Cognitive Inertia and the Implicit Association Test

CLAUDE MESSNER and JOACHIM VOSGERAU

Web Appendix A

ANALYSES OF THE LEARNING CURVES

Study 1: Further Analysis of the Learning Curves

Cognitive inertia also makes a prediction about the rate of learning in each block. Because cognitive inertia occurs only in the second block when participants have to switch from applying one categorization rule to applying the opposite categorization rule, learning should be slower in the second block (indicated by a smaller rate of learning parameter r) than the first block. Consistent with this hypothesis, the rate of learning parameter r is significantly smaller in the second block (r = 1.2965, SE = .0993) than in the first block (r = 1.6697, SE =.1340; $r_{\text{first block}} - r_{\text{second block}}$ = .3732, lower 95% limit = .0463).

The strength of cognitive inertia depends on how difficult it is to switch from one categorization rule to the opposite categorization rule, which in turn depends on the association pattern between the targets (e.g., *Coca-Cola* vs. *Pepsi*) and the attributes (*pleasant* and *unpleasant* words). Switching from a compatible categorization rule (e.g., *Coca-Cola* and *pleasant* words) to an incompatible categorization rule (*Coca-Cola* and *unpleasant* words) will lead to stronger cognitive inertia than switching in the opposite direction, from an incompatible categorization rule to a compatible categorization rule. In support of this hypothesis, the learning parameter r was smaller in the incompatible block following the compatible block (r = 1.0165, SE = .0978) than in the compatible block following the incompatible block (r = 1.6701, SE = .2130; $r_{\text{compatible after incompatible}} - r_{\text{incompatible after compatible}}$ = .6536, lower 95% limit = .1942).

Study 2: Cognitive Inertia and Learning Curves

As in study 1, we included the 24 practice trials that precede each block to capture learning effects, so that each block is represented by 21 successive response latency measures consisting of the average across 8 trials each. Again, we contrast the learning curve of the first block with the learning curve of the second block. As in study 1, order-effects did not vanish over the 168 trials within each block, as the base parameter b for the learning curve of the second block is significantly higher (b = 6.5421, SE = .0077) than the base of the learning curve of the first block (b = 6.4532, SE = .0058; $b_{\text{second block}} - b_{\text{first block}}$ =.0889, lower 95% limit = .0700).

As in study 1, learning was also slower in the second block as the rate of learning parameter r was significantly smaller in the second block (r = 1.0784, SE = .0834) than in the first block (r = 1.4817, SE = .1055; $r_{\text{first block}} - r_{\text{second block}}$ = .4033, lower 95% limit = .1397). But in contrast to study 1, cognitive inertia was equally strong in both blocks as neither Coca-Cola nor Pepsi was favored by the associations with the pleasant and unpleasant attributes. Consequently, the learning parameter r in the *Coca-Cola and unpleasant words* block following the *Coca-Cola and pleasant words* block (r = .9734, SE = .1078) was not different from the learning parameter in the *Coca-Cola and pleasant words* block following the *Coca-Cola and unpleasant words* block (r = 1.1927, SE = .1292; $r_{\text{Coca-Cola and pleasant after Coca-Cola and unpleasant}} - r_{\text{Coca-Cola and unpleasant after Coca-Cola and pleasant}}$ = .2193, lower 95% limit = -.1105).

Study 3: Cognitive Inertia and Learning Curves

As in study 1 and 2, we included the 20 practice trials that preceded each block to capture learning effects. The new blocks consisted of 20 practice trials and 200 main block trials each (220 trials per block). We then divided these 220 trials per block into 22 measurements by averaging over 10 trial intervals. Again, we contrasted the learning curve for the first block with the learning curve of the second block. As in study 1 and 2, order-effects did not vanish over the 220 trials within each block, as the base parameter b for the learning curve of the second block was significantly higher (b = 6.5694, SE = .0052) than the base for the learning curve of the first block (b = 6.5153, SE = .0046; $b_{\text{second block}} - b_{\text{first block}}$ = .0541, lower 95% limit = .0405).

Also consistent with the results of study 1 and 2, learning was again slower in the second block as the rate of learning parameter r was significantly smaller in the second block (r = 1.3536, SE = .0816) than in the first block (r = 1.9394, SE = .1449; $r_{\text{first block}} - r_{\text{second block}}$ =.5858, lower 95% limit = .2599). Again as in study 2, cognitive inertia was equally strong in both blocks as the learning parameter r in the *coffee and unpleasant words* block following the *coffee and pleasant words* block (r = 1.3548, SE = .1194) was not different from the learning parameter in the *coffee and pleasant words* block following the *coffee and unpleasant words* block (r = 1.3527, SE = .1112; $r_{\text{coffee and unpleasant after coffee and pleasant}} - r_{\text{coffee and pleasant after coffee and unpleasant}}$ =.0021, lower 95% limit = -.3177).

Web Appendix B

IAT RESEARCH: BEST PRACTICE GUIDE

Table 2 in Messner and Vosgerau (2009) gives an overview of the validity and methodological issues of the IAT that have been identified in the extant literature. In the following, we discuss the implications for conducting research with the IAT. Specifically, we will argue that

1. The IAT measures associations, not necessarily attitudes
2. The IAT is sensitive to situational factors
3. The IAT cannot be used as a diagnostic test

As a consequence, we conclude that the IAT, rather than being used as a diagnostic instrument, be used as a dependent variable to measure changes in associations between target concepts and attribute stimuli. In closing, we provide a best practice guide for IAT-research.

Associations ≠ Attitudes

The IAT measures associations which are then interpreted as implicit attitudes. However, several authors (Arkes and Tetlock 2004; Karpinski and Hilton 2001; cf. also Gawronski and Bodenhausen 2006) argue that associations, whether implicit, as measured by the IAT, or explicit, do not necessarily indicate attitudes or preferences (for a response to this criticism, see Banaji, Nosek and Greenwald 2004). For example, a human rights activist and a racist might both associate poverty more with Black than White. Such associations seem to reflect general cultural knowledge rather than individual preferences.

Other findings support the notion that the IAT measures associations between the target concepts and the attribute stimuli rather than attitudes. For example, IAT-effects have been shown to depend on which attribute stimuli are used (Blümke and Friese 2006; Govan and Williams 2004; Steffens and Plewe 2001) and how the categories are labeled (De Houwer 2001). Furthermore, some researchers found IAT effects to be influenced by the familiarity of the stimuli (Brendl, Markman, and Messner 2001; Ottaway, Hayden, and Oakes 2001). Finally, researchers have pointed out that the IAT is a relative measure of associations (Brendl, Markman, and Messner 2001; Brunel, Tietje, and Greenwald 2004; Ottaway, Hayden, and Oakes 2001), and thus does not allow for any conclusion about attitude toward one target concept alone (Brendl, Markman, and Messner 2005; Nosek, Greenwald, and Banaji 2007).

Concluding, all these findings support the notion that IAT-effects, rather than indicating attitudes toward the target concepts, denote relative associations between targets and attributes.

IAT-effects are Sensitive to Situational Factors

Contrary to the concept of implicit attitudes as relatively stable over time (Wilson, Lindsey, and Schooler 2000), IAT effects are sensitive to situational influences. For example, merely thinking about negative instead of positive flowers has been shown to change an flower-IAT effect (Foroni and Mayr 2005). Again, this finding highlights the earlier conclusion that IAT-effects, rather than indicating attitudes, denote associations between target concepts and attribute stimuli. Associations are, of course, subject to situational influences. A disadvantage of this sensitivity is that the IAT is not immune to faking (Fiedler and Blümke 2005) and social desirability (Czellar 2006). However, this sensitivity also denotes a great advantage since it means the IAT can be used to measure changes in associations. For example in marketing, the effectiveness of different advertisements can be compared by measuring corresponding changes in associations with the IAT.

The IAT Cannot be Used as a Diagnostic Test

Fiedler, Messner, and Blümke (2006) argue that the IAT cannot be used as a diagnostic test (i.e., a test that measures individual differences as claimed in Greenwald, McGhee, and Schwartz 1998) because retest reliabilities (ranging from r = .16 to r = .69 bivariate correlation, Bosson, Swann, and Pennebaker 2000; Cunningham, Preacher, and Banaji 2001) fall short of the psychometric property standards for diagnostic tests.

One reason why the retest reliability of the IAT is often insufficient lies in the fact that individual IAT-effects are influenced by a host of factors other than association strengths between target concepts and attribute stimuli. These factors are category salience asymmetries (Rothermund and Wentura 2004, for a comment, see Greenwald et al. 2005 and Rothermund, Wentura, and De Houwer 2005), stimulus response compatibilities (De Houwer 2001), response criterion shifts (Brendl, Markman, and Messner 2001), method specific variance (Mierke and Klauer 2003), general processing speed (Blanton and Jaccard 2009), and cognitive inertia (Messner and Vosgerau 2009). All these factors can vary between individuals and/or as a function of block-order, and consequently distort individual IAT-scores.

Another reason why IAT-effects might lack the precision required for diagnostic tests lies in the way IAT-effects are computed. The IAT measurement model presumes that the more participants associate one target concept (e.g. Coca-Cola) with one attribute (e.g. “pleasant”), the more they associate the other target concept (e.g. Pepsi) with the other attribute (“unpleasant”, Blanton et al. 2007; Blanton et al. 2006, cf. also Nosek and Sriram 2007). Stüttgen et al. (2009)

demonstrate that if this relationship does not hold (e.g., the more people associate Coke with pleasant the more they also associate Pepsi with pleasant), IAT-effects contain mostly measurement error rendering them almost meaningless.

Finally, Blanton and Jaccard (2006a) have argued that the IAT metric is arbitrary, and a zero IAT-effect does not necessarily correspond to a neutral implicit attitude (for a reply, see Greenwald, Nosek, and Sriram 2006, and Blanton and Jaccard 2006b). Consequently, the IAT cannot be used as a diagnostic test as such tests require a meaningful reference point against which test-scores can be compared (such as an IQ of 100).

Summarizing, low retest reliabilities, individual differences in cognitive inertia and other nuisance factors, substantial amounts of measurement error, and the absence of a meaningful reference point lead to the conclusion that the IAT cannot be used as a diagnostic test.

HOW TO USE THE IAT

Interpretation of IAT-effects

IAT-effects should be interpreted as relative association strengths between the target concepts and the attribute stimuli. Interpretations of IAT-effects as implicit attitudes are unwarranted because associations do not necessarily indicate attitudes and are subject to situational influences.

IAT as Dependent Variable

Because IAT-effects are influenced by situational factors and cannot be used for measuring individual differences, the IAT should not be used as a test of implicit attitudes but as a measure of changes in relative association strengths between target concepts and attribute stimuli (Fiedler, Messner, and Blümke 2006). In other words, the IAT should be used as a dependent variable.

Stimuli Selection

The selection of the stimuli influences whether an IAT measures the intended associations. If stimuli of one evaluative dimension overlap with stimuli of one target category on an irrelevant dimension, IAT-effects no longer indicate association strengths. An example is the familiarity of stimuli. Often negative stimuli are less familiar than positive stimuli. If the stimuli of one target category are also less familiar than those of the other target category, participants can use familiarity as a shortcut. In one block they react with one key to all familiar stimuli and with the other key to all unfamiliar stimuli. If sorting for familiarity is an easier categorization rule than the intended categorization rule, the IAT no longer measures associations but differences in familiarity (Brendl, Markman, and Messner 2001). Familiarity is only one example, and unintended shortcuts could be generated by many other factors like word length, initial letter, etc. (Steffens and Plewe 2001). Therefore we advise pretesting all stimuli to assure that they do not differ on such irrelevant dimensions.

The stimuli also define the meaning of a category. For example, in a *German* vs. *Turkish* IAT, the category *German* means something different when represented by stimuli like *Adolf Hitler* than when represented by stimuli like *Claudia Schiffer* (Govan and Williams 2004; Mitchell, Nosek, and Banaji 2003). Thus, the stimuli should represent the intended target concepts as closely and unambiguously as possible. Furthermore, the attribute stimuli should represent semantically relevant dimensions (Messner and Freytag 2009). For example, when the target concepts are soft drinks (e.g., Coca Cola and Pepsi), the attribute stimuli could be "refreshing", "tingly", and "flat". When people are the target concepts, the attribute stimuli could consist of stimuli like "honorable", "friendly", "unfair", and "rude".

Cognitive Inertia and Order-effects

Cognitive inertia causes IAT-effects to depend on the block-order. 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).

Most researchers counterbalance the order of the blocks to eliminate order-effects on the aggregate level. However, cognitive inertia is still present on the individual level, and thus increases type II errors. So, when the IAT is used as a dependent variable and block-order is counterbalanced between-subjects, no differences in IAT-effects might be obtained because of the increased type II error. A better way to reduce order-effects and an increase in type II error is thus to additionally increase the number of practice trials (Nosek, Greenwald, and Banaji 2005), or, even more effective, to counterbalance block-order repeatedly within-subjects. We advise counterbalancing block-order four times within-subjects to eliminate cognitive inertia and order effects.

Interpretation of Null-effects

In an experiment where the IAT is used as a dependent variable, a null-effect can occur for three different reasons: 1) the independent variable had truly no effect on the associations measured in the IAT, 2) the IAT was too insensitive to detect differences due to too much measurement error (e.g., because of cognitive inertia and the accompanying increase in type II error, or because the target concepts were positively correlated, that is the more people associated one target concept with pleasant, the more they associated also the other target concept with pleasant (Stüttgen et al. 2009)), and 3) the manipulation of the independent variable made one target concept more ambivalent than the other target concept (i.e., it caused one target concept to be more associated with both attributes, pleasant *and* unpleasant, than the other target concept, cf., study 2 in Messner and Vosgerau 2009). This last case is only possible when order-effects are present. If they are not, null-effects can still be caused by the two former causes. Since these causes cannot be distinguished from each other, we advise not to interpret null-effects.

Conclusion

In summary, the IAT can be a valuable research tool if researchers are aware of the methodological issues inherent in this measure and are cautious in interpreting IAT-effects. We advise to use the IAT only as a dependent variable, not as a diagnostic test of association strengths. Differences in IAT-effects should be interpreted as differences in relative association strength, not as differences in implicit attitudes. Stimuli should be thoroughly pretested to avoid ambiguities in interpreting IAT-effects. Finally, when null-effects are obtained with the IAT as a dependent variable, no conclusions must be drawn (such as the independent variable had no effect on the association strengths as measured by the IAT). As with other dependent variables, null-effects can have multiple causes (such as the measure being not sensitive enough), and are thus no evidence for the absence of an effect.

References

Arkes, Hal R. and Philip E. Tetlock (2004), "Attributions of Implicit Prejudice, Or "Would Jesse Jackson 'Fail' the Implicit Association Test?"" *Psychological Inquiry*, 15 (4), 257-78.

Banaji, Mahzarin R., Brian A. Nosek, and Anthony G. Greenwald (2004), "No Place for Nostalgia in Science: A Response to Arkes and Tetlock," *Psychological Inquiry*, 15 (4), 279-310.

Blanton, Hart and James Jaccard (2006a), "Arbitrary Metrics in Psychology," *American Psychologist*, 61 (1), 27-41.

Blanton, Hart and James Jaccard (2006b), "Arbitrary Metrics Redux," *American Psychologist*, 61 (1), 62-71.

Blanton, Hart, James Jaccard, Charlene Christie, and Patricia M. Gonzales (2007), "Plausible Assumptions, Questionable Assumptions and Post Hoc Rationalizations: Will the Real IAT, Please Stand Up?," *Journal of Experimental Social Psychology*, 43 (3), 399-409.

Blanton, Hart, James Jaccard, Patricia M. Gonzales, and Charlene Christie (2006), "Decoding the Implicit Association Test: Implications for Criterion Prediction," *Journal of Experimental Social Psychology*, 42 (2), 192-212.

Blanton, Hart and James Jaccard (2009), “Modeling the Influence of General Processing Speed in Measures of Implicit Attitude: An Alternative to Scoring Algorithms,” working paper, A&M University of Texas.

Blümke, Matthias and Malte Friese (2006), "Do Features of Stimuli Influence IAT Effects?," *Journal of Experimental Social Psychology*, 42 (2), 163-76.

Bosson, Jennifer K., William B. Swann, Jr., and James W. Pennebaker (2000), "Stalking the Perfect Measure of Implicit Self-Esteem: The Blind Men and the Elephant Revisited?," *Journal of Personality & Social Psychology*, 79 (4), 631-43.

Brendl, C. Miguel, Arthur B. Markman, and Claude Messner (2001), "How Do Indirect Measures of Evaluation Work? Evaluating the Inference of Prejudice in the Implicit Association Test," *Journal of Personality & Social Psychology,* 81 (5), 760-73.

Brendl, C. Miguel, Arthur B. Markman, and Claude Messner (2005), "Indirectly Measuring Evaluations of Several Attitude Objects in Relation to a Neutral Reference Point," *Journal of Experimental Social Psychology*, 41 (4), 346-68.

Brunel, Frédéric F., Brian C. Tietje, and Anthony G. Greenwald (2004), "Is the Implicit Association Test a Valid and Valuable Measure of Implicit Consumer Social Cognition?," *Journal of Consumer Psychology*, 14 (4), 385-404.

Cunningham, William A., Kristopher J. Preacher, and Mahzarin R. Banaji (2001), "Implicit Attitude Measures: Consistency, Stability, and Convergent Validity," *Psychological Science. Special Issue*, 121 (2), 163-70.

Czellar, Sandor (2006), "Self-Presentational Effects in the Implicit Association Test," *Journal of Consumer Psychology*, 16 (1), 92-100.

De Houwer, Jan (2001), "A Structural and Process Analysis of the Implicit Association Test," *Journal of Experimental Social Psychology*, 37 (6), 443-51.

Fiedler, Klaus and Matthias Blümke (2005), "Faking the IAT: Aided and Unaided Response Control on the Implicit Association Tests," *Basic and Applied Social Psychology*, 27 (4), 307-16.

Fiedler, Klaus, Claude Messner, and Matthias Blümke (2006), "Unresolved Problems with The "I", The "A", and The "T". Logical and Psychometric Critique of the Implicit Association Test (IAT)." *European Review of Social Psychology*, 17, 74-147.

Foroni, Francesco and Ulrich Mayr (2005), "The Power of a Story: New, Automatic Associations from a Single Reading of a Short Scenario," *Psychonomic Bulletin & Review*, 12 (1), 139-44.

Gawronski, Bertram and Galen V. Bodenhausen (2006), "Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change," *Psychological Bulletin*, 132 (5), 692-731.

Govan, Cassandra L. and Kipling D. Williams (2004), "Changing the Affective Valence of the Stimulus Items Influences the IAT by Re-Defining the Category Labels," *Journal of Experimental Social Psychology*, 40 (3), 357-65.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz (1998), "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, 74 (6), 1464-80.

Greenwald, Anthony G., Brian A. Nosek, Mahzarin R. Banaji, and K. Christoph Klauer (2005), "Validity of the Salience Asymmetry Interpretation of the Implicit Association Test: Comment on Rothermund and Wentura (2004)," *Journal of Experimental Psychology: General*, 134 (3), 420-25.

Greenwald, Anthony G., Brian A. Nosek, and N. Sriram (2006), "Consequential Validity of the Implicit Association Test: Comment on Blanton and Jaccard (2006)," *American Psychologist*, 61 (1), 56-61.

Karpinski, Andrew and James L. Hilton (2001), "Attitudes and the Implicit Association Test," *Journal of Personality & Social Psychology*, 81 (5), 774-88.

Messner, Claude and Peter Freytag (2009), "Enhancing the Predictive Validity of the IAT through the Facilitation of Meaningful Comparisons," working paper, University of Basel, Switzerland.

Messner, Claude and Joachim Vosgerau (2009), "Cognitive Inertia and the Implicit Association Test," *Journal of Marketing Research*, forthcoming.

Mierke, Jan and Karl Christoph Klauer (2003), "Method-Specific Variance in the Implicit Association Test," *Journal of Personality and Social Psychology,* 85 (6), 1180-92.

Mitchell, Jason P., Brian A. Nosek, and Mahzarin R. Banaji (2003), "Contextual Variations in Implicit Evaluation," *Journal of Experimental Psychology: General*, 132 (3), 455-69.

Nosek, Brian A, Anthony G. Greenwald, and Mahzarin R. Banaji (2007), "The Implicit Association Test at Age 7: A Methodological and Conceptual Review," in *Automatic Processes in Social Thinking and Behavior*, John A. Bargh, Ed.: Psychology Press.

Nosek, Brian A., Anthony G. Greenwald, and Mahzarin R. Banaji (2005), "Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity," *Personality and Social Psychology Bulletin*, 31 (2), 166-80.

Nosek, Brian A. and N. Sriram (2007), "Faulty Assumptions: A Comment on Blanton, Jaccard, Gonzales, and Christie (2006)," *Journal of Experimental Social Psychology*, 43 (3), 393-98.

Ottaway, Scott A., Davis C. Hayden, and Mark A. Oakes (2001), "Implicit Attitudes and Racism: Effects of Word Familiarity and Frequency on the Implicit Association Test," *Social Cognition. Special Issue*, 19 (2), 97-144.

Rothermund, Klaus and Dirk Wentura (2004), "Underlying Processes in the Implicit Association Test: Dissociating Salience from Associations," *Journal of Experimental Psychology: General*, 133 (2), 139-65.

Rothermund, Klaus, Dirk Wentura, and Jan De Houwer (2005), "Validity of the Salience Asymmetry Account of the Implicit Association Test: Reply to Greenwald, Nosek, Banaji, and Klauer (2005)," *Journal of Experimental Psychology: General*, 134 (3), 426-30.

Steffens, Melanie C. and Inga Plewe (2001), "Items' Cross-Category Associations as a Confounding Factor in the Implicit Association Test," *Zeitschrift fuer Experimentelle Psychologie*, 48 (2), 123-34.

Stüttgen, Peter, Joachim Vosgerau, Peter Boatwright, and Claude Messner (2009), "Adding Significance to the Implicit Association Test," working paper, Carnegie Mellon University.

Wilson, Timothy D., Samuel Lindsey, and Tonya Y. Schooler (2000), "A Model of Dual Attitudes," *Psychological Review*, 107 (1), 101-26.