcome measures for predictive validity studies, and attempts to further improve current indirect measures. Because knowledge about automatic processes in social cognition is still limited, each of these lines of research, just like the present research, concurrently investigates questions about psychological constructs and about their measures. The concurrent scientific investigation sometimes poses challenges to interpretation (e.g., surprising results might reflect novel findings about the construct or only about its supposed measure), but often rewards with simultaneous advances in knowledge of methodology and theory.

References

- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. *Advanced structural equation modeling: Issues and techniques*, 243, 277.
- Altemeyer, B. (1981). *Right-wing authoritarianism*. Winnipeg, MB: University of Manitoba Press.
- Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior research methods*, 46(3), 668-688.
- Bar-Anan, Y., Nosek, B. A., & Vianello, M. (2009). The sorting paired features task: A measure of association strengths. *Experimental Psychology*, 56, 329-343.
- Bluemke, M., & Teige-Mocigemba, S. (2014). Automatic processes in aggression: Conceptual and assessment issues. *Aggressive behavior*, 41, 44-50.
- Borsboom, D., Mellenbergh, G.J., & Van Heerden, J. (2004). *The concept of validity*. *Psychological Review*, 111, 1061-1071.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56, 81-105.
- Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, 16, 330-350.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological science*, 12, 163-170.
- Dasgupta, N., & Rivera, L. M. (2006). From automatic antigay prejudice to behavior: the moderating role of conscious beliefs about gender and behavioral control. *Journal of personality and social psychology*, 91, 268-280.
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of personality and social psychology*, 82, 62-68.
- Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, 25, 603-637.
- Fazio, R.H., Jackson, J.R., Dunton, B.C., & Williams, C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013-1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.

Friese, M., Hofmann, W., & Wänke, M. (2008). 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*, *47*, 397-419.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, *44*, 59-127.

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. *Handbook of research methods in social and personality psychology* (2nd ed.). New York, NY: Cambridge University Press.

Gawronski, B., & Payne, B. K. (Eds.). (2011). *Handbook of implicit social cognition: Measurement, theory, and applications*. Guilford Press.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*(4), 323.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*, 4-27.

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.

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

John, O. P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York: Cambridge University Press

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.

Krieger, L. H., & Fiske, S. T. (2006). Behavioral realism in employment discrimination law: Implicit bias and disparate treatment. *California Law Review*, *94*, 997-1062.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological review*, *118*, 97-109.

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, *13*, 335-361.

McConahay, J. B. (1983). Modern racism and modern discrimination. *Personality and Social Psychology Bulletin*, *9*, 551-558.

Moran, T., Bar-Anan, Y., & Nosek, B. A. (in press). The effect of the validity of co-occurrence on automatic and deliberate evaluation. *European Journal of Social Psychology*.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*, 565-584.

Nosek, B. A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science*, *16*, 65-69.

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

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2012). Implicit social cognition. *Handbook of social cognition*, 31-53.

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

Payne, B. K., & Lundberg, K. (2014). The Affect Misattribution Procedure: Ten Years of Evidence on Reliability, Validity, and Mechanisms. *Social and Personality Psychology Compass*, *8*(12), 672-686.

Pelham, B. W., & Swann, W. B. (1989). From self-conceptions to self-worth: on the sources and structure of global self-esteem. *Journal of personality and social psychology*, *57*, 672-680.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The meta-cognitive model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition*, *25*, 657-686.

Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.). (2012). *Attitudes: Insights from the new implicit measures*. Psychology Press.

Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology, 44*, 386–396.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Sriram, N., & Greenwald, A. G. (2009). The brief implicit association test. *Experimental Psychology, 56*, 283–294.

Uhlmann, E. L., Leavitt, K., Menges, J. I., Koopman, J., Howe, M., & Johnson, R. E. (2012). Getting explicit about the implicit: A taxonomy of implicit measures and guide for their use in organizational research. *Organizational Research Methods, 15*, 553-601.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement, 9*, 1-26.

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

Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. Guilford Press.