

Revised Top Ten List of Things Wrong with the IAT

Anthony G. Greenwald, University of Washington

Attitudes Preconference

SPSP - Austin, Texas

January 29, 2004

Outline

- *6. Top 10 Unsolved problems in IAT research
5. Top 10 Ways in which the IAT is used improperly in research
- *4. Top 10 Things not actually wrong with the IAT
- *3. Top 5 Excessive claims for what the IAT can do
- *2. An important development in measuring IAT effects
1. The original (Oct, 2001) Top 10 List of Things wrong with the IAT (and current status of these)

The Original (Oct, 2001) Top 10 List: Things Wrong with the IAT

6. IAT effects tend to increase with age of respondent

Current status: Solved

7. No strong rationale for standard data cleaning procedures

Current status: Solved [See Slides 5 & 6]

8. IAT effects are reduced with repeated administrations

Current status: Partially solved

9. IAT effects are smaller with picture stimuli than with word stimuli

Current status: Still a problem (unsolved)

10. Order of combined tasks influences the measure

Current status: Partially solved

The Original (Oct, 2001) Top 10 List: Things Wrong with the IAT

1. How the IAT measures association strengths is not yet well understood

Current status: Unsolved, perhaps approaching solution

2. IAT actually only measures relative strengths of pairs of associations

Current status: Partially solved

3. IAT must measure more than just association strengths

Current status: Partially solved

4. IAT appears to be slightly fakeable

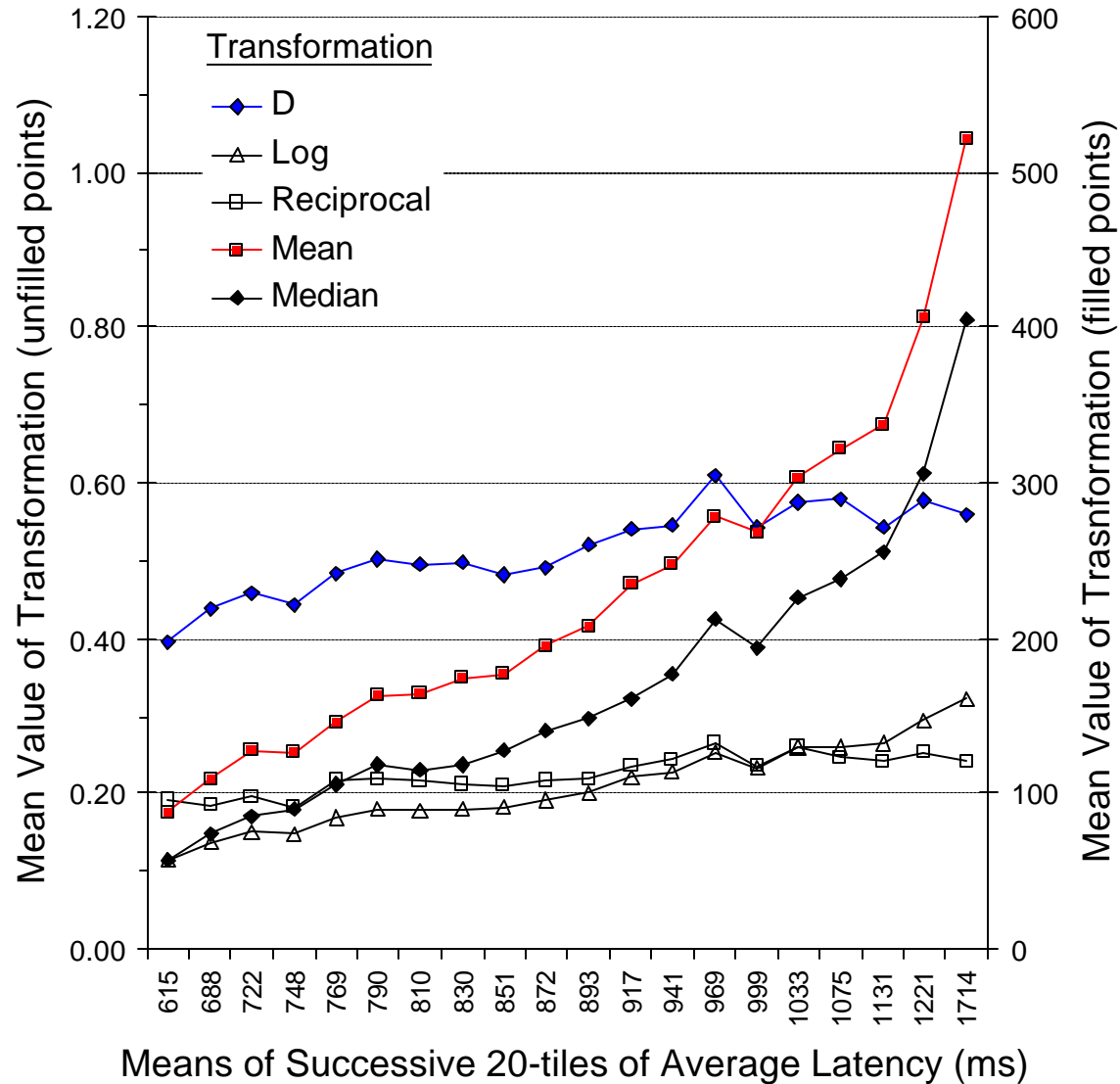
Current status: Still a problem (unsolved)

5. IAT measures are influenced by measurement context variables

Current status: Still a problem (unsolved)

“Latency Operating Characteristics” (LOCs) for IAT Scores


Election 2000 (Bush v. Gore) IAT



This slide reveals cognitive skill confound in millisecond-unit scoring of the IAT (red) and, to a lesser extent, in the log-transform measure (triangles). The new D measure (blue) is most free of this confound. The development of the new algorithm is described in the Greenwald, Nosek, & Banaji article in JPSP, Aug, 2003.

Table 4
Conventional and Improved Implicit Association Test (IAT) Scoring Algorithms Compared

Step	Conventional algorithm	Improved algorithm	Approximately equivalent alternatives for improved algorithm
1	Use data from B4 & B7	Use data from B3, B4, B6, & B7	
2	Nonsystematic elimination of subjects for excessively slow responding and/or high error rates	Eliminate trials with latencies > 10,000 ms; eliminate subjects for whom more than 10% of trials have latency less than 300 ms	
3	Drop first two trials of each block	Use all trials	
4	Recode latencies outside 300/3,000 boundaries to the nearer boundary value	No extreme-value treatment (beyond Step 2)	Delete trials with latencies below 400 ms
5		Compute mean of correct latencies for each block	Also compute <i>SD</i> of correct latencies for each block
6		Compute one pooled <i>SD</i> for all trials in B3 & B6; another for B4 & B7	Compute these pooled <i>SDs</i> just for correct responses
7		Replace each error latency with block mean (computed in Step 5) + 600 ms	Replacement = block mean + 2 × block <i>SD</i> computed in Step 5; alternately, use latency to correct response in a procedure that requires a correct response after an error
8	Log-transform the resulting values	No transformation	
9	Average the resulting values for each of the two blocks	Average the resulting values for each of the four blocks	
10	Compute the difference: B7 – B4	Compute two differences: B6 – B3 and B7 – B4	Differences can be computed in the opposite direction
11		Divide each difference by its associated pooled-trials <i>SD</i> from Step 6	
12		Average the two quotients from Step 11	

Syntax available at  <http://faculty.washington.edu/agg/>

This is a copy of Table 4 in the Greenwald, Nosek, & Banaji article in JPSP, Aug, 2003. The full article is downloadable from my home page. SPSS and SAS syntax for scoring the D measure are available as shown above.

Top 5

Excessive claims for what the IAT can do

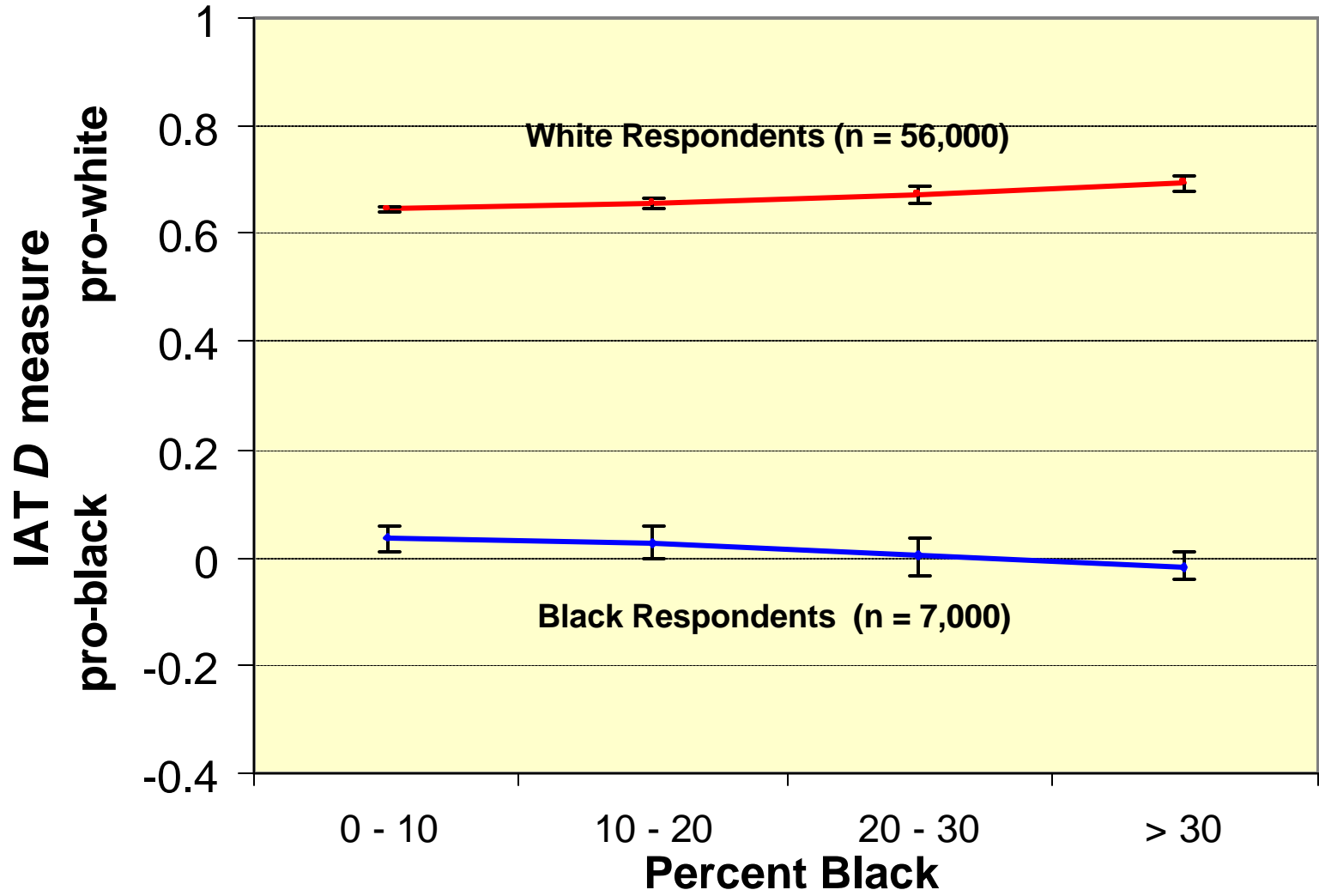
1. IAT measures can/should be used to (de-)select people for work in contexts in which intergroup biases might interfere (law enforcement, judgeship, management, jury duty)
2. The only difference between IAT and self-report is that IAT is not subject to self-presentational pressures
3. IAT measures 'true attitude'
4. IAT provides a pure measure of association strengths
5. IAT measures are unaffected by experience or measurement situation

Top 10

Things not actually wrong with the IAT

6. IAT has an arbitrary zero point (Blanton & Jaccard)
7. IAT measures cultural associations rather than personal associations (Olson & Fazio) [See Slide 9]
8. IAT measures environmental associations, rather than internalized associations (Karpinski & Hilton)
9. IAT assesses salience asymmetries (Rothermund & Wentura)
10. Positivity of IAT self-esteem indicates negativity of 'other' (Karpinski)

Implicit Bias and Percent Black displayed with 95% confidence interval



This slide shows that a standard race IAT is affected in opposite ways by an environmental variable (percentage black population in one's region). These data are from web IATs, reported in a poster at SPSP, 2004, by Jay McCauley & Greenwald.

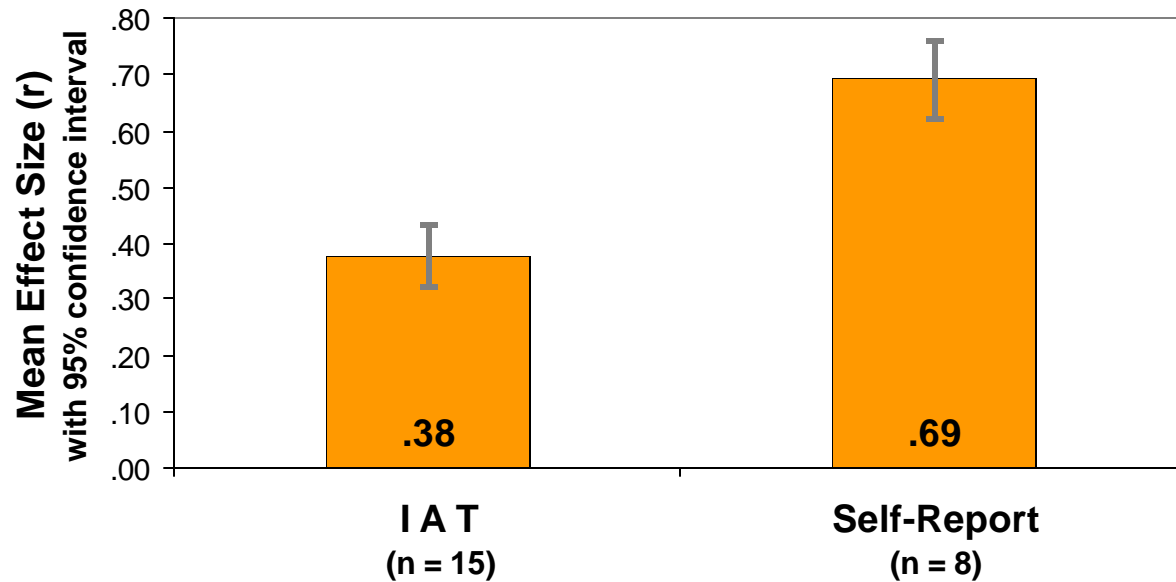
Top 10

Things not actually wrong with the IAT

1. The IAT provides a relative measure of association strengths
2. The IAT is unrelated to any interesting behavior [See Slides 11 & 12]
3. The IAT is uncorrelated with other implicit measures
4. The IAT lacks validity because it is (un)correlated with explicit measures
5. If the IAT is influenced by any non-associative factors it is an invalid measure (Blanton & Jaccard)

Predictive Validity of IAT and Self-Report: Consumer Behavior Studies

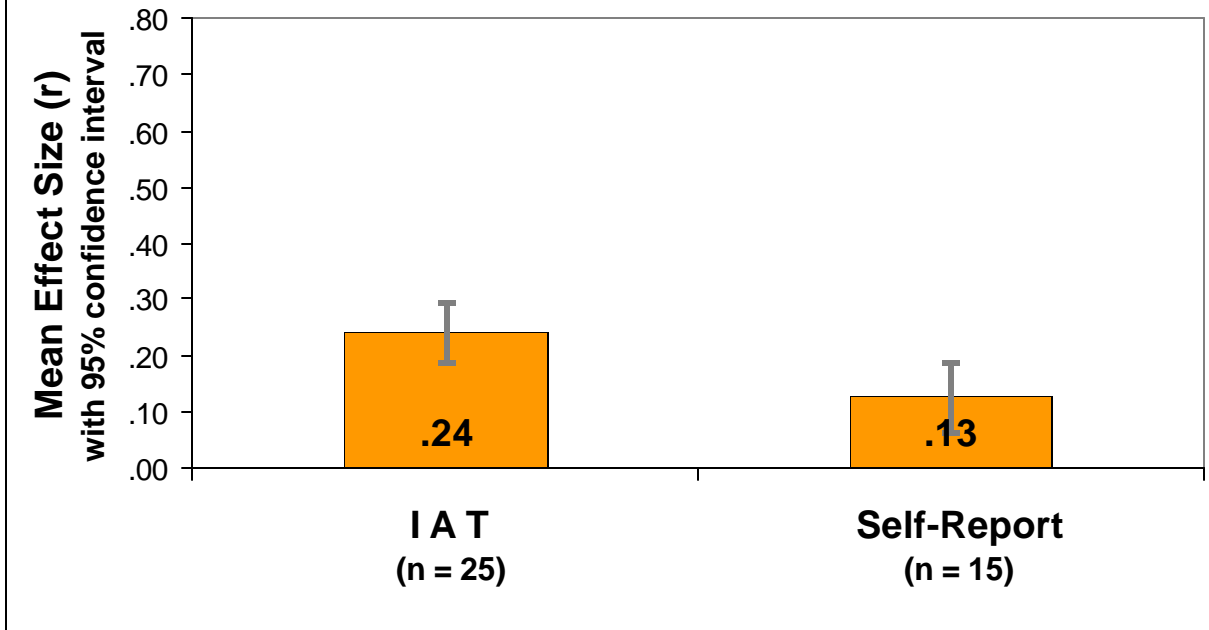
Poehlman, Uhlmann, Greenwald, & Banaji (2004)



This slide displays meta-analytic results showing substantial correlations of consumer-attitude IATs with consumer behavior. However, correlations are not as strong as with self-report measures.

Predictive Validity of IAT and Self-Report: Stereotyping and Prejudice Studies

Poehlman, Uhlmann, Greenwald, & Banaji (2004)



This slide displays meta-analytic results showing correlations of race-related attitude IATs with behavior. Here, the correlations for IATs are stronger than with self-report measures. The meta-analysis ms. is still pre-submission as of Feb. 2004.

Top 10

Ways in which the IAT is used improperly in research

6. Discarding error trials prior to data analysis
7. Making target-concept items indistinguishable in font from attribute-concept items
8. Not counterbalancing order of administration of multiple IATs when comparing magnitudes of these IAT effects
9. Randomizing the series target and attribute items rather than alternating them
10. Having subjects practice the attribute contrast before the target concept contrast

Top 10

Ways in which the IAT is used improperly in research

1. Use of non-categories (unrelated words, nonsense words) as presumably neutral categories in the IAT
2. Use of millisecond-unit IAT-effect measures (known to contain a cognitive skill artifact)
3. Treating subsets of IAT trials as measures of distinct associations
4. Using stimulus items that permit alternate interpretations of category contrasts
5. Confounding category contrasts with positive-negative valence

Top 10

Unsolved Problems in IAT Research

6. How should the IAT be used to measure implicit self-esteem?
(What do different representations of the contrast 'other' category achieve?)
7. (former #4): IAT appears to be slightly fakeable
8. (former #5): IAT measures are influenced by measurement context variables
9. (former #9): Is there a difference between properties of IATs with picture vs. word stimuli?
10. (former #8): IAT effects are reduced with repeated administrations

Top 10

Unsolved Problems in IAT Research

1. How to minimize the reactivity commonly experienced by those who take the IAT?
2. (former #1): How the IAT measures association strengths is not yet well understood
3. Can administration procedures be designed to minimize effects of the immediate research context on IAT measures?
4. How to use the IAT to measure associations involving representationally complex concepts (e.g., associations that may be at the root of health-care disparities)
5. (former #2): How should the IAT best be used to measure strengths of single associations?

CONCLUSIONS

(Jan, 2004): The IAT has benefited greatly from criticisms, even though all have not been offered in a constructive spirit.

(Oct, 2001): Nevertheless, there is room for substantial improvement in the IAT as a measure of automatic association strengths.

(Oct, 2001): The IAT is therefore presently quite useful in research on group differences, and even as a measure of individual differences.

(Oct, 2001): There is a good deal of evidence for construct validity of the IAT as a procedure for measuring automatic association strengths.