

SPECIAL REPORT

No. 196 | DECEMBER 12, 2017

The Implicit Association Test: Flawed Science Tricks Americans into Believing They Are Unconscious Racists

Althea Nagai, PhD

*The Implicit Association Test:
Flawed Science Tricks Americans into Believing
They Are Unconscious Racists*

Althea Nagai, PhD

About the Author

Althea Nagai, PhD, is Research Fellow at the Center for Equal Opportunity. The author would like to thank the Center for Equal Opportunity for its editorial support and advice throughout the publication process.

This paper, in its entirety, can be found at:
<http://report.heritage.org/sr196>

The Heritage Foundation
214 Massachusetts Avenue, NE
Washington, DC 20002
(202) 546-4400 | heritage.org

Nothing written here is to be construed as necessarily reflecting the views of The Heritage Foundation or as an attempt to aid or hinder the passage of any bill before Congress.

The Implicit Association Test: Flawed Science Tricks Americans into Believing They Are Unconscious Racists

Althea Nagai, PhD

Introduction

In 1998, University of Washington psychologist Anthony Greenwald and his colleagues developed a test that purports to uncover unconscious racism.¹ Supposedly tapping into the unconscious, the Implicit Association Test (IAT) measures disparities in millisecond response times on a computer. Based on this, Greenwald and others claim that three out of four Americans suffer from unconscious racism.²

Over the course of the past 20 years, the test has received significant media coverage in the *Washington Post*, *New York Times*, NPR, CNN, and PBS. By 2013, Greenwald and Harvard psychologist Mahzarin Banaji claimed that the “automatic White preference expressed on the Race IAT is now established as signaling discriminatory behavior.”³

But there are many scientific critics of this test, and it is far from settled science. A growing body of research suggests that the test cannot predict real-world behavior.⁴

To start, it is not clear that there are significant and reliable differences in response time, as has been asserted. When individuals take the IAT more than once, there is a good chance that results from the first and second (and subsequent) times have

very low correlations. Perhaps this is to be expected from a test measuring differences in milliseconds: One-tenth of a second can lead to highly charged accusations of racism.

Next, the difference in milliseconds can be explained by factors other than unconscious bias. There are, simply speaking, a wide variety of other explanations. Rather than unconscious racism, the test could measure the test taker’s familiarity with pairings of words and pictures. Scientists who substituted familiar versus nonsense words in place of white versus black photos or names produced the same effect as the race IAT. Some behavioral scientists suggest the race IAT measures a “figure/ground” effect, where white faces and names are the familiar and fall into the background, while black faces and names are more distinctive, thus becoming more prominent.

Some critics note that the IAT does not distinguish between cultural stereotyping, knowledge of these stereotypes, and prejudice. In a similar vein, the IAT could measure knowledge of racial disparities, which in turn could generate anger, disapproval, or dismay—not necessarily endorsement or prejudice. In some test takers, the IAT could tap into a fear of being called a racist instead of being an unconscious racist.

There are many other factors that bias the test results, including knowing the purpose of the test, faking the test results, repeatedly taking the test, being in the presence of African Americans, cognitive quickness and flexibility, physical speed, and manual dexterity.

Other social scientists have raised the serious problems related to the low level of predictive power associated with the test. The test has not been shown to significantly predict discriminatory behavior. Test results are not closely related to any other measures of discrimination: Correlations are modest at best. Even a meta-analysis by its inventors found this to be the case.

Not surprisingly, the proportion of false positives may be substantial. Estimates of false positives range from 60 percent to 90 percent.

This high probability of error has led its original proponents to conclude that it should be used with caution: “Taken together, there is substantial risk for both falsely identifying people as eventual discriminators and failing to identify people who will discriminate.”⁵ In 2015, Greenwald, Banaji, and Brian Nosek, a University of Virginia psychologist, concluded that the IAT “risk[s] undesirably high rates of false classification.”⁶

The claimed “proof” of unconscious but widespread racism can and will be used to justify any number of dubious policies. If the Implicit Association Test is used to support claims that decision makers in hiring and university admissions, housing, bank loans, and government contracting, among others, are unconsciously biased, then proponents will argue that this justifies the use of racial preferences—and even goals and quotas—to counterbalance this purported prejudice. Conversely, where such “affirmative action” is not used, or where there is any sort of racial disparity, these implicit-racism studies can be used to challenge selection decisions as discriminatory in lawsuits.⁷ These studies could be used as evidence of discrimination by law enforcement⁸ and to require minority “representation” of judges and on juries. “Unconscious bias” by teachers could be used to challenge their grading and discipline. The possibilities are endless.

Background

Since the 1950s, public opinion on race has shown a decline in racial prejudice over time, with a momentous shift in white public opinion toward the

principle of racial equality.⁹ Yet often, racial disparities in outcomes persist in income, education, home ownership, hiring, promotion, arrests and convictions, business ownership and contracting, and social mobility generally. This has led some social scientists, media commentators, and government officials to argue that there is widespread racism in our country, but it is unconscious.

Central to this movement has been an innovation in psychology that has garnered a great deal of publicity recently. Anthony Greenwald and his colleagues developed the Implicit Association Test and designed a series of experiments that purports to uncover the racism that still exists but in unconscious form.¹⁰ According to Greenwald and Banaji, 75 percent of Americans who take the IAT are found to be unconscious racists.¹¹

The IAT is an association test based on millisecond reaction time. It measures the speed with which a subject associates pleasant or unpleasant words such as “joy,” “crime,” or “work” with categories, for example, “black” or “white,” “male” or “female.” To start, Greenwald and his colleagues use the IAT to assess implicit attitudes toward socially neutral categories, such as flower versus insect, and pair these pictures with pleasant versus unpleasant words.

How the test works: Researchers instruct test takers to first hit the “positive” key when a flower appears on the computer screen and hit the “negative” key with insects, and to hit the “positive” key when pleasant words appear and hit the “negative” key with unpleasant words.

Researchers then switch the flowers/insects categories and instruct test takers to create “incompatible” pairings. Subjects are instructed to select the “positive” key when insects or pleasant words appear but select the “negative” key when flowers or unpleasant words appear.

The IAT found stronger associations, as measured by reaction speed in milliseconds, between combinations that were compatible versus those that were not. Pictures of flowers (e.g., a rose or tulip) combined with pleasant words and pictures of insects (e.g., a wasp or horsefly) combined with unpleasant words produced faster reaction times than the incompatible pairing of flowers and unpleasant words or insects and pleasant words.

From this assessment of socially neutral categories, Greenwald and his colleagues moved on to race. With this schema, they instructed test takers to hit

the “positive” and “negative” keys, creating patterns of associations as follows: For the “compatible”¹² set of pairings, test takers were instructed to pick the “positive” key when white pictures and pleasant words appeared, and to hit the “negative” key when black pictures and unpleasant words appeared. For the “incompatible” set of pairings, test takers were told to hit the “positive” key when black pictures and pleasant words appeared and to hit the “negative” key when white pictures and unpleasant words were flashed on the screen.

These combinations produced differential response times. On average, the “compatible” pairings generated faster reaction times than the “incompatible” combinations. This difference in millisecond reaction time is what led researchers to posit the existence of unconscious racism that caused the test taker to favor white over black when mixed with pleasant over unpleasant words, while taking longer when the pairings resulted in black over white when mixed with pleasant over unpleasant words.

After several years of IAT research, Greenwald, Banaji, and Nosek founded Project Implicit, a website for IAT researchers, consultants, and organizations interested in using the test and for individuals who want to take the test. Millions have accessed the test online.¹³

Major media outlets such as the *Washington Post*, *New York Times*, and CNN have profiled the IAT, with such eye-catching titles as, “Across America, Whites are Biased and Don’t Even Know It” and “What? Me, Biased?”¹⁴ It was the focus of a popular book by Banaji and Greenwald,¹⁵ featured prominently in Malcolm Gladwell’s 2005 bestseller, *Blink*, and in a 2015 film on PBS described thus: “*American Denial* sheds light on the unconscious political and moral world of modern Americans,” including “research footage, websites, and YouTube films showing psychological testing of racial attitudes.”¹⁶

The concept of implicit or unconscious bias and the use of the IAT to root it out have worked their way into public policy and our legal system. There have been suggestions to incorporate IAT technology into judicial nominations and jury selection.¹⁷ One author proposes looking for implicit bias in legislative action, advocating the use of IAT to “smoke out’ illegitimate purposes” and hidden racists among legislators, showing that race-neutral classifications, for example, tap into unconscious race bias.¹⁸ In a 2012 class action suit against the State of Iowa,

African American state employees claimed class-wide bias in hiring and promotion based on disparate impact statistics and implicit racial bias. Expert witness testimony on unconscious racism was central to their claims. The case became the first site for dueling experts on the scientific status of implicit racism. Anthony Greenwald was the expert witness for the plaintiffs, and Philip Tetlock, a psychologist at the University of Pennsylvania, was the expert for the state of Iowa. Ultimately, the judge rejected the implicit bias theory and ruled for the State of Iowa.¹⁹ The state supreme court unanimously upheld the trial judge’s ruling.²⁰

In the field of criminology, the U.S. Department of Justice has had a program of implicit bias and community policing since 2009. In light of recent events concerning race and police behavior, police departments around the country have held conferences, training sessions, and exercises to deal with the issue of unconscious racial bias among law enforcement.²¹ There is little scientific evaluation with proper design, sampling, comparison groups, controls, and statistical analysis showing that they work. Short-term effects have been shown, but results seem to dissipate over time, and, according to one critic, may actually make unconscious bias worse. In addition, it could endanger police officers by causing them to misread real threats and significantly delay reactions for fear of unconscious racism.²²

The IAT could also be used to analyze how college and university admission committees evaluate applicants, how faculty and teaching assistants grade students, how faculty hire and promote their own, and as an assessment in who studies, teaches, and practices law. Medical schools are actively moving in that direction. Prompted by the American Association of Medical College’s concern with diversity and unconscious bias, medical schools such as Stanford, Ohio State, and Johns Hopkins encourage faculty and students to take the IAT, declaring the test to be both reliable and valid, ignoring its controversy in psychology and related social sciences. Duke University has gone one step further and incorporated it into a second-year medical school course on unconscious bias.²³

In short, the search for unconscious racism has the potential for widespread educational, media, and judicial “bias training.”²⁴ Moreover, UCLA Law School professor Jerry Kang and Mahzarin Banaji advocate for permanent affirmative action. Given

the extent of unconscious racism, they argue, affirmative action should be disbanded only when unconscious racism disappears nationwide: “Fair measures that are race- or gender-conscious will become presumptively unnecessary when the nation’s implicit bias against those social categories goes to zero or its negligible behavioral equivalent.”²⁵

In the psychological and social sciences, however, there is consensus on neither unconscious racism nor the IAT. Many of the controversies focus on technical issues—measurement, validity, and reliability, to name a few. But it is precisely this technical debate that makes the study of unconscious bias and the IAT far from settled science. The strategy of its proponents is to ignore the critics or accuse them of being narrow-minded. Banaji claims that the IAT is to psychology what Galileo’s telescope was to the Copernican Revolution, also drawing an analogy of IAT research to the Copernican and Darwinian scientific revolutions. Banaji acknowledges that this would-be scientific revolution “is going to be the hardest [to accept] of all.”²⁶

The IAT findings are threatening, for the studies move us away from the familiar and comfortable. The findings undercut how we see ourselves as thoughtful beings with the free will to be moral and good. Banaji explains:

[It] will challenge our beliefs about the very nature of our own minds.... [I]t is not merely about the place of our planet amongst other planets, [sic] it’s not merely about our place in the larger set of other species, [sic] it’s about the core issue of our competence, [sic] it’s about our goodness, our ability to be moral, and to have control over our thoughts and feelings, about the most important object in our universe, other humans.²⁷

But the tide is turning for the IAT. As recently as 2012, other scholars, including University of Virginia law professors Allan King and Gregory Mitchell, point out that social science findings related to unconscious racism and the IAT are “contested research.... *This research is the subject of vigorous debate within psychology....* [E]xperts citing IAT research often *mischaracterize the findings* from this body of work and omit important limitations on the research” (emphasis added).²⁸

Before the IAT becomes entrenched in public policy and the law, its proponents should address

questions about the reliability and validity of the test. The test should be shown to predict other behaviors, and there should be a broader discussion of the social and political implications of this research.

Flaw Number One: The IAT Is Unreliable

One serious criticism of the IAT has to do with its unreliability. “Reliability” refers to the consistency of a measure—that is, the extent to which repeated applications of a measuring instrument result in roughly the same outcomes.

No measuring instrument is perfectly reliable (i.e., guaranteeing absolutely identical results time after time), but some measures are better than others. Established measuring instruments of physical traits, such as a ruler for height or a thermometer for temperature, are generally less prone to reliability issues compared to instruments in the social sciences.

The American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education have jointly published standards regarding testing.²⁹ The associations state that the reliability of a test centers on the notion that an individual’s performance is somewhat the same from one test-taking time to another. Estimating reliability should involve calculating a correlation between a test and its retest for many test takers. The test/retest correlation should be large (e.g., over 0.70) if the test is reliable.

The professional associations recognize that an individual’s scores on the same test may vary from one time to another. In the aggregate, group scores reflect some degree of measurement error—the degree to which the scores vary from the true score. In the view of these associations, however, “if a test score leads to a decision that is not easily reversed, such as rejection or admission of a candidate to a professional school or the decision by a jury that a serious injury was sustained, *the need for a high degree of precision is much greater*” (emphasis added).³⁰ In other words, the test/retest reliability should yield in the aggregate a coefficient of 0.90 or higher.³¹

Because IAT proponents argue that the IAT taps into racism on the unconscious level, and since racism is such a highly charged accusation, the IAT should be subjected to a high degree of precision. But it is not. As Texas A&M and Florida International

University psychologists Hart Blanton and James Jaccard, respectively, observe, the IAT has serious problems of test/retest reliability. The IAT measures reaction time to specific stimuli in milliseconds. Using a micro-measure of reaction time with regard to differences in unconscious racial attitudes is relatively new, according to Blanton and Jaccard, and consequently has significant problems associated with it: “[A] tenth of a second can have a consequential effect on a person’s score, and such measurement sensitivity can lead to test unreliability.”³²

According to Blanton and Jaccard, the conventionally acceptable correlation for test/retest reliability is a correlation coefficient of 0.70 and rises to 0.90 when used for individual assessment.³³ They find that Greenwald’s test/retest reliability is 0.56,³⁴ while another group of researchers found a test plus three re-tests over a two-week period caused correlation coefficients to plummet to 0.27.³⁵ Clearly, the reliability of the IAT is problematic, since the test itself has not changed since the 1990s.

Flaw Number Two: Validity—What Does the IAT Actually Measure?

Aside from its unreliability, unconscious racism and the IAT have other problems. Assume, for the sake of argument, that over time improvements have led to significant IAT test/retest reliability. Reliability is still not the same as validity. Something can be “reliable” in the technical sense of yielding similar results over time yet still not be a valid measure. Validity is a fundamental concern in science: To what extent does the object we study in fact represent the object we want to study? Do our empirical comparisons truly reflect our theoretical concepts? Are we measuring what we think we are measuring?

The number on the oven thermometer is a valid measure of the hotness of the oven, and the number on a *pH*-scale is a valid measure of the acidity-alkalinity of the soil. Astrological signs, however, are not valid measures of individuals’ personality traits.

Proponents of the IAT brag that they have millions of scores generated from their website, Project Implicit. The number of individuals taking the IAT does not address the issue of a flawed test and flawed results. There are several types of validity, and psychologists do not agree on the categories, but there is a consensus among critics regarding the IAT’s validity.

The first concern centers on construct validity, which deals with whether the measure used in fact measures what it claims to measure. That is, is the IAT a valid measure of the concept, of “unconscious racism”? IAT proponents claim that it is. In order to show that it is a valid measure of the concept, alternative explanations of the differential reaction time when faced with “white” cues or “black” cues must be ruled out.

The key question of construct validity is whether the IAT scores measure unconscious racism or something else. While alternative explanations regarding the meaning of IAT scores are either not considered at all by IAT proponents or are casually dismissed by them, published research shows that, in fact, other social-psychological processes can explain IAT scores. There are several factors that contribute to the results, including:

- Comparing familiar versus unfamiliar words, pictures, and associations;
- Knowledge of stereotypes, instead of prejudice or cultural stereotyping;
- Knowledge of racial disparities and sympathy toward African Americans for this reason;
- The fear of being labeled racist; and
- Physiological and physical factors, e.g., intelligence, physical speed, and manual dexterity.

Familiar Versus Unfamiliar Words, Pictures, and Associations

Raising the criticism of construct validity, Miguel Brendl, Arthur Markman, and Claude Messner designed IAT experiments that suggest alternative explanations to unconscious racism.³⁶ While Greenwald and his colleagues argued that the longer response times of the “incompatible” pairings of black pictures and pleasant words versus white pictures and unpleasant words tap into unconscious prejudice, Brendl, Markman, and Messner proposed that the IAT registers “familiar” versus “unfamiliar” sets of associations. The more common associations result in faster reaction times; the more distinctive or less common, the slower times.

Brendl and his colleagues used insects and non-sense syllables, then paired them with pleasant and

unpleasant words. When they used insects versus nonsense syllables in their experiment, people were quicker to associate insects with pleasant words and nonsense syllables with unpleasant ones. Yet, can one say that insects such as cockroaches and wasps have pleasant associations?

By using nonsense syllables and insects, they got the same kind of results as did Greenwald and his colleagues with the white–black IAT. They suggest that IAT responses are a function of familiarity and unfamiliarity, especially since the differences are measured in milliseconds. The “compatible” sets of associations in Greenwald’s model are basically more familiar sets of associations compared to the “incompatible” or unfamiliar sets. Yet proponents ignore these alternative interpretations.

Brendl and his colleagues’ familiar–unfamiliar explanation is similar to alternative hypotheses proposed by psychologists Klaus Rothermund of the University of Trier in Trier, Germany, and Dirk Wentura of the University of Jena in Jena, Germany. They base their alternative explanations on what is often called figure–ground issues in the psychology of perception (or “salience asymmetries”).³⁷

The design of the IAT is such that the test taker is instructed to quickly sort the photos (e.g., white faces, black faces) with highly charged pleasant and unpleasant adjectives (e.g., happy, evil, good, poor), and at the same time, must sort the faces and words with the positive key and the negative key. Then, the test taker is instructed to switch, so that the picture of the black person’s face is associated with the positive attributes and the picture of the white person’s face is paired with the negative. To now undertake this second set of instructions, the brain must first sort the prompts (i.e., is it a white face, a black face, a pleasant word, an unpleasant word) and then re-orient the manual task of pushing the correct button.

The mental tasks of re-focusing and re-orienting may very well account for the disparities in black–white IAT results—disparities that are measured in milliseconds.

The strength of Rothermund and Wentura’s explanation lies in its placement within the scientific paradigm of perceptual psychology. They criticize IAT proponents for ignoring how people process salience asymmetry. The IAT, as “a new experimental paradigm” of social cognition, should confront this alternative asymmetry explanation.

In short, the familiar/unfamiliar, figure–ground explanations provide a powerful paradigm

alternative to the IAT proponents. At best, these alternative explanations would dampen the magnitude of race-IAT disparities when properly taken into account. Just as likely, however, they highlight a more fundamental challenge to the dominant unconscious racism interpretation of the race IAT.

Racial Stereotyping, Racial Prejudice, or Knowledge of Stereotypes

The IAT is a test about race—and no one, after all, wants to be labeled a racist. Test takers know exactly what the correct response is supposed to be when photos of black and white faces are systematically associated with heavily charged positive or negative words. University of Colorado Boulder psychologists Charles Judd, Irene Blair, and Kristine Chapleau argue that the IAT taps into cultural stereotypes to which test takers had been exposed, rather than unconscious racism.³⁸ The IAT proponents conflate automatic stereotyping with automatic prejudice, particularly in highly controversial areas such as police behavior. The researchers state, “If it is the implicit activation of these stereotypes that is responsible for racially biased policing, then it seems to us that rather different interventions are called for than those that would be most appropriate if the problem was due to...highly prejudiced officers.”³⁹ Others say the IAT does not differentiate between holding cultural stereotypes and being aware of these stereotypes—and holding a stereotype would lead to treating people differently than merely being aware of stereotypes.

Knowledge of Racial Disparities, Not Unconscious Racism

In a similar vein, Gregory Mitchell and Philip Tetlock criticize proponents of the IAT for conflating knowledge of racial disparities with approval of racial disparities. According to Mitchell and Tetlock, if a test taker is aware of the following three propositions, then, by IAT proponents’ definition, he or she is an unconscious racist: (1) There are racial disparities in America; (2) A majority of Americans notice these disparities; and (3) Negative feelings have come to be associated with these widely noticed disparities.⁴⁰ Based on these premises, Mitchell and Tetlock argue that “the more people know about the past and present history of American race relations, and about current patterns of inequality, the worse they should score on the IAT.”⁴¹

Fear of Being Called a Racist, Not Being Unconsciously Racist

The fear of being called racist is yet another possible alternative explanation. Some white test takers would see the results as “confirming” their personal fears that they were racists, deep down.

In one experiment, a group of researchers led by Oberlin College psychologist Cynthia Frantz looked at whites threatened by the possibility of appearing racist and those who were not.⁴² Test takers who were more worried about appearing racist had worse IAT results. The psychologists conclude, “Ironically the IAT appears to be the most threatening to people who most want to appear nonracist⁴³—which is the opposite of what the test is supposed to pick up.

These findings led Mitchell and Tetlock to conclude that, for some, the IAT measures “*sympathy, not antipathy*” for the condition of blacks [emphasis added].⁴⁴ Tetlock and Ohio State psychologist Hal Arkes observe how the race IAT is unable to distinguish between test takers who feel a sense of injustice regarding race in America versus test takers who feel prejudice and hatred towards blacks.

The studies described above are critical to the issue of whether the IAT is a measure of unconscious racism or something else. There are, however, other validity issues that should also be addressed before concluding that the science supports the idea that unconscious racism causes the disparities in IAT results.

Physical and Physiological Factors Affecting the Validity of the IAT

There are many extraneous variables that affect an IAT score. For one, the results of the IAT can be faked. In one study, test takers who deliberately slowed their reaction time to the “compatible” set of pairings (white-positive/black-negative) obtained a significantly smaller response-time difference between the “compatible” versus the “incompatible” associations (black-positive/white-negative).⁴⁵ In another study, test takers were trained to increase their speed in the experimental pairings, also leading to a less valid score.⁴⁶ Whether slowing the expected pairings or speeding up the experimental ones, this is evidence that results on the IAT can be faked.

Repeatedly taking the IAT also reduces the disparities between the “compatible” and “incompatible” pairings. The creators of the IAT do point out

that repeated trials result in better scores and note that the scores of those taking the test for the first time cannot be compared to those who have taken the test more than once. The problem, as Greenwald and his colleagues note, is that prior experience automatically raises post-test scores: “The effect of prior experience means that scores of IAT novices cannot be compared directly with those of non-novices and, for the same reason, posttests cannot be compared directly with pretests (when the pretest is the first IAT taken)... [N]umerically less extreme IAT scores will be observed for those with prior IAT experience” (emphasis added).⁴⁷

Another external factor that improves scores is being in the presence of African Americans. When test takers are shown images of prominent African Americans such as Denzel Washington and Michael Jordan, race IAT scores improve immediately after.⁴⁸ In another study, race IAT disparities were reduced after the test taker was first assigned to a group comprised of whites and blacks.⁴⁹

Yet another condition affecting scores has to do with taking the race IAT in one’s native tongue. Bilingual individuals taking the race IAT get significantly different scores, depending on the language. One study showed that when bilingual test takers took the race IAT in their native language (e.g., Spanish), scores were significantly worse than when taking it in their non-native language (English).⁵⁰

Cognitive speed and cognitive flexibility have also been shown to affect IAT response time. Intelligence research finds that intelligence speeds performance on simple tasks. The more intelligent subjects had significantly faster reaction times than the less intelligent.⁵¹ IAT experiments that do not control for intelligence likely overestimate prejudice.

Finally, physical speed and flexibility affect IAT response time. The IAT requires task-switching, as when the “compatible” sets of pairings (white and positive words/black and negative words) are switched to the “unconventional” sets of black and positive words/white and negative words. Those who have greater physical speed and manual flexibility, e.g., younger adults, do better on the test.

Flaw Number Three: Do IAT Scores Predict Racist Behavior?

The discussion on validity has covered some of the alternative explanations for IAT results and the many factors compromising scores. All these explanations

and variables raise serious questions about the overall validity of the IAT as tapping into unconscious racism. Some have become more cautious and willing to acknowledge the malleability of IAT scores. A study by University of Massachusetts at Amherst psychologist Nilanjana Dasgupta allows for the manipulability of IAT scores but notes that other race-linked attitudes, postures, and interactions “are also quite malleable depending on the extent to which awareness, control, and motivation are at play.”⁵²

Given that IAT results may be affected by some or all of the many factors described above, how much does the race IAT correlate with other measures of prejudice or discrimination? The critics say, “Not much.”

Scientific theories have to address the issue of “predictive validity,” which deals with the association between the independent variable (e.g., SAT score, hiring-exam score, IAT results) and the dependent variable (e.g., college admission, college GPA, job-performance evaluation). What little existing research there is has found less-than-robust correlations between the IAT and other measures, including other controversial measures of racial prejudice and discrimination—“symbolic racism” and race-based “microaggressions.”

IAT Correlations with Symbolic Racism Attitudes

One way IAT researchers conduct validity studies is to correlate IAT scores and scores on a “symbolic racism attitude” scale. Symbolic racism attitudes are the positions on policy issues such as affirmative action (on which the anti-affirmative action response is considered an indicator of symbolic racism). This means that holding conservative beliefs, by definition, makes you racist. Symbolic racism proponents argue that certain attitudes toward affirmative action, work, unemployment, and welfare, among other issues, are actually racist attitudes masked as policy positions.

But in their review of such studies, Mitchell and Tetlock find that the “median result of studies of the implicit-explicit linkage yield estimates of low positive correlations between measures.”⁵³ In other words, the linkage between unconscious racism and symbolic racism is weak, and results correlating IAT scores and other measures are mixed.

Of course, another problem is that symbolic racism attitudes are methodologically suspect. Regarding

symbolic racism and other related studies, or what Stony Brook University political scientists Leonie Huddy and Stanley Feldman call “the new racism,” there is significant controversy within political science regarding the new racism’s construct validity, measurement validity, and predictive validity—the same scientific controversy surrounding the race IAT. Most noted is the criticism and research by Paul Sniderman and his colleagues, who argue that the “new racism” is confounded by conservative ideology, insofar as many issues used as indicators of “new racism” use the language of ideological individualism.⁵⁴ Huddy and Feldman, in their own study, found that conservatives opposed race-conscious scholarship programs, whether the programs favored blacks or whites. Huddy and Feldman conclude, “Racial resentment, therefore, is not a clear-cut measure of racial prejudice for all Americans and may convey ideological principles for conservatives.”⁵⁵

In sum, both IAT and symbolic racism measures are problematic, and the two do not even correlate well with one another.

Correlating IAT Scores with Micro-Behaviors

The results relating race IAT scores with “micro-behaviors” are mixed. One study examined differences in subjects and actions when interacting with a white experimenter or a black experimenter. Subjects were coded according to their degree of friendliness, comfort level, eye contact, and body posture among other things. McConnell and Leibold videotaped subjects’ responses when interacting with black experimenters and then with white experimenters and examined corresponding IAT scores. Significant correlations were found between the IAT score and the nature of interaction with white-versus-black experimenters.⁵⁶

Another study, however, produced confounding results. After face-to-face contact, black subjects awarded more positive interaction scores to whites with “more racist” IAT scores. The black subjects awarded more negative interaction scores to whites with better IAT scores (i.e., whites who were less “racist”).⁵⁷

Other studies found that higher race IAT scores were only slightly correlated with greater social discomfort and anxiety in contact with blacks and other groups. In her review of the controversies in IAT research, University of Georgia sociologist Justine

Tinkler observes that implicit measures are basically associated with behaviors “that are spontaneous and difficult to control.”⁵⁸ These include various test taker micro-behaviors—including eye contact, eye blinks, posture, smiling, speaking time, and speech errors—as part of the interaction between test taker and white/black experimenter, or white/black persons who were just in the room.⁵⁹

Mitchell and Tetlock criticize the interpretation of the micro-behaviors such as looking down or away, halting speech, not speaking, and posture as indicators of racism. They note that these behaviors are also indicators of shame—which, in turn, can be a function of unfamiliarity, uncertainty, fear of being labeled racist, or shame of societal treatment of blacks, among other factors⁶⁰—not micro-behavioral manifestations of racism.

Even a major meta-analysis of the IAT and other variables, conducted by Greenwald and his colleagues, on 188 studies with 184 independent samples and 14,900 subjects regarding the predictive validity of the IAT could not find large correlations.⁶¹ Their meta-analysis involved not just the black-white IAT but the IAT used for many other groups.⁶² They found an average correlation of 0.27 between different types of the IAT and various performance, judgment, and physical measures. For black-white race IAT studies, correlation coefficients averaged around 0.24.⁶³ In a later meta-analysis, Rice University psychologist Frederick L. Oswald and his colleagues found even lower average correlations of roughly 0.15.⁶⁴

Greenwald, Banaji, and Nosek in 2015 point out that the differences in methodology between the studies account for the difference in correlations, but they concede the point that the correlations between IAT results and other behaviors are small. But they do state in their refutation, “Statistically small effects...can have societally large effects.”⁶⁵ Except they offer no proof.

The problem of weak correlations was found as far back as 2003. Psychologists Russell Fazio of Ohio State University and Michael Olson of the University of Tennessee concluded in their 2003 review of implicit measures, “One of the most disturbing trends to emerge in the literature on implicit measures is the many reports of disappointingly low correlations among the measures.... Unquestionably, part of the problem with these disappointing correlations among various implicit measures is

their rather low reliability.”⁶⁶ Earlier they state, “In contrast to the numerous investigations concerning known-group differences, less work has been conducted concerning the prediction of behavior from IAT scores. The evidence that does exist is mixed.”⁶⁷

Not much has changed. In the 2009 *Annual Review of Political Science*, noted public opinion researchers and political scientists Leonie Huddy and Stanley Feldman examined the work on the IAT, and concluded that, while interesting, “the results of implicit racial attitudes can be confusing.”⁶⁸ They, too, note the often-contradictory results between implicit and explicit attitudes.

Huddy and Feldman argue that explicit racial attitude questions, not the results of an unconscious racism test, should suffice to uncover racism, especially when there is time for a respondent to think about policies or a politician (e.g., feelings toward President Obama).⁶⁹ Huddy and Feldman state: “[C]ontinuing disputes in psychology over the meaning of implicit attitudes serve as a cautionary note to political scientists interested in incorporating such measures into their research.”⁷⁰

Along similar lines, the lack of consistent and robust correlation between IAT results and other measures led social scientist Justine Tinkler in 2012 to also conclude that it is wrong to discount explicit attitudes and focus only on implicit attitudes.

[I]t is not accurate to interpret explicit attitudes as politically correct and dishonest and implicit attitudes as true attitudes.... [W]ith evidence that implicit and explicit attitudes affect race-related behavior in different ways, it would be a mistake to dismiss egalitarian, non-racist explicit attitudes as dishonest because they reflect social desirability.⁷¹

The weakness of the race IAT in terms of predictive validity brings us to the last methodological issue—the problem of the false positive.

Flaw Number Four: False Positives Are False Accusations of Racism

The rate of “false positives” and “false negatives” is critical in evaluating the truth-value and utility of any test. A false positive is a result when the condition is detected but is not really there. Everyday examples include the false alarm for home security systems, where the security system indicates

an intruder but is in reality the family cat. Another example from medicine is the presence or absence of cancer. One test (e.g., PSA test) indicates possible prostate cancer, but further testing (e.g., a biopsy) turns up negative. This is an incidence of a false positive regarding the first test, where a different test was subsequently used for assessment. Less common is the discussion of false negatives. This is where the test indicates the absence of a condition (e.g., cancer), but the condition is really there.

The first indicator that the IAT may produce many false positives is that the correlations of IAT scores with other indicators of racism are low. Low correlations indicate an insensitive test, and therefore a high level of false positives. In other words, many who are labeled racists by the IAT are in fact not. Mitchell and Tetlock's analyses find the false-positive rate for the IAT falls between 60 percent and 90 percent, depending on the study. Since we do not know the true state of one's character (whether a person is truly racist or not), we have to use other measures (e.g., behavioral or attitudinal responses) as substitutes to discover false positives and false negatives. Mitchell and Tetlock use survey attitudes, where at most 30 percent appear to give a "racist" response on an attitude question. The table below shows the more widespread "racist response" finding that this author took from Huddy and Feldman's "The American Racial Opinion Survey" (AROS).⁷²

Answers in the affirmative ("great deal," "some," or "little") for items five and six were used as indicators by Huddy and Feldman of overt racism. Roughly 40 percent thought that differences in standardized tests were due to racial differences in intelligence; 35 percent thought the black-white difference was a "fundamental genetic difference between the races."⁷³

For the analysis below, this report will use the 40 percent of overtly racist responses, for the sake of argument, plus assuming that 75 percent have a high enough IAT score to be unconscious racists. This 40 percent gives us the largest percentage of true positives and the smallest percentage of false positives regarding the IAT. If we have 1,000 individuals taking both the AROS survey and the IAT, let us assume the following: 400 would give the racist survey response; 750 of those taking the IAT would get a racist IAT; and the IAT would pick up 90 percent of those truly racist (i.e., racist IAT score and racist survey response).

The false positive rate is the non-racist attitude as a percentage of the racist IAT score. Since there are 390 persons with a non-racist response on the survey question but a racist score on the IAT, if we calculate the false-positive rate, we get a false positive of 52 percent.⁷⁴ This means that 52 percent of those with a racist IAT score are not racist.

Conversely, the false negatives are those persons who are truly racists according to the survey item but are not detected by the test. There are 40 of these persons, which gives us a false-negative rate of 16 percent.⁷⁵ In other words, the IAT would fail to detect roughly 16 percent of the truly racist.

Are a 52 percent false-positive rate and a 16 percent false-negative rate acceptable?⁷⁶ This is not a scientific question, but an ethical and policy one. In the real world, the IAT was suggested for screening students and faculty, job hiring and promotion, and jury selection, to name a few instances. Over half those taking the race IAT will be falsely accused, while 16 percent of the truly racist will slip by.

The error rate for the IAT is sufficiently high that even one of the founders of Project Implicit admitted, "Taken together, there is substantial risk for both falsely identifying people as eventual discriminators and failing to identify people who will discriminate."⁷⁷

Flaw Number Five: IAT Results Cannot Be Generalized to the Real World

To what extent can the race IAT findings generated from a sample of undergraduates in a college laboratory be generalized to the real world? IAT proponents occasionally acknowledge the problems of moving from scholastic research to the public policy and legal arenas. The meta-analysis conducted by Greenwald and his colleagues excluded the Internet-based IAT findings on the Harvard-sponsored Project Implicit website because the latter's test conditions are unreliable and not valid. But proponents use the Web-based numbers to give it the appearance of scientific authority.⁷⁸

Even the creators of the IAT, Greenwald and Banaji, acknowledge in their book, *Blindspot*, that the website-based findings of Project Implicit cannot be generalized to the American public.⁷⁹ In 2015, in a technical paper, Greenwald, Banaji, and Nosek concede that the scientific issues associated with the IAT mean that the test should not be used for individual assessment.⁸⁰

TABLE 1

Poll Question Designed to Reveal Racism

Q: On average, African-American students get lower scores on standardized tests than do whites. How much of a difference in test scores (Great Deal, Some, Little, or None) is due to the following:

- 1) Occurs because more blacks do not have the chance to get a good education
- 2) Can be explained by discrimination against blacks
- 3) Occurs because most blacks just don't have the motivation or will power to perform well
- 4) Occurs because most blacks do not teach their children the values and skills which are required to be successful in school
- 5) Is due to racial difference in intelligence
- 6) Occurs because of fundamental genetic difference between the races

SOURCE: Leonie Huddy and Stanley Feldman, "On Assessing the Political Effects of Racial Prejudice," *Annual Review of Political Science*, Vol. 12 (2009).

SR196  heritage.org

Elsewhere, Brian Nosek and Rachel Riskind, assistant professor of psychology at Guilford College in Greensboro, North Carolina, agree with critics that the IAT generates high false-positive rates and therefore should not be used for individual diagnostic purposes.

Although the reviewed evidence shows that this stereotype is true in the aggregate, in our view, applying this to individual cases disregards too much uncertainty in measurement and predictive validity.... [T]he present sensitivity of implicit measures do not justify this application.⁸¹

As a concrete example, Mitchell and Tetlock present a list of workplace best-practice conditions that make generalizing from IAT research to the American workplace dubious.⁸² Significant differences between the lab setting and workplace include the following:

1. Only race is considered in the lab, while supervisors look at employees' backgrounds, work histories, and past performance.
2. The test taker has no knowledge of the experimenter's views, while organizations typically publish an official view on discrimination, prejudice, affirmative action, and diversity. There is no close future contact between experimenter and test taker, and negative responses have no future consequences, while there is often future

interaction among employees and between employees and supervisors. Negative interactions between supervisors and employees have consequences.

3. Test takers do not expect to work with the experimenter in a team, and negative IAT lab results have no future teamwork consequences, while organizations expect teamwork between supervisors and employees.
4. Test takers are usually college students and lack experience in supervising others. Workplace managers and human resource personnel have experience in hiring and managing others.

For these many differences, Mitchell and Tetlock conclude that the IAT should not be used for workplace evaluations. In the workplace, there would be hiring and promotion consequences as well as legal and financial ramifications if an employee or manager is said to be an unconscious racist based on the IAT. The same kind of analysis applies to such fields as teaching, criminal justice, and health care.

Conclusion and Implications

Proponents of the IAT have not shown that the test unequivocally measures unconscious racism and have failed to rule out alternative explanations. Likewise, the IAT has not been shown to correlate with other established measures of prejudice and discrimination, and little research shows it

predicting discriminatory behavior. There are high rates of false positives and false negatives associated with the test. The IAT has not been shown to apply to real-world settings.

Notwithstanding these problems, IAT proponents seek its widespread adoption in public policy and legal arenas. Ignoring the differences between scholarly research and real-world implementation, big institutions have hired private consultants and instituted proprietary programs to correct, train, and generally root out unconscious racism and other forms of bias.

The American Association of Medical Colleges (AAMC) makes unconscious bias a focus of medical school curriculum and medical practice reform.⁸³ Despite the current scholarly consensus that the IAT should not be used for individual diagnostics, the AAMC encourages its use to discover the extent of an individual's unconscious biases, and various medical schools do the same. Prominent medical schools feature the IAT on their websites and encourage people to take the test.⁸⁴

In the AAMC's forum on diversity and inclusion, the IAT is referenced throughout, starting with the AAMC making the highly disputed claim, "The IAT has been rigorously tested for reliability, validity, and predictive validity and has been shown to be a methodologically sound instrument for measuring unconscious associations."⁸⁵

At Ohio State University, members of the medical school admission committee took the IAT in order to screen for their "implicit white preference," followed by a presentation on unconscious bias and strategies for its reduction, before screening applicants.⁸⁶

A medical school, of all places, should know all about the proper administering of tests. While citing the work of Greenwald, Banaji, and others, these elite medical schools and the AAMC seemed to have missed the part about the limitations of the IAT—that it should not be used for individual diagnostics, such as assessing the unconscious bias of individual committee members before actually screening candidates.⁸⁷

At UCLA, administration officials encouraged all to test themselves for unconscious racism, speaking of "a series of troubling racial climate incidents."⁸⁸ The university has a Vice Chancellor for Equity, Diversity, and Inclusion. The office is currently held by Jerry Kang, Professor of Law and Asian American Studies, who has written numerous articles and

guides on implicit-bias studies and the law, including a recent lecture on the implications of implicit-bias research and the Equal Protection Clause.⁸⁹ He is assisted by "diversity prevention officers," responsible for investigating bias among faculty and to ameliorate the situation by holding re-education and training sessions, pointing to the IAT as a resource. In 2017, UCLA required all members of faculty search committees to undergo training in spotting implicit bias, including mandatory viewing of UCLA's full Implicit Bias Video Series before starting the search.⁹⁰

In the corporate world, according to the *Wall Street Journal*, roughly 20 percent of large corporations currently provide some sort of unconscious bias training—a figure estimated to grow to 50 percent by 2020, which has given rise to consultants and technology to stamp out unconscious bias.⁹¹

Policy experience has shown, however, that even the best-designed and executed studies and programs (e.g., FDA studies, randomized reading research in education) can have non-generalizable results and lead to unintended consequences. Public policy application of social science findings must be significantly more demanding, analogous to FDA drug studies—and even more so when there are legal consequences. There are no scientifically based evaluation studies showing that the programs implemented to correct unconscious bias are reliable, valid, and effective and that subjecting people to these programs yields the desired real-world outcomes.

Yet many are not shy about proposing how the IAT could be used. After all, what are the societal consequences for promulgating a less-than-rigorous scientific theory of prejudice? Should the arguments of reliability and validity be limited to the concerns of methodologists and psychologists?

Researchers inserting themselves into the policy and legal arenas on behalf of social intervention is a problem, and, as Nosek and Riskind observe, many experimental scientists fail to appreciate the complexity of the policy process. Unlike the university science lab, the real world is full of intervening variables. In the eyes of the public, it further politicizes the social sciences and further erodes the disciplines' credibility.

Moreover, by emphasizing the finding that unconscious racism is spread throughout society, even among its most "enlightened" elite, the race

IAT research may sadly serve to make race relations worse. False accusations of racism are highly likely, and true instances of racism lose their salience. The real difficulty is the public cynicism and indifference that results when accusations are made, new policies are implemented, and millions of dollars are spent on the problem—with little perceived progress. Ultimately, unconscious racism, cultural stereotyping, stereotype threat—or whatever is actually measured by the IAT—is regularly overcome in everyday life.

Given the high probability of errors associated with the IAT, it should not be incorporated into public policies, such as hiring and university admissions, housing, banking, and government contracting, by law enforcement, in lawsuits, or in jury selection. Although it has been hailed by the media as uncovering a dark, secret side of the American psyche, numerous critics of the IAT have demonstrated that it simply cannot predict how test takers will act in the real world. The test fails to prove that we are a nation of unconscious racists.

—Althea Nagai, PhD, is Research Fellow at the Center for Equal Opportunity. The author would like to thank the Center for Equal Opportunity for its editorial support and advice throughout the publication process.

Endnotes

1. See, e.g., Anthony G. Greenwald, Debbie E. McGhee, and Jordan K. L. Schwartz, "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, Vol. 74 (1998), pp. 1464-1480; Anthony G. Greenwald and Mahzarin R. Banaji, "Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes," *Psychological Review*, Vol. 102 (1995), pp. 4-27; and Brian A. Nosek, Anthony G. Greenwald, and Mahzarin R. Banaji, "The Implicit Association Test at Age 7: A Methodological and Conceptual Review," in John A. Bargh, ed., *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (New York: Psychology Press, 2007), pp. 265-292.
2. See Mahzarin R. Banaji and Anthony G. Greenwald, *Blindspot* (New York: Delacorte Press, 2013), p. 47. The percentage is based on the number of people taking the race-IAT on their website. But they also note that it is also not a representative sample of the American public. Banaji and Greenwald, *Blindspot*, p. 221.
3. *Ibid.*, p. 47.
4. After this author had completed the manuscript and it was under review and editing, it was brought to her attention that two articles in the popular press make similar points and come to the same conclusions. See Jesse Singal, "Psychology's Favorite Tool for Measuring Racism Isn't Up to the Job," *NYMag*, January 11, 2017, <http://nymag.com/scienceofus/2017/01/psychologys-racism-measuring-tool-isnt-up-to-the-job.html> (accessed August 16, 2017), and Tom Bartlett, "Can We Really Measure Implicit Bias? Maybe Not," *Chronicle of Higher Education*, January 5, 2017.
5. Brian A. Nosek and Rachel G. Riskind, "Policy Implications of Implicit Social Cognition," *Social Issues and Policy Review*, Vol. 6, No. 1 (2012), p. 133.
6. Anthony Greenwald, Mahzarin R. Banaji, and Brian A. Nosek. "Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects," *Journal of Personality and Social Psychology*, Vol. 108, No. 4 (2015), pp. 553-561. The quote is from page 14 of their final draft and is available at https://faculty.washington.edu/agg/pdf/GB&N.Consequential%20small%20IAT%20effects.JPSP_final.2Sep2014.pdf (accessed August 16, 2017).
7. In a lengthy class action trial involving the State of Iowa, African American state employees claimed class-wide bias in hiring and promotion based on disparate impact statistics and expert-witness testimony of unconscious racism as central to their claims. The judge rejected the implicit bias theory and ruled for the State of Iowa. *Pippen v. Iowa*, No. CL10738 (Iowa Dist. Ct. Polk Co., April 17, 2012). See also David Kadue, Gerald L. Maatman, Jr., Jennifer Riley, and David Ross, "Iowa State Court Rejects Theory of Unconscious Bias and Disparate Impact Class Claims in Bellweather Ruling," *Workplace Class Action Blog*, April 19, 2012, <http://www.workplaceclassaction.com/2012/04/iowa-state-court-rejects-theory-of-unconscious-bias-and-disparate-impact-class-claims-in-bellwether/> (accessed June 15, 2015).
8. The issue of unconscious racial bias and police behavior has come up in the context of the recent police shootings. Many departments have implemented implicit-bias training, but the program effects have not been scientifically shown. See, e.g., Todd Essig, "Unconscious Racial Bias: From Ferguson to the NBA to You," *Forbes*, September 21, 2014, <http://www.forbes.com/sites/toddesig/2014/09/21/racism-from-ferguson-to-the-nba-to-you/> (accessed June 30, 2015); Chris Mooney, "The Science of Why Cops Shoot Young Black Men," *Mother Jones*, December 1, 2014, <http://www.motherjones.com/politics/2014/11/science-of-racism-prejudice> (accessed June 30, 2015); "Police Agencies Line Up to Learn About Implicit Bias," *Associated Press*, March 9, 2015, <http://newsok.com/article/feed/808890> (accessed August 25, 2017); and Martin Kaste, "Police Debate the Effectiveness of Anti-Bias Training," *NPR*, April 6, 2015, <http://www.npr.org/2015/04/06/397891177/police-officers-debate-effectiveness-of-anti-bias-training> (accessed June 30, 2015).
9. For example, based on the input of many quantitative sociologists and political scientists, the General Social Survey (GSS) has asked a large national sample the same set of questions about American society and politics since 1972, including questions of race. In 1972, the GSS found that roughly 15 percent of white respondents thought that black and white children should go to separate public schools. In the early 1980s, fewer than one in 10 whites felt the same. In 1985, there were so few whites believing in segregated public schools that the GSS dropped the question. On the question of intermarriage, 65 percent of whites opposed black-white marriages of a relative or family member in 1990. In 2008, that percentage dropped to roughly 25 percent. Other data present a more mixed picture. Yet, those who conduct survey analysis of racial attitudes over time generally see a decline in overt prejudice against blacks. See John Wihbey, "White Racial Attitudes Over Time: Data from the General Social Survey," *Journalist's Resource: Research on Today's News Topics*, August 14, 2014, <https://journalistsresource.org/studies/society/race-society/white-racial-attitudes-over-time-data-general-social-survey> (accessed October 10, 2017); and Justine Tinker, "Controversies in Implicit Race Bias Research," *Sociology Compass*, Vol. 6, No. 12 (November 14, 2012), pp. 987-997, <http://onlinelibrary.wiley.com/doi/10.1111/soc4.12001/full> (accessed August 28, 2017).
10. See, e.g., Greenwald et al., "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," pp. 1464-1480; Greenwald and Banaji, "Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes," pp. 4-27; and Nosek et al., "The Implicit Association Test at Age 7," pp. 265-292.
11. See Banaji and Greenwald, *Blindspot*, p. 47. The percentage is calculated from those taking the race-IAT at the website. They also note elsewhere that it is also not a representative sample of the American public. See disclaimer at note 2, *supra*.
12. The researchers labeled these as "compatible" and "incompatible" sets of pairings.
13. "Implicit Social Cognition: Investigating the Gap Between Intentions and Actions," Project Implicit, 2011, <http://projectimplicit.net/index.html> (accessed March 6, 2015). Despite its huge size, the number taking the test on the website is not scientifically relevant, because the "sample" is not scientifically drawn and cannot be generalized.

14. Chris Mooney, "Across America, Whites Are Biased and Don't Even Know It," *Washington Post*, December 8, 2014, <http://www.washingtonpost.com/blogs/wonkblog/wp/2014/12/08/across-america-whites-are-biased-and-they-dont-even-know-it/> (accessed May 4, 2015); Nicholas Kristof, "What, Me Biased?" *New York Times*, October 30, 2008, <http://www.nytimes.com/2008/10/30/opinion/30kristof.html> (accessed May 4, 2015); Nicholas Kristof, "Is Everybody a Little Bit Racist?" *New York Times*, August 27, 2014, <http://www.nytimes.com/2014/08/28/opinion/nicholas-kristof-is-everyone-a-little-bit-racist.html> (accessed May 4, 2015); and Elizabeth Landau, "You May Be More Racist Than You Think," *CNN*, April 2, 2009, <http://www.cnn.com/2009/HEALTH/01/07/racism.study/index.html?iref=24hours> (accessed May 4, 2015).
15. Banaji and Greenwald, *Blindspot*.
16. "American Denial," PBS, January 29, 2015, <http://www.pbs.org/independentlens/american-denial/> (accessed May 3, 2015).
17. Anthony Page, "Batson's Blind-Spot: Unconscious Stereotyping and the Peremptory Challenge," *Boston University Law Review*, Vol. 85, No. 155 (2005), cited in Gregory Mitchell and Philip Tetlock, "Antidiscrimination Law and the Perils of Mindreading," *Ohio State Law Journal*, Vol. 67 (2006), p. 1027.
18. Reshma M. Saujani, "The Implicit Association Test: A Measure of Unconscious Racism in Legislative Decision-Making," *Michigan Journal of Race and Law*, Vol. 8 (2003), pp. 395 and 396.
19. Kadue et al., "Iowa State Court Rejects Theory of Unconscious Bias, and *Pippen v. Iowa*."
20. *Pippen v. Iowa*.
21. See "Reducing Biased Policing Through Training," *Community Policing Dispatch*, Vol. 2, No. 2 (February 2009), http://cops.usdoj.gov/html/dispatch/February_2009/biased_policing.htm (accessed June 30, 2015); Todd Essig, "Unconscious Racial Bias: From Ferguson to the NBA to You," *Forbes*, September 21, 2014; Chris Mooney, "The Science of Why Cops Shoot Young Black Men"; Tami Abdollah, "Police Agencies Line Up to Learn About Implicit Bias," March 9, 2015, <http://www.scp.org/news/2015/03/09/50259/police-agencies-line-up-to-learn-about-unconscious/> (accessed August 16, 2017); and Martin Kaste, "Police Debate the Effectiveness of Anti-Bias Training."
22. Kirsten Weir, "Policing in Black and White," *Monitor on Psychology*, Vol. 47, No. 11 (December 2016), <http://www.apa.org/monitor/2016/12/cover-policing.aspx> (accessed August 22, 2017).
23. Ohio State University, "Proceedings of the Diversity and Inclusion Innovation Forum: Unconscious Bias in Academic Medicine," <http://kirwaninstitute.osu.edu/my-product/proceedings-of-the-diversity-and-inclusion-innovation-forum-unconscious-bias-in-academic-medicine/> (accessed June 30, 2017); American Association of Medical Colleges, "E-Learning Seminar: What You Don't Know: The Science of Unconscious Bias and What To Do About it in the Search and Recruitment Process," 2015, https://www.aamc.org/members/leadership/catalog/178420/unconscious_bias.html (accessed August 16, 2017); Johns Hopkins Medicine, "Diversity and Cultural Competence: Implicit Association Test," http://www.hopkinsmedicine.org/odcc/implicit_association_test.html (accessed August 16, 2017); and Stanford University School of Medicine, "Information About Bias: Test Yourself," 2017, <https://med.stanford.edu/facultydiversity/diversity-resources/information-about-bias.html> (accessed August 16, 2017).
24. See especially work by Jerry Kang as part of his website, "Getting Up to Speed on Implicit Bias," October 12, 2016, <http://jerrykang.net/2011/03/13/getting-up-to-speed-on-implicit-bias/> (accessed August 16, 2017).
25. Jerry Kang and Mahzarin R. Banaji, "Fair Measures: A Behavioral Realist Revision of 'Affirmative Action,'" *California Law Review*, Vol. 94, No. 4 (2006), p. 1116. See also Christine Jolls and Cass R. Sunstein, "The Law of Implicit Bias," *California Law Review* (2006), pp. 969-996; Kristin A. Lane, Jerry Kang, and Mahzarin R. Banaji, "Implicit Social Cognition and Law," *Annual Review of Law and Social Sciences*, Vol. 3 (2007), pp. 427-451.
26. Quoted in Jill D. Kessler, "A Revolution in Social Psychology," *APS Observer Online*, Vol. 14, No. 6 (July/August 2001), <http://www.psychologicalscience.org/observer/0701/family.html> (accessed January 6, 2014).
27. *Ibid.*
28. Allan G. King, Jeffrey S. Klein, and Gregory Mitchell, "Effective Use and Presentation of Social Science Evidence," *Employee Relations Law Journal*, Vol. 37, No. 4 (Spring 2012), p. 6. See also Maurice Wexler, Kate Bogard, Julie Totten, and Lauri Damrell, "Implicit Bias Evidence and Employment Law: A Voyage into the Unknown," *Bloomberg Law*, July 8, 2013, <http://about.bloomberglaw.com/practitioner-contributions/implicit-bias-evidence-and-employment-law-a-voyage-into-the-unknown/> (accessed January 19, 2014).
29. American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, *Standards for Educational and Psychological Testing* (Washington, DC: American Educational Research Association, 1999), pp. 25-30.
30. *Ibid.*, pp. 29-30.
31. For example, test-retest reliability of the SAT is a correlation of roughly 0.90-0.91.
32. Hart Blanton and James Jaccard, "Unconscious Racism: A Concept in Pursuit of a Measure," *Annual Review of Sociology* (2008), p. 289.
33. *Ibid.*, pp. 289 and 290.
34. Anthony G. Greenwald, Brian A. Nosek, and Mahzarin R. Banaji, "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm," *Journal of Personality and Social Psychology*, Vol. 85, No. 2 (2003), pp. 197-216.
35. William A. Cunningham, Kristopher J. Preacher, and Mahzarin R. Banaji, "Implicit Attitude Measures: Consistency, Stability, and Convergent Validity," *Psychological Science*, Vol. 12, No. 2 (2001), pp. 163-170.

**THE IMPLICIT ASSOCIATION TEST:
FLAWED SCIENCE TRICKS AMERICANS INTO BELIEVING THEY ARE UNCONSCIOUS RACISTS**

36. C. Miguel Brendl, Arthur B. Markman, and Claude Messner, "How Do Indirect Measures of Evaluation Work? Evaluating the Inference of Prejudice in the Implicit Association Test," *Journal of Personality and Social Psychology*, Vol. 81, No. 5 (2001), pp. 760-773.
37. Klaus Rothermund and Dirk Wentura, "Underlying Processes in the Implicit Association Test: Dissociating Salience from Associations," *Journal of Experimental Psychology: General*, Vol. 133, No. 2 (2004), pp. 139-165.
38. Charles M. Judd, Irene V. Blair, and Kristine M. Chapleau, "Automatic Stereotypes vs. Automatic Prejudice: Sorting Out the Possibilities in the Weapon Paradigm," *Journal of Experimental Social Psychology*, Vol. 40, No. 1 (2004), pp. 75-81.
39. *Ibid.*, p. 81.
40. Gregory Mitchell and Philip E. Tetlock, "Antidiscrimination Law and the Perils of Mindreading," *Ohio State Law Journal*, Vol. 67 (2006), pp. 1023-1121, especially pp. 1085-1089.
41. *Ibid.*, p. 1086.
42. Cynthia M. Frantz, Amy J. C. Cuddy, Molly Burnett, Heidi Ray, and Allen Hart, "A Threat in the Computer: The Race Implicit Association Test as a Stereotype Threat Experience," *Personality and Social Psychology Bulletin*, Vol. 30, No. 12 (2004), pp. 1611-1124.
43. *Ibid.*, p. 1622.
44. Mitchell and Tetlock, "Antidiscrimination Law," pp. 1080-1083.
45. The tactic, however, was not discovered independently by test takers. They had to be told to do so. See Do-Yeong Kim, "Voluntary Controllability of the Implicit Association Test (IAT)," *Social Psychology Quarterly* (2003), pp. 83-96.
46. Xiaoqing Hu and J. Peter Rosenfeld, "Combining the P300-Complex Trial-Based Concealed Information Test and the Reaction Time-Based Autobiographical Implicit Association Test in Conceal Memory Detection," *Psychophysiology*, Vol. 49, No. 8 (2012), pp. 1090-1100.
47. Greenwald et al., "Understanding and Using the Implicit Association Test," p. 211.
48. Nilanjana Dasgupta and Anthony G. Greenwald, "On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and Disliked Individuals," *Journal of Personality and Social Psychology*, Vol. 81, No. 5 (2001), p. 800.
49. Jay J. Van Bavel and William A. Cunningham, "Self-Categorization with a Novel Mixed-Race Group Moderates Automatic Social and Racial Biases," *Personality and Social Psychology Bulletin*, Vol. 35, No. 3 (2009), pp. 321-335.
50. Oludamini Ogunnaike, Yarrow Dunham, and Mahzarin R. Banaji, "The Language of Implicit Preferences," *Journal of Experimental Social Psychology*, Vol. 46, No. 6 (2010), pp. 999-1003.
51. Rul von Stülpnagel and Melanie C. Steffens. "Prejudiced or Just Smart?: Intelligence as a Confounding Factor in the IAT Effect," *Zeitschrift für Psychologie/Journal of Psychology*, Vol. 218, No. 1 (2010), p. 52.
52. Dasgupta, Nilanjana, "Implicit Ingroup Favoritism, Outgroup Favoritism, and Their Behavioral Manifestations," *Social Justice Research*, Vol. 17, No. 2 (2004), pp. 157 and 158.
53. Mitchell and Tetlock, "Antidiscrimination Law," pp. 1062-1063.
54. Paul M. Sniderman, and Philip E. Tetlock, "Symbolic Racism: Problems of Motive Attribution in Political Analysis," *Journal of Social Issues*, Vol. 42, No. 2 (1986), pp. 129-150; Paul M. Sniderman, Thomas Piazza, Philip E. Tetlock, and Ann Kendrick, "The New Racism," *American Journal of Political Science* (1991), pp. 423-447; and Paul M. Sniderman and Thomas Piazza, *The Scar of Race* (Cambridge, MA: Harvard University Press, 1993).
55. Summaries can be found in Leonie Huddy and Stanley Feldman, "On Assessing the Political Effects of Racial Prejudice," *Annual Review of Political Science*, Vol. 12 (2009), p. 426.
56. Allen R. McConnell and Jill M. Leibold, "Relations Among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes," *Journal of Experimental Social Psychology*, Vol. 37, No. 5 (2001), pp. 435-442.
57. J. Nicole Shelton, Jennifer A. Richeson, Jessica Salvatore, and Sophie Trawalter, "Ironic Effects of Racial Bias During Interracial Interactions," *Psychological Science*, Vol. 16, No. 5 (2005), pp. 397-402.
58. Justine Tinkler, "Controversies in Implicit Race Bias Research," *Sociology Compass*, Vol. 6, No. 12 (2012), p. 992.
59. Mitchell and Tetlock, "Antidiscrimination Law," p. 1095, footnotes 229 and 230.
60. For references regarding shame, see *ibid.*, pp. 1096 and 1097, footnotes 231-233.
61. A meta-analysis is a statistical technique that aggregates results from multiple studies for the purpose of achieving greater statistical power compared to a single experiment. One inherent bias of meta-analyses is what researchers call "the file-drawer problem," where only published studies are included, thereby overlooking unpublished studies with findings that failed to support the hypothesis studied.
62. Greenwald et al., "Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity," *Journal of Personality and Social Psychology*, Vol. 97, No. 1 (2009), pp. 17-41.
63. The correlation coefficient, r , is a measure of the strength of association between variables, ranging from -1.00 to 1.00, with 0.00 representing no relationship. A correlation coefficient of -1.00 indicates a perfect negative relationship: When one variable increases, the other decreases. Conversely, a correlation coefficient of 1.00 indicates a perfect linear relationship, where the increase in the value of one variable predicts perfectly the increase in the value of the other. While there is no rule regarding judgments of correlation size (weak, moderate, or strong), Cohen defines them as follows: low=0.10; moderate=0.30; high=0.50. See Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd Ed. (New York: Lawrence Erlbaum, 1988), pp. 78-81.

64. Frederick L. Oswald, Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E. Tetlock, "Using the IAT to Predict Ethnic and Racial Discrimination: Small Effect Sizes of Unknown Societal Significance," *Journal of Personality and Social Psychology*, Vol. 108, No. 4 (2015), pp. 562-571.
65. Greenwald, Banaji, and Nosek, "Statistically Small Effects." See also Anthony Greenwald, Mahzarin R. Banaji, and Brian A. Nosek, "Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies," *Journal of Personality and Social Psychology*, Vol. 105, No. 2 (2013), pp. 171-192.
66. Russell H. Fazio and Michael A. Olson, "Implicit Measures in Social Cognition Research: Their Meaning and Use," *Annual Review of Psychology*, Vol. 54 (2003), p. 311. See also Michael A. Olson and Russell H. Fazio, "Relations Between Implicit Measures of Prejudice: What Are We Measuring?" *Psychological Science*, Vol. 14, No. 6 (November 2003), pp. 636-639. In their experiment, Olson and Fazio found that subjects who were encouraged to classify persons by race had attitudes that correlated with higher IAT scores. Those who were not previously made to classify based on race showed no correlation between attitudes and IAT score.
67. Fazio and Olson, "Implicit Measures in Social Cognition Research," p. 308.
68. Huddy and Feldman, "On Assessing the Political Effects," p. 436.
69. *Ibid.*, pp. 436 and 437.
70. *Ibid.*, p. 437.
71. Tinkler, "Controversies in Implicit Race Bias Research," p. 993.
72. Huddy and Feldman, "On Assessing the Political Effects," p. 428.
73. Roughly 4 percent thought that a great deal of the black-white test score gap was due to racial differences in intelligence; another 20 percent thought it accounted for some of the gap, while 16 percent said it accounted for a little of the gap. More than half (53 percent), however, said it accounted for none of the gap, and another 6 percent did not know. On the question of genetics and the black-white test score gap, 3 percent said a great deal of the gap was due to "fundamental genetic differences between the races." Roughly 18 percent said genetics accounted for some of the black-white test score gap, while 14 percent said genetic differences accounted for a little of the black-white gap. Fifty-eight percent, however, said that genetics accounted for none of the gap, while 7 percent did not know. See Huddy and Feldman, "On Assessing the Political Effects," p. 428.
74. Rate of false positives=Non-racist attitude/IAT racist score=390/750=52 percent.
75. Rate of false negatives=Racist attitude/non-racist IAT=40/250=16 percent.
76. If we assume that the racist survey response is 34 percent (i.e., the percentage agreeing that the black-white test score gap is fundamentally genetic), the rate of false positives is 59 percent and the rate of false negatives is 14 percent.
77. Brian A. Nosek and Rachel G. Riskind, "Policy Implications of Implicit Social Cognition," *Social Issues and Policy Review*, Vol. 6, No. 1 (2012), p. 133.
78. An increasing number of sociological and political science survey research has been conducted using Internet-based designs. They are useful from a coverage and cost perspective. In the past, surveys used random samples based on home phone numbers and conducted either by phone, mail, or face-to-face. The challenge is drawing the sample, starting with properly specifying the universe from which the sample is to be drawn. With the decline of the landline telephone in households and an increase in refusals to participate, obtaining an appropriate sample is becoming significantly harder and more expensive. In response, researchers have turned to the Internet, but with caution. Credible survey research then adds demographic weight to the sample, in order to reflect the universe from which the sample is drawn and to represent and statistically control for potentially confounding (demographic) factors. If the sample is to represent the American public, adjustments are often made to account for variables such as gender, income, race, age, and education. Ultimately, it makes no difference how many take the Web version of the IAT—no adjustments, no use.
79. Banaji and Greenwald, *Blindspot*, p. 221.
80. Greenwald, Banaji, and Nosek. "Statistically Small Effects." After all this time, Greenwald and his colleagues finally state, "Attempts to diagnostically use such measures for individuals risk undesirably high rates of erroneous classifications." The quote is from a draft dated September 2, 2014, p. 12, and can be found at https://faculty.washington.edu/agg/pdf/GB&N.Consequential%20small%20IAT%20effects.JPSP_final.2Sep2014.pdf (accessed August 17, 2017).
81. Nosek and Riskind, "Policy Implications of Implicit Social Cognition," p. 133.
82. Mitchell and Tetlock, "Antidiscrimination Law," pp. 1109-1110.
83. Ohio State University, "Proceedings of the Diversity and Inclusion Innovation Forum: Unconscious Bias in Academic Medicine," and American Association of Medical Colleges, "E-Learning Seminar: The Science of Unconscious Bias."
84. Johns Hopkins Medicine, "Diversity and Cultural Competence: Implicit Association Test," and Stanford University School of Medicine, Office of Faculty Development and Diversity, "Information about Bias: Test Yourself."
85. Ohio State University, "Proceedings," p. 8.
86. *Ibid.*, p. 20.
87. *Ibid.*

**THE IMPLICIT ASSOCIATION TEST:
FLAWED SCIENCE TRICKS AMERICANS INTO BELIEVING THEY ARE UNCONSCIOUS RACISTS**

88. See Heather MacDonald, "The Microaggression Farce," *City Journal* (Autumn 2014), http://www.city-journal.org/2014/24_4_racial-microaggression.html (accessed June 19, 2015) for a detailed account of events at UCLA.
89. UCLA, "Vice Chancellor for Equity, Diversity, and Inclusion," <https://equity.ucla.edu/about-us/our-teams/vice-chancellor/> (accessed August 17, 2017); and Kang, "Getting Up to Speed on Implicit Bias." The site includes links to his work, including a video as vice chancellor, a guide on implicit bias research for the state courts, and articles on the implications of unconscious racism findings on judges and juries and the impact of implicit bias research on the equal protection clause, among others.
90. UCLA Equity, Diversity and Inclusion, "Faculty Search Committee Resources," <https://equity.ucla.edu/programs-resources/faculty-search-process/faculty-search-committee-resources/> (accessed August 23, 2017).
91. Textio software company sells company software that will detect hidden biases in their hiring ads and performance reviews. Similarly, Giang describes Kanjoya's software as picking up "the earliest signs of bullying, harassment, and discrimination" through subtleties in language. There is also Unitive.works, which uncovers biases in real time in the recruitment and hiring process—writing the job description, scanning resumes, and conducting interviews, for example. See Vivian Giang, "The Growing Business of Detecting Unconscious Bias," <http://www.fastcompany.com/3045899/hit-the-ground-running/the-growing-business-of-detecting-unconscious-bias> (accessed June 19, 2015); and JoAnn S. Lublin, "Bringing Hidden Biases to Light: Big Business Teaches Staffers How 'Unconscious Bias' Impact Decisions," *Wall Street Journal*, January 9, 2014, <http://www.wsj.com/articles/SB10001424052702303754404579308562690896896> (accessed June 19, 2015).



214 Massachusetts Avenue, NE
Washington, DC 20002-4999
(202) 546-4400
heritage.org