



Expanding the Scope of Implicit Personality Assessment: An Examination of the Questionnaire-Based Implicit Association Test (qIAT)

Ariela Friedman, Benjamin A. Katz, Inbal Halavy Cohen, and Iftah Yovel

Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel

ABSTRACT

Self-report questionnaires can only yield information that people are able and willing to report, but implicit assessment methods are not commonly used in mainstream personality research. The Questionnaire-based Implicit Association Test (qIAT) was designed to address the limitations associated with the conventional self-concept IAT, and it enables an indirect assessment that is based on the items of standard self-reports. The present studies examined the psychometric properties of the qIAT across two personality constructs. Study 1 ($N = 528$) provided support for the internal consistency, test-retest reliability and convergent and discriminant validity of the qIAT that measured extraversion. Study 2 ($N = 164$) supported the reliability and validity of the qIAT assessment of conscientiousness, which predicted who returned to complete the second session of the study two weeks later, for which participants were paid in advance. This same prediction effect was marginally significant in Study 3 ($N = 200$), and across both Studies 2 and 3 the qIAT prediction of the criterion behavior was incremental to the parallel self-report questionnaire. Taken together, findings support the reliability and validity of the qIAT, which enables the indirect measurement of a wide variety of distinct personality constructs, currently measured only by self-report scales.

ARTICLE HISTORY

Received 7 August 2019
Accepted 27 February 2020

KEYWORDS

Implicit personality assessment; extraversion; conscientiousness
questionnaire-based implicit association test

Self-report questionnaires are used extensively in a wide variety of settings for the assessment of numerous different constructs. These instruments assess psychological phenomena that are introspectively available to respondents and are deliberately revealed by them. However, the information that people explicitly reveal about themselves is not always accurate (Schwarz, 1999). Inaccurate responses to questionnaire items may occur unintentionally, because knowledge about the self can be biased and incomplete (Wilson, 2009). Or, people may intentionally use self-presentation strategies to bias results (Schwarz, 1999). Indirect assessment methods are not based on explicit, introspective processes, and they therefore carry an appealing promise: providing information that is not tapped by ordinary self-report instruments (Nosek et al., 2011). Indeed, rather than relying solely on self-report, assessment in social cognition research (e.g., on attitudes toward various groups of people) often includes such implicit measurement tasks (Gawronski & Payne, 2010). This is not the case, however, in the field of personality. Due to the limitations of current indirect assessment procedures that measure aspects of the self, such methods are not commonly employed in mainstream personality research (De Cuyper et al., 2017). The studies presented here focus on the Questionnaire-based Implicit Association Test (qIAT), which was designed to address these limitations.

Indirect assessment of the self-concept

Over the past several decades, researchers developed different measurement paradigms collectively referred to as implicit or indirect assessment tasks (for a review see Gawronski & De Houwer, 2014). Designed to minimize the ability of people to consciously and intentionally bias their responses, the assessment procedures in these tasks are based on processes that are relatively automatic. The extent to which these processes are indeed unintentional, unconscious, and uncontrollable is a matter of debate (De Houwer et al., 2009). Nonetheless, in contrast to self-report measures, the assessment procedures in these tasks are not based on respondents' deliberate introspection of the to-be-measured construct, or on any awareness of the connection between their responses and the assessed content (Nosek et al., 2011).

The most extensively used indirect assessment method (Nosek et al., 2011) is the Implicit Association Test (IAT; Greenwald et al., 1998), which is a double-categorization reaction-time task. Notwithstanding the general concerns that have been raised regarding the predictive utility of the IAT (e.g., Carlsson & Agerström, 2016; Schimmack, 2019), it has been the most widely used and the most reliable method for the indirect measurement of personality-related aspects of the self (e.g., $r = .74$, De Cuyper et al., 2017; $r = .75$, Krause et al., 2011). When assessing the self-concept, the score provided by the IAT reflects the extent to which people associate the self with a certain attribute (e.g., the broad

trait of extraversion; Back et al., 2009). This is typically done by measuring the relative speed with which respondents classify words related to the measured attribute (e.g., “outgoing” versus “reserved”) when they share the same response keys with words representing the self (e.g., “me”) versus other people (e.g., “them”).

Limitations of the self-concept IAT

Quantitative reviews of the IAT (Hofmann et al., 2005) and large-scale studies (Bar-Anan & Nosek, 2014) indicate that explicit-implicit correlations are weaker when measuring attitudes toward the self ($r = .21$), compared to political ($r = .36$) or race attitudes ($r = .58$, in Bar-Anan and Nosek study). As De Cuyper et al. (2017) suggested, the relatively poor performance of the self-concept IAT—and perhaps its general absence from mainstream personality research—may be explained by the notable limitations of this task. These include problematic response stimuli, the difficulties in interpreting the task’s output, and the relatively limited range of personality constructs it enables to assess.

Sub-optimal stimuli

The low explicit-implicit correlations typically observed for the self-concept IAT may be at least partially due to the nature and quality of the response stimuli in this task (see De Cuyper et al., 2017). In most cases the direct and the IAT-based indirect assessment procedures of the self are based on very different sets of indicators of the target construct—typically, questionnaire items versus single words, respectively. Also, unlike standard self-reports, the psychometric properties of the stimuli used in the self-concept IAT (i.e., the indicators of the measured construct) are rarely verified (Costantini et al., 2015). Indeed, the literature suggests that explicit-implicit correspondence increases when the two types of measurement are based on the same stimuli (De Cuyper et al., 2017; Hofmann et al., 2005).

Ambiguous outcome

The assessment of the self-concept IAT is not easily amenable to clear interpretation for several reasons. For example, the self is typically compared in this task to others, and it may not be obvious who these “others” are. Also, it is unclear whether the obtained score in the IAT reflects self-positive or others-negative associations. Most importantly, the score in the self-concept IAT merely reflects the extent to which the self is associated with the measured attribute (e.g., extraversion), but such association may result from several different relations between these concepts (de Houwer, 2014). It is commonly assumed that the belief “I am an extravert person” underlies this association, but other types of idiosyncratic relationships (e.g., being preoccupied with whether or not one is an extravert person) are also possible, some of which may even reflect the inverse direction of the construct (e.g., “I wish I were an extravert person”). In explicit self-reports, such ambiguities do not exist,

because the nature of the relationship between concepts is unequivocally inferred from the content of the propositions that comprise the scale.

Scope of assessment

The assessment process of self-report scales, which is commonly based on responses to short statements or propositions, enables a great deal of flexibility and specificity. Therefore, despite their limitations, standard explicit questionnaires allow the assessment of numerous distinct constructs related to the self. In contrast, implicit assessment, which is typically based on responses to single words, has been far narrower in scope (De Cuyper et al., 2017; Wilson, 2009). Indeed, the adequate measurement of most psychological constructs requires the use of stimuli (i.e., indicators of the target construct) that are far more semantically complex.

To summarize, the self-concept IAT is a reliable and straightforward implicit assessment method that does not rely on reflective processes. As such, it may provide information that cannot be obtained by self-report personality questionnaires. However, it also has several notable weaknesses. The assessment procedure of this task is often based on sub-optimal stimuli, the information it provides cannot be easily and unambiguously interpreted, and it enables the assessment of a much smaller scope of personality constructs than ordinary self-reports. The qIAT, which integrates key aspects of the assessment procedure of standard questionnaires into the indirect assessment method of the IAT, addresses these limitations.

The questionnaire-based implicit association test

The qIAT (Yovel & Friedman, 2013) enables an indirect assessment that is based on the items of standard self-report scales. Following earlier versions of the IAT (Greenwald et al., 1998; Sartori et al., 2008), the qIAT measures the speed of classification of original questionnaire items (e.g., “I feel comfortable around people”, “I don’t like to draw attention to myself”) to relevant categories (e.g., *extravert person vs. introvert person*, respectively), when they are classified interchangeably with *true* (e.g., “I am looking at a computer screen”) versus *false* (e.g., “I am currently playing an electric guitar”) statements related to the self (cf. Sartori et al., 2008). This assessment procedure is indirect because it is not based on people’s explicit endorsements of the questionnaire’s items (e.g., using a Likert scale). Instead, as in other variants of the IAT, the measurement in the qIAT is based on facilitation and inhibition processes resulting from the paired combinations in the different blocks of the task (e.g., *extravert person* and *true* versus *introvert person* and *false*).

The indirect assessment method of the qIAT resembles the assessment procedure of standard self-reports in several ways, and it thus addresses the limitations of the self-concept IAT. First, the qIAT and the parallel explicit assessment are based on the same set of stimuli—the items of the explicit questionnaire. Consequently, any observed explicit-implicit dissociations

Table 1. Questionnaire-based Implicit Association Tests for extraversion and conscientiousness: Task sequence and stimuli.

Sequence	N of trials	Task	Response key assignment	
			Left key	Right key
1	40	Personality categories discrimination	Extravert (Conscientious)	Introvert (Easy-going)
2	20	Self-related logical categories discrimination	True	False
3	20	Initial combined task: Practice block	Extravert (Conscientious), true	Introvert (Easy-going), false
4	40	Initial combined task: Pest block	Extravert (Conscientious), true	Introvert (Easy-going), false
5	40	Reversed personality categories discrimination	Introvert (Easy-going)	Extravert (Conscientious)
6	20	Reversed combined task: Practice block	Introvert (Easy-going), true	Extravert (Conscientious), false
7	40	Reversed combined task: Pest block	Introvert (Easy-going), true	Extravert (Conscientious), false

Note. Words in parentheses refer to the task sequence within the conscientiousness qIAT.

must be attributed to other factors (Nosek et al., 2011; Schnabel & Asendorpf, 2010). Also, to the extent that the target personality scale is well-validated, the indirect assessment of the qIAT is based on psychometrically sound items, rather than on stimuli constructed ad-hoc for a specific study (see Costantini et al., 2015). Most importantly, an indirect assessment process that is based on semantically complex stimuli (i.e., whole sentences rather than single words) allows the assessment of a much wider scope of self-related psychological constructs, currently measured only by self-reports.

The pairing combinations in the qIAT also parallel the assessment procedure of standard self-reports. In most personality questionnaires, respondents need to rate the extent to which each of several short statements is true for themselves. Similarly, the personality items in the qIAT are presented interchangeably with true versus false self-related statements. This procedure enables a much clearer interpretation of the results, compared to the broad and vague association of the assessed construct (e.g., extraversion) with the self (versus “others”) in the self-concept IAT. That is, in the qIAT, the obtained score unambiguously reflects the extent to which respondents associate the personality items (e.g., descriptions of an extraverted person) with truth about the self.

The present research

In the first investigation that tested this method, a qIAT that measured extraversion showed good split-half reliability ($r = 0.87$; Yovel & Friedman, 2013), and the qIAT extraversion score correlated with the parallel explicit extraversion scale ($r = .36$), but not with the other four Big Five scales (r 's between $-.05$ and $.09$). Similarly, Currie et al. (2017) found that a qIAT measuring shame aversion correlated with the parallel explicit measure of this trait ($r = .28$), and also with measures of relevant types of personality psychopathology (e.g., avoidant personality disorder, $r = .21$).

The qIAT was designed to facilitate the integration of implicit assessment into the field of personality. Therefore, in the present studies we tested whether the implicit assessment the qIAT yields is stable and valid, and whether it can predict behavior.¹ In Study 1, measuring extraversion, we primarily focused on the stability over time of this implicit assessment procedure. In Studies 2 and 3, we further tested

the reliability of this task and examined whether the qIAT that measured conscientiousness would predict actual behavior.

General methods

Participants

In all studies, participants were recruited via crowdsourcing platforms. To encourage valid responses, participants were allowed to complete the studies regardless of their reported native language or country of birth, but only those whose native language was English were included in the final analyses. Additionally, the following exclusion criteria were based on performance in the qIAT (see below): high error rate (20% or above) in the critical blocks of the task, high rate of extremely fast responses (at least 10% with $RT < 300$ msec), and taking a break while completing the task (based on $RT > 30$ seconds in any trial).

The questionnaire-based implicit association test

The qIAT tasks in both studies were administered using the Flash program (Macromedia Flash 10.0 Professional, 2005) for reaction-time measurement (Reimers & Stewart, 2007). Following Yovel and Friedman (2013) and Sartori et al. (2008), the task included seven blocks (see Table 1). On each trial, a sentence was presented at the center of the computer screen, and participants needed to press one of two designated response keys in order to classify it as quickly and accurately as possible (the stimuli used the tasks appear in Table 2). In Block 1 (40 trials), participants were introduced to the classification of the personality categories (e.g., “*Extravert person*” vs. “*Introvert person*”). In Block 2 (20 trials), they were introduced to the classification of the self-related logical categories, labeled *True* (e.g., “*I am in front of the computer*”) versus *False* (e.g., “*I am sunbathing at the beach*”). In the critical Blocks 3 (20 trials) and 4 (40 trials), participants performed a double-categorization task, where these two tasks appeared interchangeably and the two types of categories shared the same response keys (e.g., *Extravert person* and *True* versus *Introvert person* and *False*). In Block 5 (40 trials), participants practiced the reversed classification of the personality category, and in the critical Blocks 6 and 7, they performed the second double-categorization task, this time using the reversed trait classification (e.g., *Introvert person* and *True* versus *Extravert person* and *False*).

¹Measures, data, data processing codes and analysis syntax of the present studies are openly available as [supplementary materials](https://osf.io/z52mw/) at osf.io/z52mw/.

Table 2. Categories and stimuli used in the implicit tasks.

Logical self-related categories and stimuli (all tasks)		
	True	False
	I am doing a psychology experiment	I am playing football on the grass
	I am in front of the computer	I am sunbathing at the beach
	I am participating in an experiment on the internet	I am currently playing an electric guitar
	I am looking at a computer screen	I am climbing a steep mountain
	I am putting my fingers on the keyboard	I am buying groceries in the local grocery store
Personality categories and stimuli		
Study 1 (both sessions)	Extravert person I am the life of the party I feel comfortable around people I start conversations I talk to a lot of different people at parties I don't mind being the center of attention	Introvert person I don't talk a lot I keep in the background I have little to say I don't like to draw attention to myself I am quiet around strangers
Study 2 and Study 3 First Session	Conscientious I am exacting in my work I complete tasks successfully I carry out my plans I follow through with my plans I pay attention to details	Easy Going I need a push to get started I don't put my mind on the task at hand I find it difficult to get down to work I don't see things through I mess things up
Second Session	Conscientious I get chores done right away I do things according to a plan I make plans and stick to them I finish what I start I am always prepared	Easy Going I make a mess of things I leave things unfinished I shirk my duties I waste my time I do just enough work to get by

In all blocks, the category labels remained on the computer screen, and an error signal (a small red “X”) appeared at the center of the screen following incorrect classifications, until the correct response was made. The order of the double-categorization blocks was counterbalanced across participants.

The qIAT *D* score was calculated following Greenwald and colleagues' (Greenwald et al., 2003) improved scoring algorithm. Larger positive *D*s represented a stronger association between the measured trait and the self-related *True* category. Thus, like the total scores of the parallel self-report questionnaires, individuals higher on the assessed trait (e.g., extraversion) were expected to have larger *D*s.

Study 1

In this study we tested the basic psychometric properties of the qIAT, including stability over time. Participants completed a qIAT that measured extraversion and an explicit questionnaire that measured all Big Five domains (Goldberg et al., 2006) twice, in two separate sessions completed two weeks apart. We examined the internal consistency of the qIAT in each session, as well as the test-retest reliability of this task. To test the convergent and discriminant validity of the qIAT, we examined its correlations with the Big Five subscales in each session. We expected that the implicit extraversion assessment of the qIAT would show stability over time, reflected by the correlation between the two implicit scores obtained in the two sessions. Also, based on previous findings (Yovel & Friedman, 2013), we predicted that the qIAT would correlate with the parallel explicit extraversion subscale, and that this correlation would be larger than the correlations with the other four subscales that measure other personality dimensions.

Method

Participants

In Studies 1 and 2, participants were recruited via Amazon's Mechanical Turk platform. They were all American residents (based on IP address) who had completed more than 100 Human Intelligence Tasks (HITs) with an approval rate of at least 95%. We had no basis for estimating the expected effect sizes associated with an online version of the qIAT. Therefore, we used a conservative a priori power analysis, which was based on the mean population explicit-implicit correlation reported by Hoffman et al. (2005) for previous versions of the self-related IAT ($\rho = 0.128$). Based on this analysis, we needed a large sample of at least 477 participants to detect Pearson correlations of this magnitude between the parallel implicit and explicit measures with alpha levels set at 0.05 and a power of 0.80. Of the 528 native English speakers recruited for the study, 47 participants were excluded based on their performance in the qIAT task in the first session of the study, following the criteria detailed in the General Methods. Analyses were based on the remaining 481 participants (238 females; mean age = 36.75, $SD = 12.60$). Of these, 365 participants (75.88%) returned to take part in the second session two weeks later, and six were excluded based on their performance in the second qIAT task. Thus, analyses in the second session were based on 359 participants (173 females; mean age = 37.76, $SD = 12.89$).

Materials

Explicit measures

The standard self-report assessment included the 50-item International Personality Item Pool (IPIP; Goldberg et al.,

Table 3. Descriptive statistics, internal consistencies and intercorrelations of measures in Study 1.

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. qIAT T1	0.00	0.57	.89											
2. Extraversion T1	27.49	9.62	.50***	.93										
3. Emotional stability T1	34.5	9.13	.28***	.38***	.92									
4. Agreeableness T1	39.36	6.85	.20***	.34***	.25***	.89								
5. Conscientiousness T1	38.48	7.04	.13**	.16**	.39***	.27**	0.88							
6. Intellect T1	38.77	6.00	.09	.26***	.19***	.24**	.15**	0.83						
7. qIAT T2	-0.06	0.61	.69***	.58***	.29***	.17**	.10	.047	0.92					
8. Extraversion T2	26.86	9.78	.53***	.93***	.38***	.34**	.18**	.22***	.58***	0.94				
9. Emotional stability T2	34.87	9.23	.28***	.37***	.92***	.26***	.31***	.18**	.27***	.40***	0.93			
10. Agreeableness T2	39.35	6.89	.17**	.29***	.27***	.89***	.25***	.22***	.16**	.32***	.25***	0.90		
11. Conscientiousness T2	38.96	7.05	.15**	.16**	.34***	.26***	.92***	.12*	.09	.20***	.33***	.30***	0.90	
12. Intellect T2	38.81	6.46	.06	.19***	.16**	.25***	.10	.88***	.02	.19***	.16**	.28***	.12*	0.86

Note. Internal consistencies are presented in the diagonal. T1 = first session; T2 = second session. * $p < .05$; ** $p < .01$.

2006) representation of the Goldberg markers (Goldberg, 1992) for the Big Five factor model of personality (Agreeableness, Conscientiousness, Extraversion, Emotional Stability, Intellect; all items are presented in the online [supplementary materials](#)). Each domain was measured by a 10-item subscale, and items were rated using a 1–5 Likert scale. Internal consistencies (Cronbach's alphas) in the current study ranged between 0.83 for Intellect and 0.93 for Extraversion.

Implicit measure

The qIAT was designed to assess extraversion, using the ten items of the IPIP Extraversion subscale (e.g., “I talk to a lot of different people at parties”, “I don't like to draw attention to myself”; see Table 2).

Procedure

Participants were informed that the study would include two sessions, administered two weeks apart. In each session, they first completed the explicit standard self-report questionnaire, which was immediately followed by the qIAT. The IAT is generally robust against explicit-implicit order effects (Hofmann et al., 2005). However, in contrast to the conventional IAT, here the target constructs (i.e., extraversion) were measured explicitly and implicitly using the exact same set of relatively complex propositional stimuli (i.e., the questionnaire items). Therefore, to prevent a likely carryover effect, the administration of the self-report scale, in which each item was presented only once, always preceded the qIAT task, in which participants were exposed to each of these items many times (Yovel & Friedman, 2013). Unpublished data collected in our lab showed that administering the implicit task first may affect the explicit assessment and consequently distort the explicit-implicit correlations. Participants received \$1 for their participation in the first session. Two weeks later, they received an email reminding them to return and complete the second part of the study, for which they received \$2. In the second session, they completed the same version of the qIAT (i.e., the same order of critical blocks) they had completed in the first session.

Results

We first examined the split-half and test-retest reliabilities of the qIAT. The split-half reliabilities, based on the correlations between D scores calculated separately for odd and even trials in the task (Spearman-Brown corrected), were $r = .89$ for the qIAT in the first session, and $r = .92$ for the qIAT completed in the second session. As to test-retest reliability, the correlation between the qIAT tasks performed two weeks apart was $r_{tt} = .69$. Supporting the convergent validity of the qIAT, the correlation between the implicit and explicit extraversion measurements was $r = .50$, $p < .001$, 95% CI [.43, .56] in the first session, and $r = .58$, $p < .001$, 95% CI [.51, .64] in the second session (see Table 3). Supporting the discriminant validity of the qIAT, the correlations between the qIAT extraversion score and the other four explicit subscales were all significantly smaller (all Fisher's Z s > 5.72 , all p s $< .001$), and they ranged between $r = .09$ (with Intellect), $p = .056$, 95% CI [.00, .18], and $r = .28$ (with Emotional Stability), $p < .001$, 95% CI [.20, .36], in the first session; and between $r = .02$ (with Intellect), $p = .648$, 95% CI [-.08, .12], and $r = .27$ (with Emotional Stability), $p < .001$, 95% CI [.17, .36], in the second session.

We next examined more closely the convergent and discriminant validity of the qIAT score at the item level. These analyses showed that in both sessions of this study, the ten self-reported extraversion items had the strongest correlations with the implicit qIAT extraversion score among all 50 IPIP items, ranging between $r = .30$, 95% CI [.22, .38], and $r = .50$, 95% CI [.43, .56], in the first session; and between $r = .37$, 95% CI [.28, .46], and $r = .54$, 95% CI [.46, .61], in the second session (see Table 4). To compare, the correlations of all the remaining 40 IPIP items (i.e., the items of other four Big Five scales) with the qIAT extraversion scores ranged between $r = -.06$, 95% CI [-.15, .03], and $r = .26$, 95% CI [.17, .34], with a median of 0.12 in the first session; and between $r = -.01$, 95% CI [-.11, .09], and $r = .31$, 95% CI [.21, .40], with a median of .08 in the second session.

Discussion

The split-half reliability of the qIAT extraversion score obtained here was excellent, and the two qIAT extraversion scores obtained two weeks apart strongly correlated with each other. Thus, the present findings suggest that in terms

Table 4. Bivariate correlations of the items of the self-report extraversion scale with the total explicit extraversion score (corrected item-total correlations) and with the implicit qIAT extraversion score in each of the sessions in Study 1.

Item	First Session		Second Session	
	Self-Report	qIAT	Self-Report	qIAT
Keep in the background	.82	.42	.82	.47
Talk to a lot of different people at parties	.80	.50	.81	.53
Am quiet around strangers	.80	.43	.81	.54
Start conversations	.77	.43	.80	.44
Am the life of the party	.75	.39	.77	.51
Don't talk a lot	.74	.38	.78	.49
Don't mind being the center of attention	.70	.38	.72	.50
Have little to say	.69	.31	.72	.40
Feel comfortable around people	.69	.44	.76	.46
Don't like to draw attention to myself	.65	.30	.63	.37

Note. qIAT = questionnaire-based Implicit Association Test. For all correlations, $ps < .001$.

of both internal consistency and temporal stability, the qIAT is at least as reliable as earlier versions of the IAT (internal consistency and temporal stability estimates of $r = .74$, $r_{tt} = .41$, respectively, De Cuyper et al., 2017; $r = .79$; $r_{tt} = .63$, Gawronski et al., 2017; $r = .79$; $r_{tt} = .51$, Hofmann et al., 2005). The present findings also support the convergent and discriminant validity of the assessment provided by the qIAT. The correlations observed here between the implicit qIAT and the parallel self-report extraversion scale were strong, and in both sessions this correlation was significantly higher than the correlations between the qIAT score and all the other Big-5 scales (cf. De Cuyper et al., 2017). In addition, more detailed analyses showed that the implicit extraversion score correlated more strongly with items that measured extraversion explicitly, compared to all items of the other four Big Five subscales. In sum, the present findings support the internal consistency, test-retest reliability, convergent and discriminant validity of the extraversion qIAT. In Study 2 we focused on a different personality trait, and tested additional aspects of the reliability and validity of the qIAT.

Study 2

In this study we expanded the examination of the qIAT in several ways. First, in the initial validation study of this task (Yovel & Friedman, 2013) and in Study 1, the qIAT measured the same construct, extraversion. Here we focused on conscientiousness, a Big Five dimension that has been associated with more social desirability and self-report biases (e.g., a correlation of $r = .29$ with Impression Management subscale; Paulhus, 2002). Perhaps relatedly, the convergent validity of IAT-based measures of conscientiousness with explicit measures of this construct have been less consistent and smaller in magnitude compared to extraversion (e.g., $r = .15$ for conscientiousness, $r = .24$ for extraversion; De Cuyper et al., 2017). Second, the indirect personality assessment of the qIAT has been previously based on the items of the parallel self-report scale. This is a major strength of this task, but it could be argued that the observed explicit-implicit relationships the qIAT demonstrated are merely due to response biases associated with the fact that both measures use the exact same set of stimuli (e.g., carry-over effects). To examine this possibility, we tested the convergent validity of the qIAT

with the explicit self-report assessment, but this time the two types of measurement were based on parallel forms of the target personality construct (i.e., on two different sets of items). Finally, we tested whether the implicit conscientiousness assessment of the qIAT can predict actual behavior.

As in Study 1, the present study included two parts. Here, however, participants were paid for their participation in the second session in advance, immediately following the first session. We expected that under these conditions, more conscientious individuals would be more likely to return to complete the second session of the study. Thus, we hypothesized that conscientiousness assessments would predict who would return to complete the second session.

Method

Participants

The results in Study 1 strongly indicated that that our effect size estimation for the online version of the qIAT in that study was too low. Therefore, here we used a higher estimation ($r = .30$), which was still conservative considering the explicit-implicit correlations obtained in Study 1. An a priori power analysis showed that a sample of at least 85 would be needed to detect such moderate correlations with a power of .80. However, we employed a larger sample, because here we also attempted to predict participants' behavior using the qIAT, and we had no basis for estimating the effects size associated with this prediction. Of the 164 native English speakers who participated in this study, 20 participants were excluded from the analyses due to their performance in the qIAT tasks, following the criteria detailed in the General Methods. Analyses were based on the remaining 144 participants (85 females; mean age = 36.98, $SD = 11.66$).

Materials

The explicit and implicit measures in the present study were created using items from the five 20-item IPIP (Goldberg et al., 2006) representations of the Big Five factors in Costa and McCrae (1992) NEO Personality Domains. For each factor we created two parallel forms (Subset A and Subset B), each containing 10 IPIP items.

Table 5a. Hierarchical logistic regression analysis predicting participants' return for the second session in Study 2.

Step and variable	<i>b</i>	<i>SE</i>	Wald statistic	Odds ratio (95% confidence interval)	χ^2	<i>R</i> ²
Step 1					1.84 (<i>p</i> =.175)	.02
Self-report	.032 (<i>p</i> =.181)	.02	1.79	1.03 (.99-1.09)		
Step 2					3.18 (<i>p</i> =.075)	.05
Self-report	.02 (<i>p</i> =.389)	.02	.74	1.02 (.97-1.07)		
qIAT score	.63 (<i>p</i> =.080)	.36	3.07	1.87 (.93-3.78)		

Note. qIAT = Questionnaire-based Implicit Association Test.

Explicit measures

The standard self-report assessment included the five Subset A IPIP subscales (items are presented in the supplemental materials online). Participants rated each of the 50 items on a 1–5 Likert scale. Cronbach's alphas for these subscales ranged between .81 (for Openness to experience) and .93 (for Neuroticism), and it was .92 for Conscientiousness, the target trait in the present study.

Implicit measures

Two parallel forms of the qIAT designed to measure conscientiousness were created (e.g., “I get chores done right away”, “I do just enough work to get by”; see Table 2), based on Subset A and Subset B IPIP Conscientiousness subscales.

Procedure

The study included two sessions. In the first session, participants completed the five Subset A self-report Big Five subscales, including Conscientiousness. They then completed a qIAT task based on the items of the Subset B Conscientiousness subscale. Afterward, participants were offered the option to participate in the second session of the study two weeks later, for an additional amount of \$1. Those who declined this offer received \$1 for their participation in the first session. Participants who agreed to complete the second session received immediately following the first session \$1 *in advance* for their future participation in the second session (i.e., a total of \$2), and a reminder email was sent to them one week later. In the second session of the study, participants completed a qIAT task in which the personality items were the ten Subset A conscientiousness items.

Results

In the first session of the study, the split-half reliability of the qIAT, based on the correlation between the *D* scores calculated separately for the odd and even trials (Spearman-Brown corrected), was $r = .85$. The correlation between the implicit qIAT score and the explicit self-report measure of conscientiousness (using different sets of items) was significant, $r = .30$, $p < .001$, 95% CI [.14, .44]. As to discriminant validity, this correlation was significantly larger than the correlations of the qIAT with the explicit measures of Openness to Experience, $r = -.02$, $p = .820$, 95% CI [-.18, .14], Fisher's $Z = 3.44$, $p = .001$ and Extraversion, $r = .14$, $p = .094$, 95% CI [-.02, .30], Fisher's $Z = 2.00$, $p = .045$, but not than the correlations with Neuroticism, $r = -.25$, $p =$

.003, 95% CI [-.40, -.09], Fisher's $Z = 0.64$, $p = .521$ and Agreeableness, $r = .20$, $p = .019$, 95% CI [.03, .35], Fisher's $Z = 1.32$, $p = .188$.

One hundred and eighteen participants agreed to return to complete the second session and got paid for it in advance, but only 60 participants (50.85%) eventually returned and completed both parts of the study. For these participants, the split-half reliability of the second qIAT was $r = .81$. The correlation between the two qIAT scores administered two weeks apart *and* using parallel forms was $r_{tt} = .41$, $p = .001$, 95% CI [.17, .60]. The correlation between self-reported conscientiousness completed in the first part and the qIAT score completed in the second part (both based on the same set of items) was $r = .34$, $p = .009$, 95% CI [.09, .55].

We next examined whether the explicit and implicit conscientiousness measures completed in the first session predicted who would return for the second session. A between-group *t*-test revealed that the difference between those who returned ($M = 39.73$, $SD = 7.13$) and those who did not return ($M = 37.78$, $SD = 8.59$) on the self-reported scale, which was in the expected direction, was not significant, $t(116) = 1.35$, $p = .180$, $d = .24$, 95% CI [-.92, 4.83]. However, another *t*-test showed that the difference between these two groups on the implicit qIAT score ($M = .94$, $SD = .54$ vs. $M = .73$, $SD = .56$, respectively) was significant, $t(116) = 2.07$, $p = .041$, $d = .38$, 95% CI [.01, .41]. Taken together, these results indicate that the implicit qIAT conscientiousness measure, but not the parallel self-reported scale, predicted who would return to complete the study. As Table 5a shows, a hierarchical logistic regression analysis in which self-reported conscientiousness was entered in the first step, the qIAT score was entered in the second step, and participants' behavior (returned vs. did not return) was the dependent measure revealed that the incremental prediction of the implicit over the explicit measure was marginally significant ($p = .075$).

Discussion

The target construct in the present study was conscientiousness, a Big Five personality dimension that has been associated with more report biases compared to extraversion (D L Paulhus, 2002), the focal trait in Study 1. Supporting the reliability of the qIAT conscientiousness measure, it showed good internal consistency, which was comparable to previous studies that focused on this construct (e.g., $r = .82$, Grumm & von Collani, 2007). The two qIAT scores completed two weeks apart correlated significantly with each

other, but due to the design of the present study, it is likely that this is an underestimation of the test-retest reliability of this score. First, it confounds stability over time with parallel forms reliability, as the two qIAT tasks included two different sets of items. Perhaps more importantly, this correlation was based on a range-restricted subset of the sample, because those who returned to complete the second session of the study had higher conscientiousness qIAT scores than those who did not return.

The qIAT score correlated significantly with the parallel conscientiousness self-report scale completed in the first session of the study, even though these two measures were based on different sets of items. Thus, the present results provide additional support for the convergent validity of the assessment of the qIAT, and suggest that observed explicit-implicit relationships between the qIAT and parallel self-report scales cannot be attributed solely to time proximity or item similarity between the two modes of assessment. Finally, the qIAT conscientiousness score predicted whether participants returned to complete the second session after they had agreed and had been paid to do so. The parallel self-report conscientiousness scale did not predict this behavior, but the incremental prediction of the qIAT over the explicit scale was only marginally significant. In the next study we again tested the predictive utility of the qIAT conscientiousness measure using similar methods, attempting to replicate these findings.

Study 3

In the present pre-registered study (<https://aspredicted.org/2qm3t.pdf>), we focused on the predictive utility of the qIAT, using the same procedures used in Study 2. Attempting to replicate the findings observed in Study 2, we predicted that the implicit qIAT conscientiousness measure would predict who would return to complete the second session of the study, after agreeing to do so and getting paid for it in advance.

Method

Participants

The study was opened to 200 participants, in accordance with the recruitment strategy set forth in the pre-registration. Participants were recruited via Prolific Academic platform, and they all had previously submitted at least 50 tasks with an approval rate of 95% or above. Six participants who were not native English speakers (based on reported birth country) were not included in the analyses. An additional 27 participants were excluded from the analyses based on their performance in the qIAT tasks, following the criteria detailed in the General Methods. Analyses were based on the remaining 167 participants (97 females; mean age = 33.78, $SD = 10.08$).

Materials

The same explicit and implicit measures that were used in Study 2 were also used in the present study.

Procedure

The same procedures that were used in Study 2 were also used here. The two studies only differed in compensation, following Prolific's payment standards. In this study, participants who declined the offer to participate in the second session received £1.25 for their participation in the first session, while those who agreed to complete the second session received an additional £1.25, totaling £2.50.

Results

The split-half reliability of the qIAT in the first session of the study, based on the correlation between the D scores of the odd and even trials (Spearman-Brown corrected), was $r = .77$. Of the 151 participants who agreed at the end of the first session to complete the second session of the study (and got paid for it in advance), only 63 (41.72%) did so. The split half reliability of the qIAT in the second session was $r = .78$. The correlation between the two qIAT scores administered two weeks apart *and* using parallel forms was $r_{tt} = .28$, $p = .029$, 95% CI [.04, .49]. The correlation between the implicit qIAT score and the explicit self-report measure of conscientiousness (using different sets of items) was $r = .19$, $p < .015$, 95% CI [.04, .33]. The correlation between self-reported conscientiousness completed in the first part and the qIAT score completed in the second part (both based on the same set of items) was $r = .25$, $p = .052$, 95% CI [-.002, .47].

We next examined whether the self-report and the qIAT measures of conscientiousness completed in the first session predicted who would return to complete the second session. A between-group *t-test* revealed that the difference between those who returned ($M = 33.89$, $SD = 7.27$) and those who did not ($M = 33.26$, $SD = 7.56$) on the explicit scale was small and not significant, $t(149) = .51$, $p = .923$, $d = .09$, 95% CI [-1.80, 3.05]. The difference between these two groups on the qIAT score ($M = .84$, $SD = .47$ vs. $M = .71$, $SD = .46$, respectively) was almost significant, $t(149) = 1.73$, $p = .087$, $d = .28$, 95% CI [-.02, .28]. Thus, the effect sizes in the present study were smaller compared to the previous study, but the general pattern remained. A hierarchical logistic regression analysis (as in Study 2; see Table 5b) indicated that this was also the case for the incremental prediction of the qIAT over the self-report conscientiousness scale ($p = .095$).

Follow-up analyses

The methods used in the present study and in Study 2 were similar. We therefore aggregated the samples of these studies into a single, larger sample, which allowed the examination of the predictive validity of the two measures of

Table 5b. Hierarchical logistic regression analysis predicting participants' return for the second session in Study 3.

Step and variable	<i>b</i>	<i>SE</i>	Wald statistic	Odds ratio (95% confidence interval)	χ^2	<i>R</i> ²
Step 1					.264 (<i>p</i> =.607)	.002
Self-report	.01 (<i>p</i> =.607)	.02	.26	1.01 (.97-1.06)		
Step 2					2.79 (<i>p</i> =.095)	.03
Self-report	.01 (<i>p</i> =.813)	.02	.06	1.01 (.96-1.05)		
qIAT score	.62 (<i>p</i> =.101)	.38	2.69	1.86 (.89-3.88)		

Note. qIAT = Questionnaire-based Implicit Association Test.

Table 5c. Hierarchical logistic regression analysis predicting participants' return for the second session across studies.

Step and variable	<i>b</i>	<i>SE</i>	Wald statistic	Odds ratio (95% confidence interval)	χ^2	<i>R</i> ²
Step 1					2.23 (<i>p</i> =.136)	.01
Study	-.37 (<i>p</i> =.137)	.25	2.22	.69 (.43-1.12)		
Step 2					1.70 (<i>p</i> =.192)	.02
Study	-.26 (<i>p</i> =.322)	.26	.98	.77 (.46-1.29)		
Self-report	.02 (<i>p</i> =.194)	.02	1.69	1.02 (.99-1.06)		
Step 3					6.12 (<i>p</i> =.013)	.05
Study	-.27 (<i>p</i> =.316)	.27	1.01	.77 (.46-1.29)		
Self-report	.01 (<i>p</i> =.446)	.02	.58	1.01 (.98-1.05)		
qIAT score	.63 (<i>p</i> =.015)	.26	5.90	1.88 (1.13-3.12)		

Note. qIAT = Questionnaire-based Implicit Association Test.

conscientiousness with more statistical power. Across both studies, 269 participants agreed to complete the second session and were paid for it in advance, while 123 participants (45.73%) returned.

A between-group *t*-test showed that the difference between those who returned ($M = 36.74$, $SD = 7.75$) and those who did not return ($M = 35.06$, $SD = 8.26$) across the two studies on the explicit conscientiousness scale was almost significant, $t(267) = 1.72$, $p = .088$, $d = .21$, 95% CI [-.25, 3.62]. The difference between these two groups on the implicit qIAT score ($M = .89$, $SD = .50$ vs. $M = .72$, $SD = .50$, respectively) was significant, $t(267) = 2.79$, $p = .006$, $d = .34$, 95% CI [.05, .29]. To examine the incremental prediction of the qIAT over the self-report conscientiousness scale, we performed another hierarchical logistic regression analysis. Because we used an aggregated sample, we controlled for the study in which participants were tested by entering this dichotomous variable (Study 2 versus Study 3) in the first step. As Table 5c shows, across both studies, the qIAT conscientiousness score significantly predicted the criterion behavior over and above the parallel self-report scale ($p = .013$).

Discussion

The findings of earlier studies that examined the criterion validity of indirect measures of conscientiousness have been mixed (Costantini et al., 2015). In one particularly extensive investigation of the validity of all Big Five traits, implicit conscientiousness (measured by the self-concept IAT) did not predict an aggregate of eight relevant intentional and spontaneous behavioral indicators (e.g., number of errors in a short story, lateness in sending back a questionnaire, $r = -.08$; Back et al., 2009). In several other studies, however, self-IATs that measured conscientiousness predicted observed, particularly spontaneous behaviors, such as

mistakes in a concentration test ($r = .36$; Steffens & Schulze König, 2006) or the number of exams passed in an academic year ($r = .36$; Vianello et al., 2010). Here, we focused in Study 2 and in Study 3 on a deliberate, intentional criterion behavior. Specifically, participants who chose to take part in both sessions of the study were paid in advance for their future participation in the second session, immediately following the first session. They received a reminder email a week later, but in both studies only about half of them eventually returned to complete the second session. Results showed that those who returned had higher qIAT conscientiousness scores than those who did not return, while the difference between those groups on the parallel self-report questionnaire was not significant in either study. Thus, only the implicit assessment predicted who would return to complete the second part of the study after committing, being paid, and being reminded to do so. This effect was only marginally significant in the present replication study, but in general a similar pattern of results was observed in both studies. Moreover, across both these studies, the qIAT predicted this criterion behavior incrementally over and above the parallel self-report conscientiousness scale.

General discussion

Self-report personality questionnaires can only provide information that people are able to access via introspection and are willing to share with others (Schwarz, 1999). Despite this limitation, alternative indirect assessment methods are rarely employed in personality research (De Cuyper et al., 2017). Compared to explicit personality-assessment instruments, implicit tasks are usually based on suboptimal stimuli (Costantini et al., 2015) and yield findings (e.g., associations of the self with certain personality constructs) that are difficult to interpret (Remue et al., 2013). Compared to other types of attitudes (e.g., racial, political), implicit

measurement of attitudes about the self has generally shown inferior convergent validity with explicit measures (e.g., Bar-Anan & Nosek, 2014; Hofmann et al., 2005). The qIAT was designed to address these challenges, and the current research suggests that it does.

The indirect assessment procedure of the qIAT, which is based on the original items of standard self-report questionnaires, measures the extent to which respondents associate the trait measured by these items with truth about themselves. In the studies presented here, the qIAT showed good levels of internal consistency, stability over time, and explicit-implicit convergent validity. These psychometric properties of the qIAT were at least comparable to those reported for other implicit measures of attitudes toward the self, including the self-concept IAT (e.g., Bar-Anan & Nosek, 2014; De Cuyper et al., 2017). In addition, the results of Study 2 provided initial support for the utility of the qIAT in predicting behavior. In this study, participants who agreed to complete both parts of a two-session online study were paid in advance immediately following the first session, and the qIAT that measured conscientiousness (but not the explicit questionnaire) predicted who would later return to complete the second session two weeks later. Study 3's procedure directly replicated that of Study 2 and showed a similar pattern of results. The prediction of the qIAT, however, was only marginally significant. Across both these studies, the qIAT predicted this criterion behavior significantly above and beyond the parallel self-report conscientiousness scale.

Internal consistency

Notwithstanding its limitations, the self-concept IAT is the most widely used paradigm for the indirect measurement of personality (De Cuyper et al., 2017; Schnabel & Asendorpf, 2010). The instructions for the IAT are straightforward, and its flexible structure enables the assessment of different traits and attitudes toward the self (see Back et al., 2009). Most importantly, compared to other implicit tasks, the IAT typically shows good levels of internal consistency. Low internal consistency (i.e., high proportion of measurement error) entails poor validity, because it weakens the correlations of the implicit measure with other variables (Gawronski & De Houwer, 2014; LeBel & Paunonen, 2011). The qIAT differs from the self-concept IAT in that respondents classify questionnaire items and true versus false self-related statements rather than single words (trait-related, self vs. others). Otherwise, the two tasks are essentially similar. They are both derived from the same logic, have the same structure, and use the same scoring methodology. Accordingly, the qIAT tasks in the present studies all showed good or excellent levels of internal consistency.

Explicit-Implicit similarity: Stimuli

Implicit assessment in the present studies was based on the items of standard personality questionnaires. This central feature of the qIAT addresses a notable limitation in implicit

personality research (De Cuyper et al., 2017), where the indirect measurement has been typically based on ad-hoc stimuli with no known psychometric properties (for an exception see Costantini et al., 2015). Also, the literature suggests that explicit-implicit relationships are higher when both types of assessment methods use the same stimuli (e.g., Hofmann et al., 2005). Accordingly, the explicit-implicit correlations between the qIAT score and the parallel explicit personality questionnaire were relatively high, compared to other variants of the IAT that measured similar constructs (e.g., De Cuyper et al., 2017). This was the case even in Study 2, where the explicit and implicit measures that used the exact same items were administered two weeks apart, and completed by a subset of the initial sample. These relationships with well-validated self-report scales support the convergent validity of the qIAT, and suggest that this assessment method measures the intended target construct.

Explicit-Implicit similarity: Assessment procedure

The relatively strong explicit-implicit correlations observed here may also be attributed to the structure of the assessment procedure of the qIAT, which reflects the extent to which respondents associate the target trait with truth about the self. In contrast to the self-concept IAT, assessment in the qIAT is not based on the problematic and vague association between the self and the measured trait (cf. Remue et al., 2013). To illustrate, consider a hypothetical example, in which no response biases of any kind are involved, of a respondent who scored low on self-report extraversion items such as “*I feel comfortable around people*”. Because response biases did not contaminate this hypothetical explicit score, it can be safely assumed that this person truly believes that such items do not describe himself or herself well. Still, the self in this case may be automatically associated with extraversion in the self-concept IAT, perhaps due to persistent preoccupation with issues related to this trait (see De Cuyper et al., 2017). In the qIAT, on the other hand, it is unambiguously expected that this person would associate such personality items with false self-related items (along with associating the inverse personality items with self-related truth). This specific automatic association is clearly predicted based on the assessment procedure of this task. Thus, the person's qIAT score (but not the self-concept IAT score) unambiguously represents the levels of the target construct. In sum, the response patterns on the self-report and on the qIAT (but not necessarily on the self-concept IAT) that assess the same construct are expected to mirror each other, thus strengthening the explicit-implicit relationship between them.

A close examination of the assessment procedure of the qIAT, which minimizes the differences between the implicit and explicit modes of assessment, may shed light on the nature of the assessment this task yields (see, for example, Schimmack, 2019). Explicit questionnaires' assessments are typically based on the degree to which respondents rate their agreement with several propositions that reflect the measured construct, while the qIAT assesses the association

between similar propositions and self-related truth. Self-report questionnaires arguably assess the explicit aspects of the truth value (cf. Shidlovski et al., 2014) of personality propositions, which are based on processes such as deliberate evaluations and logical reasoning (see Gawronski & Bodenhausen, 2011). In the same vein, the qIAT assesses the spontaneously assigned or implicit truth value of the same information (see Wyer & Radvansky, 1999). Thus, based on this perspective (Shidlovski et al., 2014), the output of the indirect assessment procedure of the qIAT may reflect the implicit truth value of propositional information related to the self-concept.

Limitations and future directions

The present findings support the psychometric properties of the qIAT, but more work is needed to further examine the reliability and validity of this task. Most importantly, the findings of Study 2 and Study 3 provided initial support for the predictive validity of the qIAT (measuring conscientiousness), but future studies will need to further test the utility of the qIAT in predicting actual behavior. Also, the samples in all studies were recruited online, which might restrict the generalizability of the findings.

The predictive validity of implicit personality assessment was possibly influenced by the specificity of the measured constructs (Costantini et al., 2015). Gawronski and De Houwer (2014) suggested that a closer match between the measurement procedures and the criterion behavior may improve the prediction of implicit tasks in real-life situations. Similarly, Banse et al. (2015) recommended using verbs rather than adjectives as attribute stimuli in the IAT, in order to increase the criterion validity of this task (see also De Cuyper et al., 2017). The assessment in the qIAT is based on questionnaire items, which often describe certain well-defined behavioral tendencies that reflect the measured construct. In Studies 2 and 3, the qIAT that predicted the criterion behavior (returning to complete the second part of the study) included items that describe behavioral patterns closely related to that particular situation (e.g., “I complete tasks successfully”; “I follow through with my plans”). However, the scale we used in these studies, which measured the broad Big Five conscientiousness domain, included other items that seem less relevant to the predicted behavior (e.g., “I pay attention to details”). Costantini et al. (2015) argued that rather than measuring conscientiousness as a unidimensional variable, studies may be improved by focusing on the predictive validity of more specific, lower-level traits related to this broad personality dimension (e.g., responsibility, orderliness), which are associated with different mechanisms. This can be done with the qIAT, using the items of existing validated subscales that measure particular facets of broad and relatively heterogeneous personality variables. Future research can employ this task to assess specific and narrowly defined constructs, depending on the context of the predicted behavior.

Finally, the extent to which the qIAT score can be intentionally biased is currently not known. The assessment

procedure of the this task is largely based on the autobiographical IAT (aIAT; Sartori et al., 2008), which was introduced as a lie-detection tool. Later findings, however, suggested that the aIAT is not resistant to faking (Verschuere et al., 2009; but see Agosta et al., 2011). Implicit measures in general are not completely immune to self-presentation strategies (Gawronski & De Houwer, 2014), but perhaps a procedure that is based on reaction times to whole sentences is particularly prone to such intentional biases, compared to other implicit assessment methods.

Conclusion

The qIAT is an implicit assessment method that combines the simplicity and flexibility of the IAT with the rating procedure, semantic complexity, and established validity of the items of existing, ordinary self-report scales. The findings of the present studies, which focused on two different Big-5 dimensions, support the reliability and validity of this task. The qIAT provides access for the indirect measurement of a wide variety of distinct psychological phenomena tapped by standard self-report questionnaires. As such, it may facilitate the integration of implicit assessment into mainstream personality research.

Acknowledgments

The authors would like to thank Gidi Aviram, Ruti Mayo, Tali Kleiman, and Yarden Weis for their helpful comments. This research was supported by the Israeli Science Foundation Grant 886/18 to Iftah Yovel

Open Scholarship



This article has earned the Center for Open Science badges for Open Data, Open Materials and Preregistered through Open Practices Disclosure. The data and materials are openly accessible at <https://osf.io/z52mw/>, <https://osf.io/z52mw/> and <https://osf.io/z52mw/>.

Funding

This work was supported by Israel Science Foundation Grant 886/18 to Iftah Yovel.

References

- Agosta, S., Ghirardi, V., Zogmaister, C., Castiello, U., & Sartori, G. (2011). Detecting fakers of the autobiographical IAT. *Applied Cognitive Psychology, 25*(2), 299–306. doi:10.1002/acp.1691
- Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*(3), 533–548. doi:10.1037/a0016229
- Banse, R., Messer, M., & Fischer, I. (2015). Predicting aggressive behavior with the aggressiveness-IAT. *Aggressive Behavior, 41*(1), 65–83. doi:10.1002/ab.21574
- Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods, 46*(3), 668–688. doi:10.3758/s13428-013-0410-6

- Carlsson, R., & Agerström, J. (2016). A closer look at the discrimination outcomes in the IAT literature. *Scandinavian Journal of Psychology*, 57(4), 278–287. doi:10.1111/sjop.12288
- Costa, P. T. J., & McCrae, R. R. (1992). NEO-PI-R professional manual: Revised NEO personality and NEO Five-Factor Inventory (NEO-FFI). Psychological Assessment Resources. [Database]
- Costantini, G., Richetin, J., Borsboom, D., Fried, E. I., Rhemtulla, M., & Perugini, M. (2015). Development of indirect measures of conscientiousness: Combining a facets approach and network analysis. *European Journal of Personality*, 29(5), 548–567. doi:10.1002/per.2014
- Currie, C. J., Katz, B. A., & Yovel, I. (2017, September). Explicit and implicit shame aversion predict symptoms of avoidant and borderline personality disorders. *Journal of Research in Personality*, 71, 13–16. doi:10.1016/j.jrp.2017.08.006
- De Cuyper, K., De Houwer, J., Vansteelandt, K., Perugini, M., Pieters, G., Claes, L., & Hermans, D. (2017). Using indirect measurement tasks to assess the self-concept of personality: A systematic review and meta-analyses. *European Journal of Personality*, 31(1), 8–41. doi:10.1002/per.2092
- de Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, 8(7), 342–353. doi:10.1111/spc3.12111
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, 135(3), 347–368. doi:10.1037/a0014211
- Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model. Theory, evidence, and open questions. *Advances in Experimental Social Psychology*, 44, 59–127. doi:10.1016/B978-0-12-385522-0.00002-0
- Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. *Handbook of Research Methods in Social and Personality Psychology*, 1(519), 283–310. <http://users.ugent.be/~jdhouwer/chapterbertram.pdf>
- Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures. *Personality and Social Psychology Bulletin*, 43(3), 300–312. doi:10.1177/0146167216684131
- Gawronski, B., & Payne, B. K. (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. Guilford Press. [Database]
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26–42. doi:10.1037/1040-3590.4.1.26
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96. doi:10.1016/j.jrp.2005.08.007
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. doi:10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. doi:10.1037/0022-3514.85.2.197
- Grumm, M., & von Collani, G. (2007). Measuring Big-Five personality dimensions with the implicit association test – Implicit personality traits or self-esteem? *Personality and Individual Differences*, 43(8), 2205–2217. doi:10.1016/j.paid.2007.06.032
- 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(10), 1369–1385. doi:10.1177/0146167205275613
- Krause, S., Back, M. D., Egloff, B., & Schmukle, S. C. (2011). Reliability of implicit self-esteem measures revisited. *European Journal of Personality*, 25(3), 239–251. doi:10.1002/per.792
- LeBel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin*, 37(4), 570–583. doi:10.1177/0146167211400619
- Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, 15(4), 152–159. doi:10.1016/j.tics.2011.01.005
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 37–48). Erlbaum. doi:10.1097/00005053-195311720-00010
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The Role of Constructs in Psychological and Educational Measurement* (pp. 49–69). Erlbaum. doi:10.1097/00005053-195311720-00010
- Reimers, S., & Stewart, N. (2007). Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities. *Behavior Research Methods*, 39(3), 365–370. doi:10.3758/BF03193004
- Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M. A., De Raedt, R., Houwer, J. D., ... Barnes-Holmes, D. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, 27(8), 1441–1449. doi:10.1080/02699931.2013.786681
- Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science*, 19(8), 772–780. doi:10.1111/j.1467-9280.2008.02156.x
- Schimmack, U. (2019). The implicit association test: A method in search of a construct. *Perspectives on Psychological Science*. doi:10.1177/1745691619863798
- Schnabel, K., & Asendorpf, J. B. (2010). The self-concept: New insights from implicit measurement procedures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 408–425). Guilford Press.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105. doi:10.1037/0003-066X.54.2.93
- Shidlovski, D., Schul, Y., & Mayo, R. (2014). If I imagine it, then it happened: The Implicit Truth Value of imaginary representations. *Cognition*, 133(3), 517–529. doi:10.1016/j.cognition.2014.08.005
- Steffens, M. C., & Schulze König, S. (2006). Predicting spontaneous big five behavior with Implicit Association Tests. *European Journal of Psychological Assessment*, 22(1), 13–20. doi:10.1027/1015-5759.22.1.13
- Verschuere, B., Prati, V., & Houwer, J. D. (2009). Cheating the lie detector. *Psychological Science*, 20(4), 410–413. doi:10.1111/j.1467-9280.2009.02308.x
- Vianello, M., Robusto, E., & Anselmi, P. (2010). Implicit conscientiousness predicts academic performance. *Personality and Individual Differences*, 48(4), 452–457. doi:10.1016/j.paid.2009.11.019
- Wilson, T. D. (2009). Know thyself. *Perspectives on Psychological Science*, 4(4), 384–389. doi:10.1111/j.1745-6924.2009.01143.x
- Wyer, R. S., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychological Review*, 106(1), 89–118. doi:10.1037/0033-295X.106.1.89
- Yovel, I., & Friedman, A. (2013). Bridging the gap between explicit and implicit measurement of personality: The questionnaire-based implicit association test. *Personality and Individual Differences*, 54(1), 76–80. doi:10.1016/j.paid.2012.08.015