TPM®

A FAIRER COMPARISON BETWEEN THE IMPLICIT ASSOCIATION TEST AND THE SINGLE CATEGORY IMPLICIT ASSOCIATION TEST

OTTAVIA M. EPIFANIA
PASQUALE ANSELMI
EGIDIO ROBUSTO
UNIVERSITY OF PADOVA

The Implicit Association Test (IAT) and the Single Category Implicit Association Test (SC-IAT) are the most common procedures for the assessment of implicit attitudes, stereotypes, and self-concept. The IAT has been found to outperform the SC-IAT in predicting behavioral choices. However, the comparison between the IAT and the SC-IAT might have been affected by many differences in both the administration and the scoring of the two procedures. This study was aimed at developing a common scoring procedure for the IAT and the SC-IAT that would allow a fair comparison of the two. Results supported the higher accuracy of the IAT in predicting the behavioral choice, regardless of the scoring procedure. Implications of the results and limitations of the study are discussed.

Keywords: IAT; SC-IAT; Implicit measures; Implicit preference; Behavioral outcomes; *D*-score.

Correspondence concerning this article should be addressed to Ottavia M. Epifania, Department FISPPA, Section of Applied Psychology, University of Padova, Via Venezia 14, 35131 Padova (PD), Italy. Email: marinaottavia.epifania@phd.unipd.it

Within the past two decades, implicit measures of attitudes, stereotypes, and self-concept have been widely used for the investigation of different topics, ranging from consumers behaviors (e.g., Diamantopoulos, Florack, Halkias, & Palcu, 2017; Karnal, Machiels, Orth & Mai, 2016; Kudo & Nagaya, 2017) to addiction behaviors (e.g., Chen et al., 2018; Keough, O'Connor, & Colder, 2016; Li, Langham, Browne, Rockloff & Thorne, 2018; Montes, Olin, Teachman, Baldwin, & Lindgren, 2018) and social cognition (e.g., Anselmi, Vianello, & Robusto, 2013; Anselmi, Vianello, Voci, & Robusto, 2013; Anselmi, Voci, Vianello & Robusto, 2015; Greenland, Xenias, & Maio, 2017). Among these procedures, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and the Single Category-IAT (SC-IAT; Karpinski & Steinman, 2006) are two of the most commonly used. Both are based on the speed and accuracy with which stimuli are sorted into their reference categories. The main feature distinguishing the IAT from the SC-IAT is the number of categories for the target objects, namely two for the former and one for the latter, and hence the measure of the implicit preference that can be obtained. While the IAT results in a comparative measure of the preference toward one of the target objects in comparison to its contrasted object, the SC-IAT results in an absolute positive or negative evaluation of the target object. This study deals with the comparison of the predictive capacity of both the IAT and SC-IAT.

Karpinski and Steinman (Study 1, 2006) provided a direct comparison between the IAT and the SC-IAT. These authors investigated the capacity of a Coke-Pepsi IAT, a Coke SC-IAT, and a Pepsi SC-IAT in predicting the choice between Coke and Pepsi. Results showed that both the Coke-Pepsi IAT and the Pepsi

SC-IAT allowed for predicting the soda choice, while the Coke SC-IAT was unrelated to the choice. This result suggests that the soda choice is guided more by Pepsi preference/positive evaluation than by Coke dislike/negative evaluation. Nonetheless, the comparison between the IAT and the SC-IAT predictive ability has been poorly investigated. To the best of our knowledge, only Karpinski and Steinman (2006) provided a systematic comparison between the two implicit measures, but their study was limited by some drawbacks. Firstly, the sample size was rather small, and results should hence be interpreted with caution. Secondly, issues related to both administration and scoring of the two implicit measures might have affected the comparison. Concerning the administration, the IAT and the SC-IAT differed in the number of both trials and stimuli used to represent each category. The SC-IAT employed more stimuli than the IAT, for both the evaluative dimensions (twenty-one stimuli for each SC-IAT evaluative dimension versus five stimuli for each IAT evaluative dimension) and the object stimuli (seven stimuli for each SC-IAT target object category and five stimuli for each IAT target object category). Furthermore, while the administration of the SC-IAT included a response time window, for which after 1,500 ms the stimulus on the screen disappeared, the IAT did not have such a constraint. The SC-IAT also included feedback for each response, signaling to the respondents whether they hit the correct or incorrect response key, while the IAT did not. The procedures differed also in terms of the labels used to represent the positive and negative attribute categories (*pleasant* and *unpleasant* for the IAT, and *good* and *bad* for the SC-IAT) and on the response keys used to sort the stimuli. Regarding the scoring, the IAT scores were computed according to the *D*-score procedure in Greenwald, Nosek, and Banaji (2003), while for the SC-IATs an ad-hoc *D* measure was adapted from the usual IAT *D*-score.

Given the differences between administration and scoring, the comparison between the IAT and the SC-IAT capacity to predict a behavioral outcome might have been unfair. To the best of our knowledge, no scoring procedure employing the same criteria on both the IAT and the SC-IAT exists, nor any attempt to align the two implicit procedures to allow for a fairer comparison between their predictive capacity. It would be interesting to compare the predictive capacity of the two implicit measures by using the same scoring procedure and keeping the administration as similar as possible while acknowledging their key features (e.g., block types and usual length of the blocks). If when using the same scoring method and when reducing the differences related to the administration procedures there are still differences in the predictive capacity of the IAT and the SC-IAT, these differences can be reasonably attributed to the implicit procedure itself.

The aim of this study was hence to provide a fairer comparison of the predictive capacity of the two measures. This was done by aligning both administration (e.g., stimuli, response time window, feedback) and scoring of the two procedures.

The Implicit Association Test and the Single Category Implicit Association Test

The IAT is a computerized task in which stimuli related to contrasted objects, like images of *flowers* or *insects*, are mapped into their reference categories together with two evaluative attribute stimuli (e.g., *good* and *bad* words). The structure of a usual IAT is reported at the top of Table 1. Three of the IAT blocks (Blocks 1, 2, and 5) are mere practice blocks. The remaining blocks (Blocks 3 and 4 and Blocks 6 and 7) constitute the two associative conditions of the IAT. The categorization task is supposed to be easier, in terms of higher accuracy and slower response times, in the condition that is congruent with the respondents' automatically activated association (the so-called compatible condition) than the task against their automatically activated association (the so-called incompatible condition). The difference between respondents'

TABLE 1
Structures of a Flowers-Insects IAT (top) and a Flowers SC-IAT (bottom)

Block	Trial	Function	Left key	Right key
		IAT		
1	20	Practice	Good	Bad
2	20	Practice	Flowers	Insects
3	20	Associative practice	Good + Flowers	Bad + Insects
4	40	Test	Good + Flowers	Bad + Insects
5	20	Practice	Insects	Flowers
6	20	Associative practice	Good + Insects	Bad + Flowers
7	40	Test	Good + Insects	Bad + Flowers
		SC-IAT		
1	24	Practice	Good + Flowers	Bad
2	72	Test	Good + Flowers	Bad
3	24	Practice	Good	Bad + Flowers
4	72	Test	Good	Bad + Flowers

Note: IAT = Implicit Association Test; SC-IAT = Single Category Implicit Association Test.

performance in the two associative conditions is known as the IAT effect and the *D*-score (Greenwald et al., 2003) is the most common measure of this effect. Although many algorithms are available for the computation of the *D*-score (Greenwald et al., 2003; see Table 2 for further details), the core idea underlying them is to standardize the difference between the average response time in the compatible and incompatible conditions by the standard deviation of the pooled trials of both. The differences between the algorithms mainly concern the penalty for incorrect responses and the deletion of fast responses. Regardless of the specific strategy chosen for its computation, the *D*-score is a convenient measure to summarize the IAT effect. Sticking with the Flowers-Insects IAT example, a positive *D*-score indicates faster responses in associating flowers with positive attributes and insects with negative attributes than the opposite. Conversely, a negative *D*-score indicates faster responses in associating insects with positive attributes and flowers with negative attributes than the opposite. However, one issue concerning the IAT and its *D*-score is that it is not possible to disentangle the attitude toward one of the two objects (e.g., toward flowers) from that toward the contrasted object (e.g., toward insects) (Karpinski & Steinman, 2006; Stefanutti, Robusto, Vianello & Anselmi, 2013). The identification of an absolute attitude in the IAT would require the use of appropriate formal models for the decomposition of the IAT effect, such as the Discrimination-Association Model (DAM; Stefanutti et al., 2013; Stefanutti, Vianello, Anselmi, & Robusto, 2014). Furthermore, there are some cases in which a complementary category for the target object is not clearly identifiable. For example, when assessing self-esteem (Karpinski & Steinman, 2006), the object category *I* is contrasted with the object category *Others*.

The SC-IAT (Karpinski & Steinman, 2006) has been proposed as an alternative to the IAT when the aim is to obtain an absolute measure toward one target object. Its structure is reported at the bottom of Table 1. As stated above, the SC-IAT includes only one category for the target object, along with the two evaluative attribute categories, and hence it is supposed to provide a measure of the absolute attitude (either positive or negative) toward the target object. As for the IAT, the categorization tasks take place under two different associative conditions, and the SC-IAT effect is determined by the difference in respondents' performance between the two associative conditions. A modified version of the IAT *D*-score is used to capture this effect (see Table 2 for further details). Since the difference is usually computed between the average response time in the condition where the target object is mapped with the negative attribute and the condition in which the target object is associated with the positive attribute, a positive score would indicate faster response times in the latter condition than in the former. Conversely, a negative score would indicate faster response times in associating the target object with the negative attribute than with the positive attribute. For instance, a Flowers SC-IAT would result in an absolute measure of the attitude (positive or negative) toward flowers. If the research interest is on the preference for flowers or insects, two SC-IATs should be designed, one for the assessment of the attitude toward flowers (i.e., a Flowers SC-IAT), and one for the assessment of the attitude toward insects (i.e., an Insects SC-IAT). The preference for flowers or insects would then result from a comparison of the two SC-IAT effects.

METHOD

To test the predictive capacity of the new scoring procedures, a Dark-Milk chocolate IAT, one Milk chocolate SC-IAT, and one Dark chocolate SC-IAT were developed.

The decision to use chocolate as the target object was driven by different reasons. Firstly, chocolate preference should not be sensitive to social desirability, and hence participants would have no concerns in reporting their actual chocolate preference. Moreover, it offers the chance to ask for a behavioral choice disguised as a reward for participation.

Participants

Participants were recruited at the University of Padova. One-hundred and sixty-one people (female = 63.55%, age = 23.95 ± 2.83) volunteered to take part in the study, with no compensation. Participants were informed about the confidentiality of the data. They were asked for their consent to take part in the study and had the chance to withdraw their participation at any time. The majority of the participants were students (94.08%), including both undergraduates, master, and Ph.D. students. Only two participants reported having a Ph.D. title, while most reported having a bachelor's degree (43.42%), immediately followed by those who reported having a high school diploma (32.24%), and a master's degree (23.03%). The Inquisit 3.0 software was used to administer the IAT and the SC-IATs.

Materials and Procedure

For the chocolate stimuli, seven images of chocolate were modified to represent either Dark or Milk chocolate, resulting in seven images for each type of chocolate. Three independent judges evaluated the stimuli regarding their properties, specifically whether they were clearly identifiable as Dark or Milk chocolate images. The three judges agreed on the representativeness of the category to which the stimuli belonged. All the chocolate images were presented on a white background. The stimuli and the Inquisit script to run the experiment can be retrieved in an online repository (https://osf.io/cnq4u/). In the Dark-Milk chocolate IAT, both dark and milk chocolate images were used, while in the two SC-IATs only either dark (Dark chocolate SC-IAT) or milk (Milk chocolate SC-IAT) chocolate images were used. In all the implicit procedures, the evaluative attribute categories were composed of 13 stimuli each. They were labeled as *Positive* (i.e., "good," "laughter," "pleasure," "glory," "peace," "happiness," "joy," "love," "wonderful," "beautiful," "excellent," "heaven," "marvelous") or *Negative* (i.e., "evil," "bad," "horrible," "terrible," "annoying," "pain," "failure," "hate," "nasty," "disaster," "agony," "ugly," "disgust"), while the target object categories were labeled as *Dark* or *Milk.* The response key "E" was used to sort the stimuli belonging to left-side categories and the response key "I" was used to sort the stimuli into the right-side categories. Unlike the SC-IAT procedure in Karpinski and Steinman (2006), the SC-IAT practice Blocks 1 and 3 were composed of 20 trials, instead of 24, as for the practice blocks of the IAT. Neither the IAT nor the SC-IATs included a built-in correction or a response time window, and hence participants did not receive any feedback on their performance. Participants were instructed to be as fast and accurate as they could in performing the tasks.

Participants were explicitly asked to evaluate Dark and Milk chocolates via two items ("How much do you like Dark/Milk chocolate?") on a 6-point Likert-type scale (0 = *Not at all*, 5 = *Very much*). While the order of presentation of the implicit measures was counterbalanced across participants, the explicit assessment was kept constant at the end of the experiment. As a reward, participants were offered with a dark or milk chocolate bar. The experimenter registered respondents' choices after they left the laboratory.

Dark-Milk chocolate IAT: The critical blocks were composed of 60 trials each (20 practice + 40 test), defining the Dark-Good/Milk-Bad condition (DGMB), and the Milk-Good/Dark-Bad condition (MGDB).

Dark chocolate SC-IAT: The critical blocks were composed of 72 trials each, defining the Dark-Good/Bad (DG) and the Good/Dark-Bad (DB) conditions.

Milk chocolate SC-IAT: As for the Dark chocolate SC-IAT, the critical blocks were composed of 72 trials each, defining the Milk-Good/Bad (MG) and the Good/Milk-Bad (MB) conditions.

Classical and Modified Scoring Procedures

All the *D*-scores procedures in Greenwald et al. (2003) not including the built-in correction for error responses were computed for the IAT, while the procedure described in Karpinski and Steinman (2006) was followed for the computation of the SC-IAT *D.* Given that the SC-IAT administration procedure did not include a time response window and there are no guidelines concerning the upper response times treatment, no upper time responses deletion was applied. An overview of the classical and modified scoring procedures is illustrated in Table 2.

While the classical procedure for the SC-IAT includes a default lower tail treatment, the lower tail treatment for the IAT depends on the *D*-score procedure employed. To have a comparable score for the two implicit measures, a common lower tail treatment for both procedures was set, according to which responses with a latency lower than 350 ms are discarded. Since it is not uncommon to find SC-IATs with no response window in their administration procedure, a common upper tail treatment for response times was proposed for both the implicit measures. For the SC-IAT upper tail treatment, it might be argued that deleting responses faster than 1,500 ms (i.e., the response time window cut-off) may be a more appropriate threshold for slow responses. Nonetheless, the presence of the response time window itself produces an urge to respond that is missing when the response time window is not included in the administration procedure (Karpinski & Steinman, 2006).

TABLE 2
Classical and modified effect size scores

	Classical scores					Modified scores							
	IAT				SC-IAT								
	D-score 3	*D*-score 4	*D*-score 5	*D*-score 6	*D*	m1	m2	m3	m4	m5	m6	m7	m8
Included Blocks	IAT: Blocks 3-4 and Blocks 6-7; SC-IAT: Block 2 and Block 4												
Lower tail treatment	None	None	< 400 ms	< 400 ms	< 350 ms	< 350 ms							
Upper tail treatment	> 10,000 ms				RTW	> 10,000 ms							
Error treatment	M (C) + 2 SD	M (C) + 600 ms	M (C) + 2 SD	M (C) + 600ms	M (C) + 400 ms	Mean (C) + 2 *SD* (C)				Mean (C) + 2 *SD*			
Denominator	Pooled trials *SD*					Pooled trials *SD*		Cohen's pooled *SD*		Pooled trials *SD*		Cohen's pooled *SD*	
Trials for the denominator	All trials				Correct trials	Correct trials	All trials	Correct trials	All trials	Correct trials	All trials	Correct trials	All trials

Note: IAT = Implicit Association Test; SC-IAT = Single Category Implicit Association Test; RTW = response time window; *M* (C) = mean (or standard deviation) computed only on correct response trials. The response time window upper tail treatment was not applied on this data because it was not included in the administration procedure.

The only difference between the error treatment of the modified scoring procedures was related to the trials for the computation of the standard deviation. In the first four modified procedures, the *SD* is computed considering only the latencies of the correct responses, while for the latter four procedures the *SD* is computed on all the latencies, hence including those for error responses. The decision to use only standard deviations as penalty strategies is because the latencies in the SC-IATs tend to be faster than latencies in the IAT. Therefore, assuming that 600 ms would be a reasonable time for correcting the error response might be a too strong assumption for the SC-IAT data. Conversely, the penalty used in the SC-IAT (400 ms) might not be enough to acknowledge the response time needed to correct the error response in the IAT.

The pooled-trials standard deviation used in the classical measure was compared with Cohen's pooled standard deviation. The pooled-trials standard deviation and Cohen's pooled standard deviation were computed considering either only the correct responses or all the trials. The variability due to incorrect responses is not accounted for in the former case, while it is addressed in the latter.

Finally, the IAT modified procedures were computed as the difference between the two associative conditions, rather than the mean of the standardized average response time differences between the practice and test blocks.

Both classical and modified procedures for the IAT were computed so that positive scores indicated faster responses in associating Milk chocolate with positive attributes and Dark chocolate with negative attributes, and hence a likely preference for milk chocolate and/or a dislike for dark chocolate. Conversely, negative scores indicated faster responses in associating Dark chocolate with positive attributes and Milk chocolate with negative attributes, suggesting a probable preference for Dark chocolate and/or a dislike for Milk chocolate.

For the SC-IATs, both the classical and modified procedures were computed so that positive scores indicated faster responses in associating the target chocolate with positive attributes than with negative attributes, hence showing a likely positive attitude toward the target chocolate. Conversely, negative scores indicated faster responses when the target chocolate was associated with negative attributes, and hence a probable negative attitude toward the target chocolate.

Data were analyzed with R (R Core Team, 2018). The IAT *D*-scores were computed by means of the DscoreApp (Epifania, 2019).

RESULTS

Data from nine participants were discarded. Eight of them explicitly reported not understanding the tasks they were asked to perform in either the IAT or one of the SC-IATs, while one of them had too many fast responses, specifically on the Dark chocolate SC-IAT (more than 30% of responses with a latency lower than 350 ms). The final sample was composed of 152 participants (female = 63.82 %, age = 24.03 ± 2.82). Milk chocolate was chosen by 48.03% of the participants.

The median for the explicit evaluation of Dark chocolate was 3 (first quartile = 2, third quartile = 5). The median for the explicit evaluation of Milk chocolate was 4 (first quartile = 3, third quartile = 4).

No SC-IAT trials were eliminated because of latencies greater than 10,000 ms, while three IAT trials were eliminated because of latencies greater than 10,000 ms. The IAT had the lowest percentage of trials with latencies faster than 400 ms (1.39%) and faster than 350 ms (0.19%). The percentage of trials with a latency faster than 350 ms was 1.00% in the Milk chocolate SC-IAT and 0.90 % in the Dark chocolate SC-

IAT. The two SC-IATs showed a similar percentage of trials faster than 400 ms (4.40% and 4.32% for the Dark chocolate SC-IAT and the Milk chocolate SC-IAT, respectively).

All the implicit measures had the same overall percentage of correct responses (95%). Average response latencies were similar in the two SC-IATs (Dark chocolate SC-IAT = 679.45 ± 328.72 ms, Milk Chocolate SC-IAT = 675.90 ± 322.31 ms). IAT average response time was 862.03 ± 496.50 ms. Descriptive statistics for all the measures, along with their correlation with explicit measures, are reported in Table 3.

The SC-IATs measures, both classical and modified, have smaller effect sizes than the IAT measures. The modified measures for the IAT, particularly the ones using Cohen's pooled standard deviation, displayed larger effect sizes than the classical IAT measures, while the modified measures for the SC-IATs were more similar to each other.

The explicit dark chocolate evaluation was negatively and moderately correlated with the explicit milk chocolate evaluation (r = –.386, p < .001). Both the IAT and the Dark SC-IAT classical measures significantly correlated with both explicit chocolate evaluations, while the Milk SC-IAT classical measure correlated with neither Dark nor Milk chocolate explicit evaluation. The IAT modified measures significantly and moderately correlated with both the explicit chocolate evaluations. The modified measures of both the Dark chocolate SC-IAT and the Milk chocolate SC-IAT significantly correlated with the Milk chocolate explicit evaluation, while only the first four modified measures of the Dark chocolate SC-IAT significantly correlated with the Dark chocolate explicit evaluation. Moreover, the correlation between the Dark chocolate explicit evaluation and the Milk chocolate SC-IAT was near zero for both the classical and the modified scores.

Pearson correlations were computed between the classical and modified scores and the explicit chocolate evaluations to check for the consistency of the scores. Correlation coefficients between the classical IAT *D*-scores ranged between .993 and .998 (all *p*s < .001), while correlations between classical IAT measures and the *D*-Dark classical measure ranged between –.214 and –.209 (all *p*s < .01). No correlations were found between classical *D*-scores and the classical *D*-Milk (correlations ranged between –.043 and –.030, all *p*s > .05). Classical *D*-Dark and *D*-Milk positively correlated with each other (r = .148, p > .05), but the correlation was not significant. IAT modified scores correlations with each other ranged between .974 and .999 (all *p*s < .001), while their correlations with the modified *D*-Dark scores ranged between –.313 and –.278 (all *p*s < .001). The modified *D*-Milk scores and modified *D*-scores did not correlate with each other (correlation coefficients ranged between –.007 and .008, all *p*s > .05). Correlations between modified *D*-Dark scores ranged between .983 and 1.00 (all *p*s < .001), while the correlation between the modified *D*-Milk scores ranged between .987 and .999 (all *p*s < .001). The correlation between the modified *D*-Milk and *D*-Dark had the same direction than the correlation between the classical SC-IAT scores and ranged between .145 and .204. Interestingly, the correlation between all the *D*-Milk modified scores and the *D*-Dark modified from 5 to 8 (i.e., the scores in which the error responses are replaced with the mean added with twice the standard deviation computed on all trials) showed slightly stronger and significant correlations, ranging from .182 and .204 (all *p*s < .010). The complete correlation matrix between all the classical and modified scores is reported in an online repository (https://osf.io/cnq4u/).

Behavioral Choice

Both the classical and modified scores were regressed on the behavioral chocolate choice, coded as 0 for the Dark chocolate choice (DCC) and 1 for the Milk chocolate choice (MCC). For the IAT, each score

TPM Vol. 27, No. 2, June 2020
207-220

TABLE 3
Descriptive statistics of the scores and correlations with explicit chocolate evaluations r_{Milk}

	Modified	*M (SD)*	*Min*	*Max*	r_{Milk}	r_{Dark}	Classical	*M (SD)*	*Min*	*Max*	r_{Milk}	r_{Dark}
IAT	IAT m1	0.64 (0.62)	–1.91	1.72	.401***	–.361***	*D*-score 3	0.41 (0.41)	–1.29	1.25	.421***	–.376***
	IAT m2	0.64 (0.60)	–1.86	1.69	.402***	–.366***	*D*-score 4	0.39 (0.39)	–1.26	1.27	.414***	–.376***
	IAT m3	0.73 (0.73)	–2.25	2.84	.388***	–.344***	*D*-score 5	0.40 (0.41)	–1.29	1.29	.419***	–.373***
	IAT m4	0.72 (0.70)	–2.12	2.59	.390***	–.351***	*D*-score 6	0.39 (0.39)	–1.26	1.32	.412***	–.373***
	IAT m5	0.64 (0.63)	–2.29	1.72	.398***	–.354***						
	IAT m6	0.64 (0.60)	–1.85	1.69	.402***	–.364***						
	IAT m7	0.72 (0.75)	–2.34	2.85	.386***	–.340***						
	IAT m8	0.72 (0.71)	–2.11	2.60	.391***	–.350***						
Dark SC-IAT	Dark m1	–0.06 (0.35)	–0.98	1.07	–.223**	.176*	*D*-Dark	–0.05 (0.31)	–0.74	0.78	–.185*	.168*
	Dark m2	–0.06 (0.34)	–1.03	0.94	–.226**	.173*						
	Dark m3	–0.06 (0.36)	–1.01	1.07	–.219**	.174*						
	Dark m4	–0.06 (0.35)	–1.05	0.94	–.222**	.171*						
	Dark m5	–0.06 (0.36)	–1.00	1.13	–.195*	.158						
	Dark m6	–0.06 (0.35)	–0.95	0.99	–.201*	.157						
	Dark m7	–0.06 (0.36)	–1.04	1.13	–.191*	.156						
	Dark m8	–0.06 (0.35)	–0.97	1.00	–.197*	.155						
Milk SC-IAT	Milk m1	0.16 (0.39)	–1.92	1.22	.166*	.037	*D*-Milk	0.15 (0.33)	–0.93	1.21	.134	.063
	Milk m2	0.16 (0.39)	–1.93	1.13	.166*	.037						
	Milk m3	0.16 (0.41)	–1.92	1.50	.171*	.035						
	Milk m4	0.16 (0.40)	–1.94	1.38	.170*	.036						
	Milk m5	0.16 (0.38)	–1.39	1.23	.154	.050						
	Milk m6	0.16 (0.37)	–1.40	1.14	.154	.050						
	Milk m7	0.17 (0.39)	–1.39	1.51	.160*	.048						
	Milk m8	0.16 (0.39)	–1.40	1.39	.159*	.048						

Note: IAT = Implicit Association Test; SC-IAT = Single Category Implicit Association Test; r_{Milk}= Spearman correlation with explicit Milk chocolate attitude; r_{Dark}= Spearman correlation with explicit Dark attitude. Positive IAT scores indicate a preference for Milk chocolate over Dark chocolate. Positive SC-Dark scores indicate a positive evaluation of Dark chocolate. Positive SC-Milk scores indicate a positive evaluation of Milk chocolate.
*p < .05. **p < .01. ***p < .001.

was regressed on the choice. Since the choice is a dichotomous task in which Dark chocolate is contrasted with Milk chocolate, it is plausible that the positive and negative attitudes toward both types of chocolate play a role in determining the actual choice. However, the scores obtained from each of the SC-IATs would convey unique information regarding the positive or negative evaluation of only one type of chocolate, therefore lacking part of the information that might be included in the choice task. It can also be argued that, since the two SC-IATs are two distinct experiments, it would not be reasonable to use the linear combination of their scores to predict the actual choice. Grounding on these considerations, both the Dark chocolate SC-IAT and the Milk chocolate SC-IAT and their linear combination were used to predict the choice.

Nagelkerke's (1991) R^2 and model accuracy of prediction (Faraway, 2016) were used as criteria to investigate the scores best accounting for the actual choice. Specifically, model general accuracy (i.e., the ratio between the number of chocolate choices correctly identified by the model and the total number of choices), DCC accuracy (i.e., the ratio between the number of DCCs correctly identified by the model and the total number of observed DCCs), and MCC accuracy (i.e., the ratio between the number of MCCs correctly identified by the model and the total number of observed MCCs) were computed. Results of the logistic regressions are reported in Table 4.

TABLE 4
Results for the choice prediction

	β (*SE*)	Deviance (*df* = 150)	Nagelkerke's R^2	General accuracy	DCC accuracy	MCC accuracy
D-score 3	2.23*** (0.54)	188.22	.18	.64	.63	.66
D-score 4	2.26*** (0.55)	188.80	.18	.64	.66	.63
D-score 5	2.18*** (0.53)	188.92	.18	.63	.63	.63
D-score 6	2.22*** (0.55)	189.48	.17	.64	.65	.63
D Dark	–0.70 (0.54)	208.77	.01	.53	.65	.41
D Milk	0.35 (0.49)	209.99	.00	.53	.78	.26
IAT m1	1.35*** (0.34)	191.07	.16	.64	.65	.63
IAT m2	1.38*** (0.35)	190.91	.16	.62	.62	.63
IAT m3	1.07*** (0.28)	192.62	.15	.63	.65	.62
IAT m4	1.12*** (0.29)	192.09	.15	.64	.66	.63
IAT m5	1.35*** (0.34)	190.95	.16	.64	.65	.63
IAT m6	1.38*** (0.35)	190.73	.16	.64	.63	.64
IAT m7	1.07*** (0.28)	192.51	.15	.63	.65	.62
IAT m8	1.12*** (0.29)	191.93	.15	.64	.66	.63
Dark m1	–0.73 (0.48)	208.07	.02	.53	.62	.44
Dark m2	–0.72 (0.49)	208.24	.02	.53	.62	.42
Dark m3	–0.70 (0.47)	208.18	.02	.53	.62	.44
Dark m4	–0.69 (0.48)	208.33	.02	.52	.62	.41
Dark m5	–0.62 (0.47)	208.65	.02	.52	.63	.40
Dark m6	–0.63 (0.48)	208.72	.02	.52	.63	.40
Dark m7	–0.60 (0.46)	208.74	.02	.51	.63	.38
Dark m8	–0.60 (0.47)	208.80	.01	.51	.63	.38
Milk m1	0.33 (0.42)	209.86	.01	.53	.77	.26

(Table 4 continues)

Table 4 (continued)

	β (*SE*)	Deviance (*df* = 150)	Nagelkerke's R^2	General accuracy	DCC accuracy	MCC accuracy
Milk m2	0.33 (0.43)	209.88	.01	.53	.77	.26
Milk m3	0.32 (0.40)	209.83	.01	.52	.76	.26
Milk m4	0.32 (0.41)	209.85	.01	.53	.77	.26
Milk m5	0.31 (0.44)	209.99	.00	.50	.75	.23
Milk m6	0.31 (0.44)	209.99	.00	.52	.76	.26
Milk m7	0.30 (0.42)	209.95	.00	.50	.75	.23
Milk m8	0.30 (0.42)	209.96	.00	.51	.76	.25

Note. *SE* = standard error; DCC = Dark chocolate choice; MCC = Milk chocolate choice; IAT = Implicit Association Test. The null deviance for all the models is 210.48 (*df* = 151). β are the log-odds for the probability of choosing Milk chocolate.
****p* < .001.

Regardless of the score used, whether classical or modified, the IAT outperformed both SC-IATs in predicting the chocolate choice. Models including the IAT scores showed the highest values of Nagelkerke's R^2. Besides, the IAT was the most accurate in predicting both chocolate choices. Both SC-IATs displayed lower values of Nagelkerke's R^2, particularly the models including only the Milk chocolate SC-IAT. Classical and modified IAT scores tended to have similar values of both Nagelkerke's R^2 and accuracy of prediction, with the latter ones displaying slightly lower values. All the modified measures for the Dark chocolate SC-IAT resulted in slightly higher values of Nagelkerke's R^2, while just the first four modified Milk chocolate SC-IAT measures showed slightly higher values than the classical score. Regarding the accuracy of prediction, the modified SC-IAT scores showed a slightly worse performance. Results of choice prediction including the linear combination of the SC-IAT scores are reported in Table 5.

TABLE 5
Choice prediction: SC-IAT scores linear combination

	β_{dark} (*SE*)	β_{milk} (*SE*)	Deviance (*df* = 149)	Nagelkerke' R^2	General accuracy	DCC accuracy	MCC accuracy
Classical	–0.77 (0.55)	0.46 (0.50)	207.93	.02	.55	.67	.41
m1	–0.81 (0.49)	0.44 (0.43)	206.99	.03	.56	.66	.45
m2	–0.80 (0.50)	0.44 (0.44)	207.19	.03	.53	.66	.40
m3	–0.78 (0.48)	0.43 (0.41)	207.07	.03	.54	.66	.41
m4	–0.77 (0.49)	0.43 (0.42)	207.28	.03	.54	.67	.40
m5	–0.71 (0.48)	0.44 (0.45)	207.70	.02	.55	.68	.40
m6	–0.72 (0.49)	0.44 (0.45)	207.78	.02	.54	.68	.38
m7	–0.68 (0.47)	0.42 (0.43)	207.77	.02	.55	.68	.40
m8	–0.69 (0.48)	0.42 (0.44)	207.85	.02	.54	.68	.38

Note: SE = standard error; DCC = Dark chocolate choice; MCC = Milk chocolate choice. The null deviance for all the models is 210.48 (*df* = 151). β are the log-odds for the probability of choosing Milk chocolate.

The linear combination of both SC-IAT scores predicted the chocolate choice better than the two SC-IAT scores considered separately, but it was still outperformed by the IAT. The *D*-Dark coefficients, both classical and modified, tended to be higher than the coefficients of the *D*-Milk, both classical and modified. Nonetheless, the linear combination of the first four modified scores resulted in higher values of Nagelkerke's R^2, when compared with both the classical scores and the last four modified scores. All the scores resulted in similar general accuracy of prediction and DCC accuracy, while a slightly higher variability on the MCC accuracy was observed. Specifically, the linear combination of the modified measures m6 and m8 showed the worst performance, while the combination of the modified measures m1 resulted in the highest MCC accuracy.

As a final analysis, the incremental validity of the IAT and the two SC-IATs with respect to the self-report chocolate evaluations was investigated. Four hierarchical multiple logistic regressions for predicting the chocolate choice were specified for each of the scoring procedures. In the first step, the Dark and Milk chocolate explicit evaluations were included. The IAT *D*-score entered in the second step. The *D*-Dark entered in the third step, and the *D*-Milk entered in the fourth step. This was done for both the classical and modified scores. Nagelkerke's R^2 was used as a criterion to decide whether the added predictor was useful to account for the chocolate choice. Nagelkerke's R^2 at the first step (i.e., the model including only the explicit chocolate evaluations) was 0.83. From the second step on, Nagelkerke's R^2 remained 0.84 for both the modified and the classical scores. It is reasonable to argue that the implicit measures do not add anything to the prediction afforded by the explicit measures. However, this result should be interpreted with caution because, in the present study, the explicit chocolate evaluation was asked right before the behavioral choice.

Discussion and Conclusions

By keeping the administration procedures as similar as possible and by using the same scoring procedures, it was possible to obtain interesting insights on the IAT and the SC-IAT, and their predictive performance of a behavioral outcome.

The IAT showed the highest correlation coefficients with both the explicit chocolate evaluations, while both the SC-IATs seemed to be mostly related to just one of the explicit chocolate evaluations. The IAT outperformed the SC-IATs on the prediction of the behavioral choice, both when the SC-IAT scores were considered as single measures and when their linear combination was considered. Considering all the above, it can be argued that the IAT has a higher predictive capacity than the SC-IAT. However, the higher predictive capacity of the IAT might also be due to the characteristics of the choice task itself. Since participants were presented with two different bowls of chocolate and were invited to take one chocolate bar, their like and/or dislike for both types of chocolate were concurrently playing a role in determining their choice. A measure able to include the comparative evaluation of the chocolate types, like the IAT, might hence best account for the real chocolate preference, and result in better accuracy of prediction, while measures dealing with only one of the components of the chocolate evaluation, like the SC-IAT, might be less effective. Even when SC-IAT scores were considered concurrently, and hence both the chocolate evaluations were included, the general accuracy of prediction was similar to the one obtained when the single measures were considered in the choice prediction. Nonetheless, it resulted in slightly better accuracy of prediction for the MCC, especially when compared to the performance of the Milk chocolate SC-IAT scores alone. This result supports the claim for which a measure including the comparative evaluation between two contrasted objects, instead of their absolute evaluation, results in better prediction of the choice between alternative options. Finally, the positive evaluation of one type of chocolate does not imply the negative evaluation of the other, because people can equally like both types of chocolate. Given that the better performance of the IAT might have been due to both the choice task and the type of preference assessed, it would be interesting to compare the performance of the two implicit measures in cases in which a clearly contrasted category is not identifiable, like in the *self-esteem* case. In this case, the SC-IAT might perform better than the IAT. Moreover, the SC-IAT might also perform better than the IAT in predicting a behavioral choice when the choice is not strictly dichotomous like in this study. For instance, respondents might be left free to choose between one type of chocolate, both types, or none, or even between different types of candy bars, including Dark and Milk chocolate bars.

Both in the study by Karpinski and Steinman (2006) and in the present one, the predictive validity of the implicit measures was assessed for nonsocially relevant stimuli, like soda and chocolate preference. Future studies should investigate the IAT and SC-IAT predictive validity with respect to socially relevant stimuli, such as members of stigmatized social groups. In pursuing this aim, different behavioral indicators might be used as a dependent variable. An example is the willingness to affiliate with members of the stigmatized social group. Another one is having or not having contacts with members of the stigmatized social groups.

Finally, Karpinski and Steinman (2006) administered the SC-IAT with a response time window of 1,500 ms. In the present study, no response time window was used. We could have applied an *a posterior* threshold for upper tail responses as if a response time window had been used in the administration of the implicit measure itself. However, we decided not to do so because using a response time window affects respondents' performance. Indeed, Karpinski and Steinman (2006) observed that the response time window influenced participants' speed by producing a sense of urgency in providing the response that was missing when the response time window was not included. Future research on the systematic comparison between the two implicit measures might include a response time window for both the IAT and the SC-IAT.

References

Anselmi, P., Vianello, M., & Robusto, E. (2013). Preferring thin people does not imply derogating fat people. A Rasch analysis of the implicit weight attitude. *Obesity*, *21*(2), 261-265. doi:10.1002/oby.20085

Anselmi, P., Vianello, M., Voci, A., & Robusto, E. (2013). Implicit sexual attitude of heterosexual, gay and bisexual individuals: Disentangling the contribution of specific associations to the overall measure. *Plos One*, *8*(11), e78990. doi:10.1371/journal.pone.0078990

Anselmi, P., Voci, A., Vianello, M., & Robusto, E. (2015). Implicit and explicit sexual attitudes across genders and sexual orientations. *Journal of Bisexuality*, *15*(1), 40-56. doi:10.1080/15299716.2014.986597

Chen, L., Zhou, H., Gu, Y., Wang, S., Wang, J., Tian, L., . . . Zhou, Z. (2018). The neural correlates of implicit cognitive bias towards Internet-related cues in Internet addiction: An ERP study. *Frontiers in Psychiatry*, *9*, 421. doi:10.3389/fpsyt.2018.00421

Diamantopoulos, A., Florack, A., Halkias, G., & Palcu, J. (2017). Explicit versus implicit country stereotypes as predictors of product preferences: Insights from the stereotype content model. *Journal of International Business Studies*, *48*(8), 1023-1036. doi:10.1057/s41267-017-0085-9

Epifania, O. M. (2019, April). *DscoreApp*. Retrieved from http://fisppa.psy.unipd.it/DscoreApp/

Faraway, J. J. (2016). *Extending the linear model with R* (2nd ed.). Boca Raton, FL: CRC Press.

Greenland, K., Xenias, D., & Maio, G. R. (2017). Effects of promotion and compunction interventions on real intergroup interactions: promotion helps but high compunction hurts. *Frontiers in Psychology*, *8*, 528. doi:10.3389/fpsyg.2017.00528

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (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

Karnal, N., Machiels, C. J., Orth, U. R., & Mai, R. (2016). Healthy by design, but only when in focus: Communicating non-verbal health cues through symbolic meaning in packaging. *Food Quality and Preference*, *52*, 106-119. doi:10.1016/j.foodqual.2016.04.004

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*(1), 16-32. doi:10.1037/0022-3514.91.1.16

Keough, M. T., O'Connor, R. M., & Colder, C. R. (2016). Testing the implicit and explicit cognitions underlying behavioral inhibition system-related drinking in young adults. *Alcoholism: Clinical and Experimental Research*, *40*(5), 1065-1074. doi:10.1111/acer.13036

Kudo, D., & Nagaya, K. (2017). Effects of matching and mismatching messages on purchase avoidance behavior following major disasters. *Psychology & Marketing*, *34*(3), 335-346. doi:10.1002/mar.20992

Li, E., Langham, E., Browne, M., Rockloff, M., & Thorne, H. (2018). Gambling and sport: Implicit association and explicit intention among underage youth. *Journal of Gambling Studies*, *34*(3), 1-18. doi:10.1007/s10899-018-9756-0

Montes, K. S., Olin, C. C., Teachman, B. A., Baldwin, S. A., & Lindgren, K. P. (2018). Hazardous drinking has unique relationships with implicit and explicit drinking identity. *Addictive Behaviors*, *87*, 155-161. doi:10.1016/j.addbeh.2018.07.011

Nagelkerke, N. J. D. (1991). A note on the general definition of the coefficient of determination. *Biometrika*, *78*(3), 691-692.

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Stefanutti, L., Robusto, E., Vianello, M., & Anselmi, P. (2013). A Discrimination–Association Model for decomposing component processes of the Implicit Association Test. *Behavior Research Methods*, *45*(2), 393-404. doi:10.3758/s13428-012-0272-3

Stefanutti, L., Vianello, M., Anselmi, P., & Robusto, E. (2014). GRace: A MATLAB-based application for fitting the discrimination-association model. *The Spanish Journal of Psychology*, *17*, E73. doi:10.1017/sjp.2014.67