Looking Past the Claims: a Psychometric Analysis of the Implicit Association Test

Hart Blanton, Texas A&M University, USA

I examine the theory, methods and analytic strategies surrounding the IAT to determine the psychometric model that a researcher embraces (knowingly or unknowingly) whenever this measure is used. This analysis reveals a causal model that is unusually restrictive, a measurement model that is not empirically supported, and a set of measurement assumptions that are typically violated. This framework can help explain why IAT scores might have undesirable test properties and it points to potential misapplications of the measure in marketing contexts. Recent attempts to address these limitations like new scoring algorithms introduce systematic confounds. Alternative new methods are suggested.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/14262/volumes/v36/NA-36

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
SESSION OVERVIEW

The Implicit Association Test has become a popular tool for measuring implicit attitudes, presumably because the IAT is characterized by a unique combination of impressive effect size and effortless implementation. The IAT attracted massive media attention (e.g., in “Blink” by Gladwell 2005), and was featured on TV shows (e.g. “The Oprah Winfrey Show” on 06/06/2007) as a tool to uncover implicit racism. Originally introduced to social psychology, the IAT has been extended to clinical psychology, organizational psychology, the law, and marketing. Marketing research companies have started to use the IAT as a measure of “true” preferences and brand associations.

Despite the huge interest sparked among academics and practitioners, critical examinations of the IAT’s methodology and predictive validity have received only limited attention. A notable exception are Brunel, Tietje, and Greenwald (2004) who discuss some of the methodological issues surrounding the IAT in marketing. However, the two most important issues, order-effects and assumptions behind the measurement model of the IAT, have not been investigated. Because methodologies developed in clinical research settings should undergo thorough critical examination before being used in applied disciplines such as Marketing, the symposium is aimed at investigating the role and consequences of order-effects and violations of IAT assumptions.

In the first paper, Messner and Vosgerau explain the logic of the IAT design and empirically demonstrate a procedural effect (an order effect of the IAT-blocks) and its impact on correlations of the IAT with predictor criteria. Counterbalancing or keeping order constant are shown not to alleviate the problem. In the second paper, Blanton explores the underlying measurement model and assumptions for the interpretation of IAT-effects. Stuettgen and Boatwright argue in the third paper that these assumptions (as identified by Blanton) are likely to be violated in marketing contexts, and examine via simulations how violations of the IAT-assumptions impact the validity of the IAT.

Concluding, order-effects and measurement assumptions are crucial determinants for the validity of the IAT in marketing research. The three papers explore the consequences of order-effects and violations of measurement assumptions for IAT-effects, and suggest solutions to overcome these issues.

EXTENDED ABSTRACTS

“Order Effects in the IAT”
Claude Messner, University of Basel, Switzerland
Joachim Vosgerau, Carnegie Mellon University, USA

Marketing researchers are becoming increasingly interested in non-conscious influences on consumer behavior. A central concept of non-conscious influences are implicit attitudes, which are defined as “...introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (Greenwald and Banaji 1995, p. 8). Implicit attitudes are thought of being potentially better predictors of behavior than explicit (self-reported) attitudes, because consumers might be unwilling to reveal their attitudes when the attitude concerns stigmatized behavior (e.g., racists attitudes), or because they are unable to reveal their attitudes as they might lack the ability for correct introspection (Brunel et al. 2004).

The Implicit Association Test (IAT) has become the most popular tool for measuring implicit attitudes, presumably because the IAT is characterized by a unique combination of impressive effect size and effortless implementation. The IAT was initially introduced to social psychology (Greenwald, McGhee, and Schwartz 1998), and has since been applied in clinical psychology, organizational behavior, the law, and marketing. Outside the academic realm, the IAT has attracted massive media attention as a tool to uncover implicit racism, and is used by marketing research firms as a measure of “true” preferences and brand associations.

We focus on an internal validity problem of the IAT that has not been investigated to date, so called order-effects. Consider an Coca Cola versus Pepsi IAT in which participants favor Coke over Pepsi. In the ‘compatible’ block participants learn to exploit the association between Coca Cola and the positive words, but in the subsequent ‘incompatible’ block must respond in the opposite direction, that is Coca Cola and negative words. Switching from applying one categorization rule to applying an opposite categorization rule is cognitively demanding and requires time and practice. So, ceteris paribus, cognitive inertia leads to slower responses in the second block, no matter whether it is ‘compatible’ or ‘incompatible’. Order-effects now accrue from the interplay of cognitive inertia (slower responses in the second block) and the IAT-effect (faster responses in the ‘compatible’ block). When the faster ‘compatible’ block comes first, cognitive inertia slows down responses in the subsequent ‘incompatible’ block, thereby augmenting the difference in response latencies between the two blocks (i.e. enlarging the IAT-effect). In contrast, when the ‘incompatible’ block precedes the faster ‘compatible’ block, cognitive inertia slows down responses in the faster ‘compatible’ block (i.e., decreasing the IAT-effect).

In study 1 we conducted an Coca Cola versus Pepsi IAT and counterbalance the order of the blocks. As expected, when the ‘compatible’ block came first an IAT-effect in favor of Coca Cola was found (t(24)=4.41, p<.01), but no IAT-effect was found when the ‘incompatible’ block came first (t(25)=0.05, p=96). Learning within blocks confirmed that the order-effect was due to cognitive inertia. Learning in the second block was always slower than learning in the first block, no matter which block came first and which second.

In study 2, we show that cognitive inertia can be so strong that it reverses the IAT-effect. We chose positive and negative stimuli such that Coca Cola was more strongly associated with either than Pepsi. The stronger association of Coca Cola with positive and negative words implied no IAT-effect at the aggregate level (neither in favor of Coca Cola nor Pepsi). However, an IAT-effect in favor of Coca Cola was found when the Coca Cola & positive words block came first (t(22)=2.91, p<.01), but an IAT-effect in favor of Pepsi was found when the Coca Cola & negative words block was administered first (t(24)=3.66, p<.01). And as in study 1, learning was always slower in the block that came second.

IAT-effects crucially depend on the selection of the positive and negative stimuli. Thus, the previous cognitive inertia effects might be a result of the positive and negative stimuli that were used.
We therefore conducted study 3 with the original positive and negative stimuli from Greenwald et al.’s (1998) study 1 and 2. The targets were coffee versus black tea. As in study 2, no IAT-effect was observed on the aggregate level. However, in accordance with cognitive inertia, an IAT-effect in favor of coffee was found when the coffee & positive words block came first (t (29)=3.34, p<.01), but an IAT-effect in favor of black tea was found when the coffee & negative words block was administered first (t (29)=2.68, p<.01). Again, learning was always slower in the block that was administered second.

Order-effects introduce so much noise in IAT-scores that they can wipe out (study 1) and reverse IAT-effects (study 2 and 3). Order-effects thus distort individual IAT-scores and render correlations with predictor variables almost meaningless. The common practice of counterbalancing block-order between-subjects cancels out order-effects on the aggregate level, but on the individual level IAT-scores are still contaminated with cognitive inertia effects.

We eliminate cognitive inertia-effects in study 4 by manipulating block-order repeatedly within- instead of between-subjects, using the same stimuli as in study 1. One group was administered the block-sequence CICIC where C denotes the block in which Coca Cola is paired with the pleasant stimuli, and I denotes the block where Coca Cola is paired with the unpleasant stimuli. The other group was administered the complementary block-sequence ICICI. When the first two blocks were used to compute IAT-effects, we found the same order-effect as in study 1, that is an IAT-effect in favor of Coca Cola when the compatible block came first (t (26)=5.33, p<.01), and no IAT-effect when the incompatible block came first (t (25)=0.46, p=.65). In contrast, when the last two blocks were used to compute IAT-effects, no difference in the IAT-effects was observed between the two block-sequences (F (1, 51)=0.67, p=.42).

Concluding, when block-order is manipulated between-subjects IAT-scores should only be interpreted on the group level. When individual IAT-scores are of interest, for example when IAT-scores are to be correlated with predictor variables, block-order should be manipulated repeatedly within-subjects.

“Looking Past the Claims: A Psychometric Analysis of the Implicit Association Test”
Hart Blanton, Texas A&M University, USA

The past decade has seen a dramatic increase in the use of implicit attitude measures, and the most popular measure by far has been the Implicit Association Test (IAT). Despite evidence of weak IAT-criterion relationships across many behavioral domains, strong claims have been made regarding the validity of this measure and its ability to improve our understanding and prediction of social behaviors. Despite the excitement surrounding this new measure, closer inspection reveals a lack of psychometric justification. To date, no formal psychometric model has been published that can justify its scoring conventions. In my talk, I examine the theory, methods and analytic strategies surrounding the IAT to determine the psychometric model that a researcher embraces (knowingly or unknowingly) whenever this measure is used.

I will begin by presenting empirical data showing that the IAT has a misspecified measurement model. The equivalent in traditional questionnaire development would occur if a researcher were to incorporate a multi-dimensional scale into his or her research program and then incorrectly interpret results as though the scale were unidimensional in nature. This type of error pervades the IAT literature.

As I show, IAT scores are influenced by three latent factors. Two of these factors assess distinct and empirically uncorrelated association strengths. The third assesses a person’s general processing speed. Despite the complex factor structure of the IAT, researchers commonly interpret even very small correlations between the IAT and psychological criteria as evidence of validity. Interestingly, however, the IAT’s previously unexamined factor structure can help explain why (1) the IAT rarely provides strong criterion prediction and (2) IAT scores are often independent of explicit self-reports.

I next show that researchers who use the IAT impose a restricted causal model into their studies, one that typically is not germane to their research questions. The causal model built into the IAT results from its construction as a measure of relative implicit attitudes. Specifically, association strengths for two attitude objects are assessed simultaneously (e.g., attitude towards smoking and attitudes towards candy). Further, this assessment is done in such a way that researchers cannot justify treating one “attitude object” as the focal object being evaluated and the other as the “scale” or “metric” against which this object is evaluated. These features require a researcher to adopt strict causal assumptions that, when not met, can cause one to unwittingly model relationships that are not of theoretical interest. As I will show, failure to consider this fact has likely resulted in faulty inferences regarding the nature of implicit attitudes and the types of influences they exert on behavioral decisions (Stuetzgen and Boatwright in this symposium present simulation results that quantify the consequences of embracing such faulty assumptions).

I present a study where we assess (1) attitudes towards apples and (2) attitudes towards oranges. We show that (1) the IAT measurement strategy quite literally mixes apples with oranges (along with general processing speed), (2) the IAT measurement strategy cannot be used to represent a person’s true “fruit attitude structure,” and (3) the IAT causal model can cause researchers to draw faulty inferences about the influence of apples and oranges in the liking of apples sauce, orange marmalade and fruit salad. Although this study was conducted somewhat tongue and cheek the implications for consumer research are quite serious.

I then review more traditional data on the IAT, including its low test-retest reliability and its responsiveness to social context and laboratory manipulations. I also explore the implications of these features for past claims that the IAT assesses stable trait constructs that influence behavioral criteria.

Finally, I turn my attention to attempts to address shortcomings by IAT researchers. IAT researchers recently have introduced a new IAT scoring algorithm, termed the “D Score” that ostensibly addresses many of the problems highlighted here. I explain how this method works and show that it not only fails to address the above but actually introduces new confounds and further undermines the measurement model.

I close by suggesting new measurement techniques that can begin to address the limitations in the IAT and that will provide marketers with a clearer guide to testing theories about consumer behavior.

“Assessing the Assumptions Behind the IAT”
Peter Stuetzgen, Carnegie Mellon University, USA
Peter Boatwright, Carnegie Mellon University, USA

Imagine you go to a restaurant with your friends. One of your friends orders Coke, but the waitress tells him that they only have Pepsi. She asks him whether he would like Pepsi instead. What do you think your friend would do? Order Pepsi or order a different soft drink? What would you do in this situation?

We asked $130$ undergraduate students enrolled in a marketing class these two questions (“What do you think your friend would do”, and “What would you do”). $88$% expected their friend to order Pepsi instead, and $84$% would also choose Pepsi them-
selves (both ratios are significantly different from 50%, chi-square >12.90, p<.001; we ran these scenarios also in the opposite direction, Coke being substituted for Pepsi. Results are similar).

While it is hardly surprising that most consumers are happy to substitute Coke with Pepsi or vice versa, consumer researchers have to assume exactly the opposite when using the IAT. The calculation of the IAT-effect requires the following relationship to hold: if Coke is someone’s favorite choice, Pepsi should be one of her/his least preferred options (cf. Blanton’s talk). In the IAT, averaging the within-block response latencies assumes a positive correlation between the two association strengths measured within one block, and subtracting the two blocks from each other assumes that the blocks are negatively correlated. Combined, this means that the more a consumer associates one of the target constructs (e.g., Coke) with one of the evaluative poles (e.g., positive), the more s/he should associate the other target construct (e.g., Pepsi) with the opposite evaluative pole (e.g., negative). In the example in the beginning, this assumption would imply that the consumer switched to a different soft drink altogether rather than substituting it with Pepsi.

This assumption on the relationship between the attitudes towards Coke and Pepsi is (1) never observable (since implicit attitudes are not directly observable) and (2) likely to be violated in marketing contexts where two products or brands are compared. In the case of Coke versus Pepsi, the more a consumer associates Coke with positive the more s/he also associates Pepsi with positive. The reason is that attitudes toward Coke and Pepsi are mainly driven by attitudes toward the product category “cola-drinks”. The more a consumer likes cola-drinks, the more s/he will like Coke, and also Pepsi. This is illustrated in the opening scenario where, even though a consumer might prefer Coke over Pepsi (or vice versa), s/he is happy to substitute one for the other. Given these two problems with this assumption, we examine via simulations its importance for the IAT’s validity.

In particular, we define 8 variables: four variables which represent the implicit association strengths to be measured in the IAT (e.g., Coke/positive, Coke/negative, Pepsi/positive, and Pepsi/negative), and another four variables that represent the IAT-measurements of these implicit attitude strengths as reaction times. We then simulate random draws from the 8x8 correlation matrix with a Gibbs sampler. By implementing constraints in the simulations of the correlation matrix we can therefore control for whether the assumption is satisfied or not. We derive a formula to calculate a measure for the IAT’s validity from these correlation matrices. This measure is simply the correlation between the IAT-effect as calculated from the observed reaction times with the IAT-effect as calculated from the simulated implicit attitudes. If the IAT was a perfect measure, this correlation should be equal to 1.

The results confirm the intuition that the IAT is on average as good a measure for the relative attitude between the two target constructs as response latencies are for each individual association strength (which are measured by the correlation between the observed reaction times and the simulated implicit attitudes).

However, the IAT assumption determines crucially whether the IAT is a better instrument than the individual response latencies, or a worse instrument. In particular, when the assumption is satisfied (i.e., the correlation between the two implicit attitudes is negative), the validity of the IAT is relatively high, even if response latencies are only a crude measure of implicit attitudes. This is because in this case the calculation of the IAT-effect combines the observed measures in a way to minimize the measurement errors contained in each measure individually. In contrast, when the assumption is violated, the validity of the IAT drops to low levels even if response latencies are very good measures of implicit attitudes. Thus, in this case interpreting the IAT-effect in the standard way is questionable.

These results suggest that the question of whether the assumption behind the IAT is satisfied or not may be more important for the validity of the IAT than the currently highly debated issue of whether response latencies are an accurate measure of implicit attitudes. Concluding, we offer advice on how to use the IAT in marketing in order to ensure its validity. In particular, whether the assumption is satisfied or not is not necessarily an inherent feature of the two target constructs to be compared, but may differ with the sample taking the IAT. For instance, in Maison et al.’s (2004) study of consumer preferences between Coke and Pepsi, the IAT performed well in predicting subsequent choice for those participants who could clearly distinguish and had clear preferences between the two brands. We conclude with the demand that IAT studies in marketing should include a discussion about how likely this crucial assumption is to be satisfied for a given sample and target categories.

REFERENCES