McConnell, A. R., & Rydell, R. J. (2019). The Implicit Association Test: Implications for understanding consumer behavior. In F. R. Kardes, P. M. Herr, & N. Schwarz (Eds.), Handbook of research methods in consumer psychology (pp. 143-159). Routledge.

8

THE IMPLICIT ASSOCIATION TEST

Implications for Understanding Consumer Behavior

Allen R. McConnell and Robert J. Rydell

Introduction

In the two decades since the publication of the original paper on the Implicit Association Test (IAT; Greenwald, McGhee, & Schwarz, 1998), it has attracted considerable interest from researchers and laypeople alike. According to Web of Science (https://clarivate.com/products/web-of-science/) statistics, the term "implicit association test" has been used in more than 2,000 publications, with those articles being cited by more than 40,000 other papers at a rate increasing each year since 1998. Based on Google Scholar data (www.google.com/scholar), there are six empirical IAT papers with at least 1,000 citations, and the original Greenwald et al. (1998) paper has been cited almost 10,000 times. Further, the Project Implicit website (implicit.harvard.edu) allows people from around the world to complete demonstration IATs via the Internet on issues such as racial bias, political preferences, and gender roles. Since its inception, visitors to the website have completed more than 20 million IATs (K. Ratliff, personal communication, May 7, 2018). Moreover, in recent years, "implicit bias" has been an important element of prominent news stories ranging from discrimination at coffee shops (Park, 2018) to medical training bias (Dembosky, 2015) to voting in U.S. presidential elections (Scott, Lee, & Merica, 2016).

Clearly, understanding and measuring implicit bias are important. Yet what can the IAT tell us? All too often, whether it is in graduate student committee discussions or consulting meetings with deep-pocket clients, we hear statements like "Can we measure that with the IAT?" or "We should run an IAT because it will give us a pure measure." Frankly, we cringe when we hear such utterances because they typically reflect a naivety toward using the IAT. We believe the IAT has great value, but there are important considerations involved in determining whether it is an appropriate tool for researchers. In our chapter, we review the IAT and address important questions one should consider before using it. For example, under what conditions does the IAT show better reliability and validity? What are the limits of the IAT, and what misperceptions exist about it? What trade-offs are involved with various forms of the IAT? How can IAT data be considered at different levels of analysis? We consider these questions with an eye to identifying insights for understanding consumer behavior.

What Is the IAT?

The IAT is a methodology that assesses the relative strength of associations between two conceptual *dimensions*, each of which is instantiated by dichotomous *categories*. For example, consider a marketing researcher who is interested in people's preferences for wine over beer. In this case, the two dimensions would be evaluation (with dichotomous categories of "positive" and "negative") and beverage type (with dichotomous categories of "wine" and "beer").

In such an IAT, participants would view a series of stimuli presented on a computer monitor, one at a time, and categorize each stimulus using one of two response buttons (left and right keys, either on a keyboard or a button box). The stimuli associated with the evaluative dimension could be adjectives (e.g., wonderful, awful, great, horrible), symbols (e.g., smile or frown emoticons), or images (e.g., photos of people with positive or negative facial expressions). Similarly, the stimuli for the beverage dimension could be words (e.g., pilsner, Bordeaux, ale, chardonnay), symbols (e.g., outlines of beer mugs or wine glasses), or images (e.g., photos of beer or wine). For an IAT, what matters is not the content of respondents' classifications (each stimulus is unambiguously positive or negative, or unambiguously wine or beer), but, rather, the speed with which people can make a series of classifications across different blocks of trails.

Table 8.1 illustrates how such a beverage-preference IAT might be structured. Overall, such an IAT might consist of 180 trials, presented in seven blocks (each with 20 or 40 trails), with a brief self-paced rest and orientation period placed between them. In Block 1, participants might see 20 beverage stimuli presented one at a time, and they would press the left key to classify those stimuli that are wines and press the right key to classify those stimuli that are beers. On each trial, each stimulus item appears on the computer monitor with labels presented to remind participants about what the left and right keys represent. Participants are told to make each response as quickly as possible while remaining accurate, and that occasional errors are okay. When participants err (e.g., categorize stout as a wine), a red X appears on the screen, and respondents must press the correct response to advance to the next trial. Typically, errors are infrequent (usually less than 5 percent of trials) because the stimuli are unambiguous in nature, and participants have unlimited time to respond. In Block 2, 20 new evaluation-specific stimuli would be presented (e.g., adjectives with a clear positive or negative connotation), and participants would categorize each stimulus as either positive (left key) or negative (right key).

For Blocks 3 and 4, the trials become more complicated because each response key is mapped onto two different categories that cross dimensions. During these critical *combination blocks*, one stimulus item is still presented on each trial, but now those stimuli could be beverage-related or evaluation-related. In this example, participants would press the left response key if the stimulus is a wine or a positive adjective, and they would press the right response key if the stimulus is a beer or a negative adjective. Typically, two of these combination blocks would be presented back to back, with a brief rest period in between (often, 20 trials in Block 3, followed by a break, followed by 40 additional trials in Block 4). In Block 5, the stimuli for each trial would only be composed of evaluative adjectives (i.e., no beverage stimuli), and participants would categorize negative adjectives by pressing the left response key and categorize positive adjectives by pressing the right response key. Finally, in Blocks 6 and 7 (with 20 and 40 trials, respectively), another combination block with beverage-related or evaluation-related stimuli would be presented, but now participants would use the left response key to classify each stimulus as either a wine or as negative in connotation and use

Table 8.1 IAT to Assess Preferences between Wine and Beer

Block	*Type of Block*	*Left Response Key*	*Right Response Key*
Block 1	Beverages	Wine	Beer
Block 2	Evaluations	Positive	Negative
Blocks 3 & 4	Pro-wine combination	Wine or positive	Beer or negative
Block 5	Evaluations (reversed)	Negative	Positive
Blocks 6 & 7	Pro-beer combination	Wine or negative	Beer or positive

the right response key to classify each stimulus as either a beer or as positive in connotation. A typical IAT such as the example in Table 8.1 takes average participants about 8–10 minutes to complete.

The critical data from an IAT concern the speed with which people render classifications in the combination blocks. In other words, one compares the mean latency for performing Blocks 3 and 4 with the mean latency for performing Blocks 6 and 7. To the extent that one combination block is completed *more quickly* than the other combination block, it reflects a stronger association for the combination block that was performed faster. Thus, if participants perform Blocks 3 and 4 (where wine and positive share the same key) more quickly than Blocks 6 and 7 (where beer and positive share the same key), it suggests a *relatively* stronger evaluative preference for wine *compared with* beer. If participants perform Blocks 6 and 7 more quickly than Blocks 3 and 4, it suggests a *relatively* stronger evaluative preference for beer *compared with* wine. The larger the difference between the combination blocks, the stronger the association. Greenwald, Nosek, and Banaji (2003) provided a refinement to IAT scoring, calculating an effect size measure (D) that captures the mean latency difference between the combination blocks divided by twice their pooled standard deviation (along with exclusion criteria for treating outliers). IAT effect sizes are impressive in magnitude, with Greenwald et al. (2003) observing strong relative associations involving Whites–good (and Blacks–bad; D = 1.00, N = 6,811), men–science (and women–liberal arts; D = 1.04, N = 10,475), and young–good (and old–bad; D = 1.38; N = 10,537).

When IATs are used in research, there are counterbalancing factors included as between-subject factors. Issues such as whether an IAT effect is an artifact of a particular response being associated with the left key or with the right key (e.g., for half of the participants in a beer–wine IAT, beer would be associated with the left key and wine would be associated with the right key) or whether combination block presentation order matters (e.g., for half of the participants, Blocks 3 and 4 have "wine or negative" on one key and "beer or positive" on the other key) are ruled out by counterbalancing. Thus, concerns about an IAT effect being the product of a "particular order or arrangement" or "particular key combinations" or "different for right-handed and left-handed people" are nullified (Greenwald et al., 1998). Some factors, such as combination block order, can produce minor differences in IAT effect sizes (Greenwald & Nosek, 2001), but counterbalancing can ensure that overall mean differences do not reflect procedural artifacts.

The example illustrated in Table 8.1 explores relative preference for wine compared with beer, but the IAT can be adapted to a wide range of consumer behavior issues. For instance, a marketing researcher in the U.S. auto industry might replace beverages with automobile types (e.g., U.S. domestic cars vs. foreign cars). Such a researcher might also be interested in other perceptions of automobiles beyond preference and replace the "evaluation" dimension with "luxury" (e.g., using adjectives such as expensive, cheap, pricey, budget) to assess the relative strength of association between foreign cars and luxury. Conversely, one could retain the "evaluation" dimension and study people's relative preferences for pets (e.g., dogs vs. cats), computers (e.g., Apple vs. Windows), or soft drinks (e.g., Coke vs. Pepsi). Another researcher may be less interested in preferences for Coke or Pepsi products, but, instead, be interested in people's associations of age with these soft drinks and employ the dimensions of soft drinks (e.g., Coke vs. Pepsi) and age (e.g., young vs. old). In short, the IAT can assess a vast array of associations of interest to marketing researchers, and, although evaluation is frequently examined because of its obvious connection to attitudes and persuasion, other dimensions unrelated to liking can be employed, such as luxury, age, health, quality, or patriotism, to name just a few.

The IAT has its origins in social psychology, examining issues such as racial prejudice (e.g., McConnell & Leibold, 2001), self-esteem (e.g., Greenwald & Farnham, 2000), preferences for physically attractive people (e.g., McConnell, Rydell, Strain, & Mackie, 2008), and views of one's own shyness (e.g., Asendorpf, Banse, & Mücke, 2002). However, the IAT has been increasingly used in marketing contexts, such as studying preferences for brands from one's own country compared with foreign brands (Maison, Greenwald, & Bruin, 2001), preferences for Apple or Windows computers (Brunel, Tietje, & Greenwald, 2004), preferences for White or Black advertisement spokespeople (Kareklas, Brunel, & Coulter, 2014), how soft drinks preferences can be systematically altered (Gibson, 2008), or how food preferences for chocolate and fruit assessed by the IAT predict people's actual enjoyment when eating those foods and why people mispredict their enjoyment of those foods (McConnell, Dunn, Austin, & Rawn, 2011). It is also true that many IAT studies in the consumer behavior space are unpublished by consultants and companies in order to maintain competitive advantages. Thus, the published literature does not provide a true sense of how frequently methods such as the IAT are used in marketing contexts.

IAT Reliability and Validity

Of course, the IAT is only a useful measure to the extent that it has reliability and validity. Making things more complicated in assessing these qualities is that each particular IAT (e.g., race bias, soft drink preference, computer preference) is unique. That being said, there is good evidence that the IAT not only demonstrates reliability and validity, but that it exhibits these properties to a much greater degree than many other association-based measures, such as evaluative priming measures or go/no-go association tasks (Nosek, Greenwald, & Banaji, 2007).

For example, Bosson, Swann, and Pennebaker (2000) examined several association-based measures of self-esteem (e.g., sequential priming, Stroop task), including the IAT (i.e., relative association between "me" and "pleasant" and "not me" and "unpleasant"). The self-esteem IAT showed the strongest split-half internal consistency ($r = .69$) among all of the association-based measures used in their study. Another indicator of reliability is whether multiple administrations of an IAT yield similar results. Research by Schmukle and Egloff (2004) assessed associations between the "self" and "anxiety" using an IAT at two different time periods and saw significant correlations between administrations ($r = .50$), with each IAT administration showing internal consistency as well (alphas .80–.84). In a study of racial attitudes using IAT administrations separated by two months, Gawronski, Morrison, Phills, and Galdi (2017) observed good relations ($r = .42$). These correlations for IATs are lower than explicit measures (where social desirability concerns for appearing egalitarian can inflate correspondence), but they are reliable.

Although these correlations and coefficient alphas meet psychometric standards for reliability, the IAT is not a perfect measure. One issue of interest is whether IAT scores can be "faked." For example, people may prefer to appear to be nonracist or to report pro-healthy food preferences for self-presentational reasons. Yet, despite these social pressures, people do reveal relative racial bias against African Americans (compared with White Americans) on race-based IATs (e.g., Greenwald et al., 1998) and show preferences for fattening (compared with healthy) foods on food IATs (e.g., McConnell et al., 2011). Similarly, people instructed to deliberately fake having positive attitudes toward homosexuals were unable to alter their IAT scores on a sexual orientation preference IAT (Banse, Seise, & Zerbes, 2001). However, some studies have shown that IATs can be susceptible to deception (Fiedler & Blümke, 2005; Klauer & Teige-Mocigemba, 2007), though faking is less problematic for implicit measures such as the IAT than for explicit measures (Egloff & Schmukle, 2002). Also, people can alter IAT scores by deliberately responding more slowly across all trials and, thus, eliminating the magnitude of the difference between combination trial blocks, though such efforts can be detected and corrected statistically to some degree (Cvencek, Greenwald, Brown, Gray, & Snowden, 2010). Also, people can alter their IAT scores better toward novel, fictitious social groups where strong, preexisting beliefs do not exist (e.g., De Houwer, Beckers, & Moors, 2007). Overall, although faked IATs can be a concern, they appear difficult to modulate for well-established attitude dimensions and for domains with greater self-relevance, and are far more difficult to falsify than explicit measures such as verbal questionnaires, where appearing socially desirable is easier to accomplish (Steffens, 2004).

With evidence that the IAT exhibits reasonable reliability, the next question to consider is whether it reveals validity. In other words, do IAT effects correspond to outcomes (e.g., other measures of similar constructs, prediction of behavior) that are related to what an IAT purports to capture. For example, McConnell and Leibold (2001) had participants complete explicit measures of racial bias (e.g., feeling thermometers, semantic differentials) for Blacks and Whites separately, and they observed that a race IAT reliably predicted these race differences on these explicit measures of prejudice ($r = .42$), providing evidence of predictive utility.

Although evidence that an IAT correlates with a similar explicit measure is one indicator of validity, the evidence would be even stronger if the IAT predicted independent assessments of behavior rather than self-reports. Indeed, McConnell and Leibold (2001) also had participants interact with Black and White experimenters in two separate interactions, and these interactions were rated by the experimenters (who were unaware of participants' responses on the other measures) and they were video-recorded, allowing third-party judges to examine participant behaviors toward both Black and White experimenters. Indeed, assessments of participants' differential behavior toward Black and White experimenters showed significant correlations to participants' racial IAT scores. For example, the differences in the quality of participant interactions as assessed by the Black and White experimenters predicted participants' racial IAT scores ($r = .39$), and third-party judge ratings of differences in participants' interactions between the Black and White experimenters predicted IAT scores (e.g., differences in speaking time, smiling, speech errors, extemporaneous social comments, all $rs \geq .32$). Thus, participants' racial IAT scores predicted their own differential behavior toward Black and White experimenters in ways that the experimenters themselves, and naive third-party observers, could detect.

Meta-analyses exploring the predictive utility of IAT responses more generally have also found evidence that the IAT is a valid measure. For example, one meta-analysis by Hofmann, Gawronski, Gschwendner, Le, and Schmitt (2005), with 126 studies involving the collection of both IAT data and explicit self-report data (e.g., self-reported preferences for consumer products, social groups, or self-esteem), showed a mean correlation of .24. In the domain of consumer attitudes, the correlation was .34. Hofmann et al. (2005) speculated that the stronger correlations observed for consumer products may reflect weaker self-presentational concerns for consumer behavior (e.g., preferences for Coke over Pepsi) than for social attitudes (e.g., preferences for Whites over Blacks). In line with this reasoning, Gibson (2008) observed a strong correlation between a Coke–Pepsi IAT and explicit reports of preferences between Coke and Pepsi ($r = .51$).

A more recent meta-analysis by Greenwald, Poehlman, Uhlmann, and Banaji (2009) examined published work involving participants who completed both IATs and other measures (e.g., self-reported attitudes, behaviors) involving relevant attitude objects. They found that IAT scores, on average, predicted explicit evaluations reliably ($r = .21$) across 155 relevant studies, with stronger relations observed in less socially sensitive domains such as consumer preferences ($r = .32$) and political party preferences ($r = .54$) than in more sensitive domains such as race bias ($r = .12$) or sexual orientation preference ($r = .17$). In socially sensitive domains such as racial prejudice where explicit measures of prejudice and collected behavioral data were assessed (e.g., interactions with Black and White confederates), the IAT was a significantly better predictor of behavior than were explicit measures of prejudice, reflecting the challenges of assessing socially sensitive attitudes using explicit measures. Overall, the meta-analyses of Hofmann et al. (2005) and Greenwald et al. (2009) provide considerable evidence that IATs predict explicit self-reports and relevant behaviors, and that the IAT can be a better predictor of behaviors than self-reports in contexts where self-presentational concerns exist.

The findings demonstrating IAT validity have been questioned by some individuals. For example, Blanton et al. (2009) criticized the IAT–behavior link reported by McConnell and Leibold (2001), although McConnell and Leibold (2009) pointed out several problematic aspects of the Blanton et al. critique (e.g., Blanton et al. analyzed an incorrect IAT measure, and IAT scores significantly predicted several *different* behavioral indicators whereas Blanton et al. only expressed concerns about one specific outcome). In addition to attacks on specific papers, some have questioned meta-analysis findings. For instance, Oswald, Mitchell, Blanton, Jaccard, and Tetlock (2013) criticized the Greenwald et al. (2009) meta-analysis, claiming a lower aggregate correlation between the IAT and outcomes (r = .15) than what Greenwald et al. reported. Oswald et al. (2013) considered a number of studies not included by Greenwald et al. (2009), many of which had no theoretical expectation of a predictive relation with the IAT. For example, Oswald et al. included studies where race IATs (which assess relative positivity toward Whites *compared with Blacks*) were used to predict behavior toward a single White target individual. As Greenwald, Banaji, and Nosek (2015) noted, one should not expect correspondence between an IAT that looks at *relative* bias between two races and outcomes involving only one-race targets. Further, Greenwald et al. (2015) noted that small but systematic biases can affect minority group members repeatedly over time, and the cumulative impact of small biases directed at many different individuals from the same social group results in meaningful group-level differences in treatment within a society.

Overall, there is solid evidence that IATs are reliable and valid. Across many domains, IAT scores are consistent and predict meaningful outcomes. In consumer behavior research, the correspondence between the IAT and explicit measures appears especially strong because social desirability concerns are often low. Further, there is recent evidence by Hehman, Calanchini, Flake, and Leitner (2018) that IAT reliability and predictive utility increase as one moves from individual-level analyses to aggregate-level analyses. For example, in looking at race IAT data from approximately 2 million U.S. residents, Hehman et al. (2018) found that the correlation among White participants between race IATs and explicit prejudice measures grew as analyses moved from the individual level (r = .21) to the county level (r = .25) to the state level (r = .84). These increases in predictive utility for the IAT reflect that larger samples have less error than any single individual's responses. Although aggregate-level analyses may not be sensible in many situations, these data suggest that, at least conceptually, there is a stable construct for the IAT to capture. Later in our chapter, we return to the implications of using aggregate-level analyses.

What Does the IAT Assess?

What is the IAT supposed to assess? The IAT was designed to measure associations between concepts in memory in ways that are relatively impervious to cognitive control. That is, people should not be able to alter their responses, even if they wanted to do so (and it is this issue that has led to interest in whether the IAT can be faked). Because the IAT is not a direct assessment of respondent knowledge (e.g., people are not asked directly to report their attitudes), and controlling responses on the IAT is difficult, it was thought to be relatively immune to socially desirable responding. Consistent with many models of attitudes (e.g., Fazio, 2007), if one can assess associations directly and bypass more controlled thought during assessment, measures of these associations should more strongly predict behavior (Greenwald et al., 2009). The most common form of associations assessed by the IAT has been between evaluation (e.g., good and bad) and attitude objects (e.g., Whites and Blacks, Coke and Pepsi), and thus the IAT is most often used as an indirect measure of attitudes. However, as noted above, the IAT can assess associations between any set of concepts (e.g., self and shyness; Asendorpf et al., 2002).

How does the IAT measure associations? Consider the IAT presented in Table 8.1 that assesses relative preference for wine compared with beer. To the extent that participants perform the pro-wine combination block more quickly than the pro-beer combination block, it reflects stronger associations in memory between wine and good (and beer and bad) relative to wine and bad (and beer and good). Moreover, because people are making judgments about the categories to which the target stimuli belong, the IAT is thought to measure associations between the *categories* themselves and not the exemplars of the categories presented (Gawronski, 2009). In other words, our example IAT would capture evaluations of beverage categories (i.e., wine and beer) rather than the exemplar stimuli presented (e.g., chardonnay, pilsner).

In a perfect world, the IAT would only assess associations between the dimensions examined. This reasoning is rooted in associative theories of memory, especially many theories of attitudes (e.g., Fazio, 2007; Gawronski & Bodenhausen, 2006; McConnell & Rydell, 2014; Petty & Briñol, 2014; Wilson, Lindsey, & Schooler, 2000) that make assumptions about associative information in memory. These approaches posit that experiences are stored in memory (e.g., I enjoyed this wine, I feel bloated after drinking beer), giving rise to generalized evaluative associations. This view of memory and abstractions, although potentially problematic (e.g., Conrey & Smith, 2007; Schwarz, 2007), underlies the logic of the IAT. Although mental representations can be conceived in many ways (e.g., connectionist processes, object-evaluation associations), most researchers who use the IAT do not try to specify the exact processes they are measuring with the IAT, even though these questions are important for interpretation of their findings.

Unfortunately, it is not clear where the associations measured by the IAT originate or how they may change over time. At a basic level, these associations can be created and altered by repeated pairings between an attitude object and evaluative information, whether that involves impressions of people (e.g., Rydell & McConnell, 2006) or conditioning involving soft drinks (e.g., Gibson, 2008). Sometimes, though, repeated pairings fail to change associations assessed by the IAT (e.g., Gregg, Seibt, & Banaji, 2006), and occasionally one new datum triggers reinterpretation of previous information that alters IAT effects considerably (Mann & Ferguson, 2015). Thus, training people to create or alter the associations in memory toward an attitude object may or may not be effective (Gawronski et al., 2018). In short, there are more underlying IAT effects than simply the number of associative pairings. Evaluations assessed by the IAT sometimes form quickly (e.g., Gregg et al., 2006), sometimes change quickly (e.g., Wittenbrink, Judd, & Park, 2001), and these changes sometimes persist (e.g., Dasgupta & Greenwald, 2001), but often they do not (see Blair, 2002). Sometimes associations measured by the IAT appear due to early socialization experiences (e.g., Rudman, 2004), years of accumulated life experiences (e.g., Jellison, McConnell, & Gabriel, 2004), or pervasive cultural values (e.g., McConnell et al., 2008); however, sometimes recent experiences play a greater role (e.g., Castelli, Carraro, Gawronski, & Gava, 2010). Context can matter as well. For example, in consumer behavior research, Humphreys and LaTour (2013) found that more positive contexts (e.g., reading stories that glorify instead of delegitimize online gambling, using "gaming" instead of "gambling" to describe the online casino industry) led participants to associate the gambling industry more with entertainment (e.g., fun, pleasure) than with crime (e.g., sin, mobster) on an IAT, and these effects were stronger for people with no previous gambling experience. These findings indicate that IAT results are somewhat malleable rather than capturing immutable automatic associations.

Given the ambiguities about what implicit measures assess, some researchers have attempted to understand the multiple processes captured by measures such as the IAT by using multinomial modeling techniques, known generally as the quad model (Sherman et al., 2008). According to this model, responses on the IAT potentially have four components: an automatic association activated in memory (AC); determining an appropriate response in a given context or for a given task (D); using self-regulation to overcome the bias of an automatically activated response from memory (OB); and guessing (G). This approach has helped determine some important factors that influence IAT responding, and we will highlight research involving OB later in our chapter. One limit of the quad model is that, although it helps to quantify facets of IAT responding, it does not help researchers answer questions about the origins of IAT responses. For example, it certainly can be helpful to determine if shifts on IAT scores are due to changes in associations (AC) or changes in motivation (OB). However, such insights do not speak to why those changes occur, though using the quad model in concert with experimental manipulations may help to do so.

Approaches such as the quad model require that participants make more errors on the IAT than is typically the case to calculate stable parameter estimates. These researchers often modify IAT tasks, such as by introducing short response windows for participant classification judgments (e.g., 225–450 ms; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) to increase error rates. When response windows are not used, some researchers will use 20 trials in Blocks 3 and 6 (as described above) but increase the number of trials in Blocks 4 and 7 to 100 presentations each to increase the number of errors during critical combination blocks. These modifications are not used in most IAT research that we discuss in our chapter, but they are necessary for researchers interested in these modeling techniques (see Sherman et al., 2008).

Common Questions and Misconceptions about what the IAT Measures

When the IAT and other implicit measures were developed, there was some hope that they would help researchers capture nonconscious associations (i.e., knowledge that people are unaware of or to which people do not have conscious access), and this view was adopted in the consumer behavior space too (e.g., Perkins, Forehand, Greenwald, & Maison, 2008). However, there is no compelling evidence that the IAT cleanly measures associations that people are unaware of. As noted, there is a general, modest correlation between implicit and explicit measures (Greenwald et al., 2009; Hofmann et al., 2005), but the lack of perfect correspondence does not mean people are unaware of information measured by an IAT. Self-presentation, measurement error, and differences in assessment techniques could all contribute to this modest correlation (Gawronski, 2009). Also, asking people to introspect on their feelings about an attitude object can increase the correlation between the IAT and explicit measures (e.g., Gawronski & LeBel, 2008), which seems to indicate that people have some influence on the information being assessed by the IAT. Thus, the empirical evidence to date seems inconsistent with the idea that the IAT assesses *purely* nonconscious associations.

As mentioned above, research has shown that nonassociative factors can influence the IAT (Sherman et al., 2008). Also, sometimes attempts to alter IAT responses can affect IAT scores (e.g., Fiedler & Blümke, 2005), although such attempts do not always work (e.g., Banse et al., 2001), and in general the IAT is more robust to falsification than explicit measures (Cvencek et al., 2010; Steffens, 2004). However, motivational factors are more likely to influence IAT responses. That is, people are often motivated to appear unbiased (e.g., not wanting to look racist on a race IAT) and thus attempt to self-regulate their IAT responses (e.g., Conrey et al., 2005). Sometimes these attempts are effective, and sometimes they are not. For example, research has shown that older adults are relatively more prejudiced toward Blacks than younger adults on the IAT (Gonsalkorale, Sherman, & Klauer, 2009), and this appears to reflect poor self-regulation among older individuals rather than their having stronger racial prejudice (see Sherman et al., 2008). Put another way, younger adults are able to control their responses on the IAT to some degree, appearing less prejudiced. Strong claims about IAT findings should ideally be supported by more sophisticated examination of the data (e.g., Conrey et al., 2005), using multiple implicit measures (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002), or experimental manipulations (e.g., Dasgupta & Greenwald, 2001). The IAT, like any implicit measure, is not process pure.

We believe this is an important point that many researchers fail to appreciate when using implicit measures such as the IAT. Ideally, one would like to assume that a method such as the IAT provides a pure measure of nonconscious processes (i.e., associations stored in memory). Yet, people must consciously follow experimenter instructions and press buttons to categorize stimuli, which means that any measure derived from an IAT (or any implicit measure, such as sequential priming) must be a blend of automatic (e.g., associations in memory) and controlled (e.g., how people engage with the task) factors. Thus, IAT scores reflect not only associations in memory (often, the focal intent), but other factors such as context and motivations (Sherman, 2009). All measures are subject to this limitation, but, because many people equate IAT scores with associations in memory, the issue of process purity is especially important for the IAT. Process dissociation techniques that attempt to model automatic and controlled components provide researchers with one way to address this issue (Payne, 2008; Sherman et al., 2008).

In addition to concerns about the degree to which IAT scores assess underlying associations in memory, one should consider the nature of the associations one wants to capture with the IAT. In the example provided earlier, a researcher might have confidence that relative preferences between wine and beer could be identified using an IAT in part because these are consumer behaviors relatively low in social desirability concerns (Greenwald et al., 2009; Hofmann et al., 2005). Yet, there might be considerations that would limit the utility of using this beverage-preference IAT. First, there is the issue of *ambivalence*, which is when an attitude object is associated with both positivity and negativity. For example, consider someone who loves white wine but hates red wine. Overall, this person has ambivalence about wine and, thus, may not show a preference on the IAT, despite having very strong attitudes about wine. In such a circumstance, one might want to construct more nuanced IATs, such as white wine versus red wine. There is another form of ambivalence that might exist—strong but equivalent evaluations of both categories along a dimension. Consider two people, one who does not like alcohol at all (and, thus, has negative attitudes toward both wine and beer) and another person who loves all forms of alcohol (and, thus, has positive attitudes toward both wine and beer). These two people would, theoretically, show no implicit preference on our wine–beer IAT, despite having starkly different feelings about these beverages. Because the IAT is a *relative* measure, it is not clear what an IAT score of zero means (and, further, arguing that an IAT score of zero is "neutral" is also problematic; see McConnell & Leibold, 2009, and Nosek et al., 2007). Ambivalence must be assessed in some other way (e.g., de Liver, van der Pligt, & Wigboldus, 2007; Petty, Tormala, Briñol, & Jarvis, 2006), and researchers need to consider possible nuances involved in people's attitudes (e.g., Is "wine" too broad a category? Are wine and beer polar opposites in the same way that gender or smartphone platform choices are viewed by most people?). In the next section, we explore different versions of the IAT that can offer some approaches to potentially deal with issues such as ambivalence.

There are misconceptions about what the IAT can and cannot assess, and there are limits to interpreting IAT scores. Yet many researchers simply use the IAT without understanding these complexities or without addressing them. Further, using the IAT does not shield researchers from even more basic considerations, such as whether stable evaluative representations exist (e.g., Schwarz, 2007) or when attitudes should predict behavior (e.g., Fazio, 1986; McConnell & Rydell, 2014). Ultimately, we believe that the IAT can be an effective tool for researchers, but using it without appreciating these important issues is fraught.

Variants of the IAT

The traditional IAT: The most commonly used version of the IAT is the one originally introduced by Greenwald et al. (1998), the traditional IAT. As described in Table 8.1, this version presents participants with judgments involving two dimensions (in our example, beverages and evaluations), and, for each dimension, there are two categories that anchor the dimension (e.g., wine–beer, good–bad). On the critical combination blocks, two categories share the same response key (e.g., "wine or good," "beer or bad"), and the difference in mean latencies between these combination blocks establishes the strength of association between the dimensions (in our example, whether one has a relative preference for wine or beer).

The relativistic nature of the traditional IAT underscores an inherent ambiguity in interpreting IAT scores (Nosek & Banaji, 2001). When considering the wine–beer IAT, any particular IAT score reflects many different interpretations. Does a "pro-wine" IAT score reflect someone who really loves wine or someone who really hates beer? Similarly, does a "pro-beer" IAT result reflect someone who loves beer or someone who hates wine? Does a near-zero IAT score reflect someone who loves both beverages, hates both beverages, or is indifferent to both? And, as mentioned previously, what if someone loves white wines but hates red wines, or loves light pilsners but hates heavy ales? The relativistic nature of the traditional IAT fuses category dimensions together, and there is no way to "break IAT responses down" to particular trials in particular blocks to recover separate components of respondents' associations (Nosek et al., 2007). In response to the limits of the traditional IAT, other versions have been developed.

The single-category IAT: The single-category IAT (SC-IAT; Karpinski & Steinman, 2006) can be a useful method. In the SC-IAT, there are still two dimensions, but it uses only three categories instead of four. For example, consider someone who believes that people's preferences for wine and beer may be independent and, thus, wants to assess associations with wines independent of associations with beers. In this case, a researcher could use an SC-IAT with one combination block where wine is associated with positivity (e.g., left key is "wine or positive," right key is "negative") and the other combination block where wine is associated with negativity (e.g., left key is "positive," right key is "wine or negative").

Similar to the traditional IAT, SC-IAT scores reflect the difference in response latencies on the combination blocks, based on the D algorithm (Greenwald et al., 2003) used in the traditional IAT. The combination block that is performed more quickly indicates the nature of the association (e.g., greater positive associations with wine if the combination block that pairs "wine or positive" is performed more quickly than the combination block that pairs "wine or negative"). The obvious advantage of the SC-IAT is that associations involve a single category (e.g., wine) and not a relative association between categories (i.e., wine *relative to beer*). In situations where dimension categories are not polar opposites, the SC-IAT may be a better choice than the traditional IAT. However, the traditional IAT may be preferable in situations where relative associations would be important for predicting relative outcomes, such as when race-based associations using the traditional IAT predict differences in behaviors toward Black and White social targets (Greenwald et al., 2015; McConnell & Leibold, 2001).

Karpinski and Steinman (2006) provide good evidence for SC-IAT utility in domains such as self-esteem, racial attitudes, and soda preferences. For instance, participants completed a traditional IAT to assess relative preferences between Coke and Pepsi (similar to Gibson, 2008), separate SC-IATs for Coke and for Pepsi, and explicit measures of Coke and Pepsi attitudes, all of which were used to predict participant preferences for drinking Coke or Pepsi. Karpinski and Steinman (2006) found that soda preferences for Coke over Pepsi were (in multiple regression analyses) uniquely predicted by pro-Coke preferences assessed by traditional IATs, and by people's SC-IATs showing more negativity associated with Pepsi (significant unique effect) and SC-IATs showing more positivity associated with Coke (marginal unique effect). In self-esteem, racial prejudice, and soda preference domains, they found SC-IATs made unique contributions in predicting preferences and behaviors above and beyond traditional IATs. Although SC-IATs can be susceptible to concerns such as faking (e.g., Stieger, Goritz, Hergovich, & Voracek, 2011), they offer an attractive alternative for assessing associations for single categories.

SC-IATs have been used in consumer research. For example, Brough, Wilkie, Ma, Isaac, and Gal (2016) used SC-IATs to assess associations with environmentalism. Using separate SC-IATs involving gender, they found that environmentalism (green products such as recycling bins and plug-in cars) was associated with femininity (using female names) but not masculinity (using male names). Moreover, they found that male, but not female, participants were more willing to support green causes when green branding affirmed masculine stereotypes (e.g., an organization named "Wilderness Rangers" with a howling wolf as its symbol) than when branding was conventional (e.g., the same organization named "Friends of Nature" with a tree as its symbol).

The personalized IAT: Olson and Fazio (2004) developed a version of the IAT to try to differentiate people's personally held attitudes from evaluations consensually held by others in their culture but not personally endorsed. The modification they made to the IAT was simple. Instead of using categories for an evaluative dimension such as "good" and "bad," they used the categories "I like" and "I dislike" and target stimuli for which people typically hold strong, yet idiosyncratic, evaluations (e.g., country music, pickles). In theory, this approach would provide a better assessment of people's own evaluations, because the associations assessed would not be contaminated by cultural knowledge that respondents may not personally endorse. Assessing people's own evaluations rather than attitudes influenced by cultural beliefs would produce a better measure and increase attitude–behavior correspondence (Fazio, 2007).

There are some important issues with the personalized IAT that merit discussion. First, there is no way to determine whether an error is made on trials where target stimuli are not objectively good or bad (e.g., the logic of the personalized IAT is to use targets such as "pickles" because some people really love them and some people really hate them), which has implications for assessing whether participants are performing the task correctly and for what constitutes an errant trial. Second, and more important, it is not clear that cultural and personal associations are stored differently or perhaps distinctly with some sort of memorial tag sensitive to IAT instructions (Nosek & Hansen, 2008; Petty & Briñol, 2014). The perspective that people store personally endorsed evaluations differently than non-endorsed evaluations seems problematic, and, instead, it is more likely that object-specific associations are simply stored in memory. Issues involving validity or personal endorsement would seem better dealt with after the association has become activated (Gawronski & Bodenhausen, 2006; Petty et al., 2006; Strack & Deutsch, 2004). Nonetheless, Olson and Fazio (2004) found that the personalized IAT predicted behavior and responses better than a traditional IAT. Further, Han, Olson, and Fazio (2006) found that a traditional IAT was influenced by knowledge obtained from observing others' opinions, whereas the personalized IAT was not. On the other hand, Nosek and Hansen (2008) observed that scores on traditional IATs were unrelated to cultural knowledge. In our view, it is an open question as to what the personalized IAT assesses that is different than the traditional IAT.

Levels of Analysis

Group-level analyses: IAT data can be considered at a number of levels of analysis, which, in turn, can speak to many interesting issues. First, one can examine general patterns of association revealed by IATs across participants, focusing on group-level analyses. For example, Greenwald et al. (1998) found that people, on average, show preferences for flowers (e.g., roses, tulips) over insects (e.g., wasps, horseflies) and for musical instruments (e.g., flutes, pianos) over weapons (e.g., guns, hatchets) on IATs. Greenwald et al. (1998) also showed reliable preferences that systematically differed between social groups. Specifically, they observed differences on a Korean–Japanese preference IAT (using Korean and Japanese surnames to capture the race dimension) as a function of whether the participants were Korean-Americans or Japanese-Americans, with each group showing a relative preference for their own social ingroup. Group-level IAT findings such as these show consistent associations among collections of participants (e.g., most people prefer flowers to insects; most Korean-Americans prefer Korean names to Japanese names, whereas Japanese-Americans show the opposite preference).

Individual-level analyses: A second approach to IAT data, focusing on individual-level analyses, can be used to make predictions about specific respondents. Rather than trying to document reliable differences between groups, researchers can use individual differences in IAT scores to predict other outcomes for the individual. Consider the case of racial bias as assessed by the IAT. Greenwald et al. (1998) showed that White participants showed a relative preference for Whites over Blacks on a racial IAT, establishing that typical White participants have a relatively more negative association with Blacks than with Whites. However, rather than documenting mean biases for a collection of respondents, researchers can use any given participant's IAT score to predict other relevant behaviors and beliefs. Indeed, McConnell and Leibold (2001) adopted this individual-level analysis approach. Specifically, they assessed race IAT scores for White participants (similar to Greenwald et al., 1998) and found that the relative degree of implicit racial bias exhibited by participants predicted discrepancies in their behavior (e.g., amount of smiling, speech errors), with more negative treatment being revealed in interactions involving a Black experimenter compared with interactions with a White experimenter. When using the IAT for individual-level analyses, the focus is on predicting a particular person's behavior or beliefs rather than assessing general group-level associations.

Aggregate-level analyses: Finally, some recent and intriguing work using the IAT has focused on how IAT bias in identifiable locations (e.g., zip codes, states, geographic regions), when aggregated, predicts behaviors, attitudes, and health outcomes for those locations. Often, these locations are tied to core-based statistical areas (CBSAs), which are geographic areas defined by the U.S. Office of Management and Budget. Using data organized around CBSAs, researchers can correlate mean IAT scores from a CBSA with federal statistics (e.g., census data, health statistics) for the same CBSA while controlling for potential confounds (e.g., differences in income, education, crime rates). For example, research by Hehman, Flake, and Calanchini (in press) used race IATs from approximately 1.8 million U.S. respondents who visited the Project Implicit website (which collects location data for visitors) to predict the disproportionate use of lethal force by police officers against Black citizens. Specifically, they observed that, in CBSAs with relatively greater implicit racial bias against Blacks, police officers were more likely to use lethal force disproportionately against Blacks in those locations, while controlling for factors such as crime rates, population density, Black population, explicit prejudice, housing segregation, and racial differences in education and income levels. This approach adopts the perspective that people in a community are influenced by prevailing norms (in this case, implicit prejudice), and, by indexing those norms (i.e., location-specific IATs), one can predict outcomes tied to these norms (e.g., racial bias in lethal force used in those locations). Although this work is correlational in nature, the idea that IATs can tap into a "community mind" is intriguing.

This aggregate-level approach with IAT data has been used in predicting other behavioral outcomes. For example, state-level racial bias on IATs have been shown to predict state-level Medicaid expenditures, with states revealing more implicit bias against Blacks (using Project Implicit website data) showing fewer expenditures for these medical programs that are often perceived to be especially beneficial to African-American citizens (Leitner, Hehman, & Snowden, in press). Finally, another study examined how IAT-measured racial bias can predict mortality rates among Blacks (Leitner, Hehman, Ayduk, & Mendoza-Denton, 2016). In this study, Leitner et al. (2016) relied on race IAT data obtained from Project Implicit, and these researchers found that mortality rates for Blacks from circulatory-related diseases were greater in U.S. counties revealing more implicit bias against Blacks, suggesting that African-Americans who live in locales with greater implicit prejudice may experience more stress, which, in turn, puts them at greater risk for cardiac-related illnesses and subsequent death.

Not only do aggregate-level data offer insights for real-world problems such as lethal force, entitlement spending, and stress-related mortality, but, as noted earlier, data aggregated from thousands of respondents, compared with single individuals, have lower variability because individual-level errors are minimized in larger sample sizes. Indeed, the correlation between explicit and implicit racial bias grows stronger as the level of data analysis grows in size from individuals, to CSBAs, to states (Hehman et al., 2018; Leitner et al. in press). Research using aggregate-level IAT data is nascent, but it holds promise for understanding important real-world behavior (e.g., law enforcement, public policy, health disparities) and for obtaining parameter estimates of constructs with reduced measurement error and greater predictive utility.

Conclusions

In this chapter, we reviewed the promise and pitfalls of using the IAT. In just two decades, the impact of the IAT and of implicit bias is undeniable. The findings discussed show that researchers can view the IAT as a potentially useful tool for their work. However, they must be cognizant of its strengths and limits. Although much of the work we reviewed comes from the social psychology literature, the IAT can shed light on important issues for marketing scholars and practitioners alike. Indeed, we discussed findings showing that the IAT provides insights for consumer behavior in areas ranging from spokesperson preferences (e.g., Kareklas et al., 2014) to improving perceptions of online gambling (e.g., Humphreys & LaTour, 2013) to promoting sustainable consumption (e.g., Brough et al., 2016).

More broadly, we encourage consumer behavior researchers to think carefully about the questions they ask and how the IAT fits in their conceptual and theoretical frameworks. The IAT is a measure of associations and, thus, it is best at capturing knowledge that itself is associative rather than propositional in nature (see Gawronski & Bodenhausen, 2014; McConnell & Rydell, 2014; Strack, & Deutsch, 2004). Yet it is not a process-pure measure, and, thus, researchers should not view IAT scores as an unfettered index of mental associations.

Many consumer behavior decisions (e.g., buying toothpaste) are relatively low in social desirability, and, thus, motivational factors that often modulate IAT performance in social psychology research such as racial or sexual orientation prejudice (e.g., Conrey et al., 2005; Gonsalkorale et al., 2009) may have less impact in marketing contexts. Indeed, meta-analyses examining the correspondence between IAT scores and explicit attitude measures show stronger correlations in consumer behavior research than in most domains examined in social psychology (Greenwald et al., 2009; Hofmann et al., 2005). However, sometimes socially sensitive domains are a focus in consumer behavior research using the IAT. For example, the IAT has been used to consider the role of race in spokesperson preferences (e.g., Kareklas et al., 2014) or the role of political ideologies in reactions to Native American brand imagery (e.g., Angle, Dagogo-Jack, Forehand, & Perkins, 2017). Thus, although motivational concerns may be relatively lower in marketing research, they can be important. Also, it may be the case that motivational factors such as perceptions of product status or perceptions of the morality of a product's manufacturer will impact IAT performance, and future work should consider the interplay of implicit associations and product-specific motivations in consumer behavior.

Relatedly, researchers should consider what sort of consumer behaviors they wish to understand and predict. For example, because implicit measures are often more predictive of spontaneous behaviors than thoughtful, deliberative action (Fazio, Jackson, Dunton, & Williams, 1995; McConnell & Leibold, 2001; Rydell & McConnell, 2006), measures such as the IAT may be especially (and, at times, uniquely) predictive of impulse purchases compared with high-effort decisions such as buying a car. Further, contexts involving consumption and in-the-moment experiences (e.g., dining) may be especially good candidates for IAT insights. For example, McConnell et al. (2011) found that people's accuracy in predicting what foods they would enjoy eating is limited, and that IAT measures of food preferences were uniquely able to predict the errors that people made in anticipating how much they would enjoy eating particular foods (e.g., participants who showed especially strong preferences for chocolates over apples on an IAT were *especially surprised*, in the moment, by how much more than they expected they enjoyed eating chocolates rather than apples). Thus, the IAT can help researchers map consumer blind-spots in consumption experiences, allowing marketers to anticipate when consumers' expectations fail them. Moreover, even in contexts where high-effort deliberation is likely (e.g., buying a home, disclosing sensitive information on the Internet), decisions implemented under conditions of cognitive depletion (e.g., fatigue, stress, strong emotions, distraction) will increase the influence that associative knowledge has on consumer behaviors and judgment (e.g., Dinev, McConnell, & Smith, 2015; Hofmann, Strack, & Deutsch, 2008).

In addition to considering circumstances under which consumer decisions are made (e.g., impulse buys, consumption experiences, cognitive resources), researchers should think about the nature of consumer choices. For example, in situations where decisions reflect an either/or decision (e.g., buying an Apple or Android smartphone), the traditional IAT may be especially adept at predicting purchase decisions, owner satisfaction, and long-term loyalty. However, in cases where the decision space is composed of many options (e.g., buying a Ford, GM, Toyota, BMW, Kia), any form of IAT (e.g., traditional, SC-IAT) may be of little use. Also, although positivity toward products is important in predicting consumption, attitudes can be poor predictors of behavior (Fazio, 1986). Undoubtedly, purchase decisions reflect a multitude of considerations (e.g., perceptions of quality, prestige, style), and using implicit measures such as the IAT can assist marketing researchers to "map constellations of associations" with products that can enhance predictions of purchase behaviors above and beyond mere product positivity.

Finally, we believe there will be interesting possibilities to consider aggregate-level data in consumer behavior as new techniques (e.g., web-based IAT administration, new modeling statistics) make such efforts more tractable. We are very impressed by recent research showing how aggregate-level IAT data can predict outcomes ranging from law enforcement behavior (Hehman et al., in press) to stress-related mortality (Leitner et al., 2016). As marketing efforts and multinational brands expand their global reach, thinking about predictive utility across regions and countries will shed important big-data light on consumer behavior.

In sum, researchers who approach the IAT with sophistication will possess a useful tool for understanding consumer behavior. Two decades of work demonstrates the value of the IAT for social scientists interested in predicting and understanding human action, and its potential for consumer behavior is considerable. We look forward to seeing how marketing scholars in the next decade leverage the IAT for important product applications and theory development.

References

Angle, J. W., Dagogo-Jack, S. W., Forehand, M. R., & Perkins, A. W. (2017). Activating stereotypes with brand imagery: The role of viewer political identity. *Journal of Consumer Psychology*, *27*, 84–90.

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality & Social Psychology*, *83*, 380–393.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift fur Experimentelle Psychologie*, *48*, 145–160.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality & Social Psychology Review*, *6*, 242–261.

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong claims and weak evidence: Reassessing the predictive utility of the IAT. *Journal of Applied Psychology*, *94*, 567–582.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality & Social Psychology*, *79*, 631–643.

Brough, A. R., Wilkie, J. E. B., Ma, J., Isaac, M. S., & Gal, D. (2016). Is eco-friendly unmanly? The green-feminine stereotype and its effect on sustainable consumption. *Journal of Consumer Research*, *43*, 567–582.

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology*, *14*, 385–404.

Castelli, L., Carraro, L., Gawronski, B., & Gava, K. (2010). On the determinants of implicit attitudes: When the present weighs more than the past. *Journal of Experimental Social Psychology*, *46*, 186–191.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: The quad model of implicit task performance. *Journal of Personality & Social Psychology*, *89*, 469.

Conrey, F. R., & Smith, E. R., (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition, 25*, 718–735.

Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S., & Snowden, R. J. (2010). Faking the Implicit Association Test is statistically detectable and partially correctable. *Basic & Applied Social Psychology, 32*, 302–314.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality & Social Psychology, 81*, 800–814.

De Houwer, J., Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the Implicit Association Test. *Journal of Experimental Social Psychology, 43*, 972–978.

de Liver, Y., van der Pligt, J., & Wigboldus, D. (2007). Positive and negative associations underlying ambivalent attitudes. *Journal of Experimental Social Psychology, 43*, 319–326.

Dembosky, A. (2015, September 7). Training doctors to spot their own racial biases. *CNN*. Retrieved from www.cnn.com/2015/09/07/health/healthcare-racial-bias/index.html

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice. *Journal of Personality & Social Psychology, 82*, 835.

Dinev, T., McConnell, A. R., & Smith, H. J. (2015). Informing privacy research through information systems, psychology, and behavioral economics: Thinking outside the "APCO" box. *Information Systems Research, 26*, 639–655.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality & Social Psychology, 83*, 1441–1455.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 1, pp. 204–243). New York: Guilford.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25*, 603–637.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality & Social Psychology, 69*, 1013–1027.

Fiedler, K., & Blümke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology, 27*, 307–316.

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology, 50*, 141–150.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Gawronski, B., & Bodenhausen, G. V. (2014). Implicit and explicit evaluation: A brief review of the Associative Propositional Evaluation model. *Social & Personality Psychology Compass, 8*, 448–462.

Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology, 44*, 1355–1361.

Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures: A longitudinal analysis. *Personality & Social Psychology Bulletin, 43*, 300–312.

Gawronski, B., Rydell, R. J., De Houwer, J., Brannon, S. M., Ye, Y., Vervliet, B., & Hu, X. (2018). Contextualized attitude change. *Advances in Experimental Social Psychology, 57*, 1–52.

Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the Implicit Association Test. *Journal of Consumer Research, 35*, 178–188.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology, 45*, 410–414.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the implicit association test can have societally large effects. *Journal of Personality & Social Psychology, 108*, 553–561.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality & Social Psychology, 79*, 1022–1038.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality & Social Psychology, 74*, 1464–1480.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift fur Experimentelle Psychologie, 48*, 85–93.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality & Social Psychology, 85*, 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality & Social Psychology, 97*, 17–41.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality & Social Psychology, 90*, 1–20.

Han, H. A., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology, 42*, 259–272.

Hehman, E., Calanchini, J., Flake, J. K., & Leitner, J. B. (2018). *Establish construct validity evidence for region-level measures of explicit and implicit racial bias*. Manuscript under editorial review.

Hehman, E., Flake, J. K., & Calanchini, J. (in press). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological & Personality Science*.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality & Social Psychology Bulletin, 31*, 1369–1385.

Hofmann, W., Strack, F., & Deutsch, R. (2008). Free to buy? Explaining self-control and impulse in consumer behavior. *Journal of Consumer Psychology, 18*, 22–26.

Humphreys, A., & LaTour, K. A. (2013). Framing the game: Assessing the impact of cultural representations on consumer perceptions of legitimacy. *Journal of Consumer Research, 40*, 773–795.

Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and related behaviors and beliefs among gay and straight men. *Personality & Social Psychology Bulletin, 30*, 629–642.

Kareklas, I., Bruner, F. F., & Coulter, R. A. (2014). Judgment is not color blind: The impact of automatic color preference on product and advertising preferences. *Journal of Consumer Psychology, 24*, 87–95.

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality & Social Psychology, 91*, 16–32.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43*, 648–655.

Leitner, J. B., Hehman, E., Ayduk, O., & Mendoza-Denton, R. (2016). Racial bias is associated with ingroup death rate for Blacks and Whites: Insights from Project Implicit. *Social Science and Medicine, 170*, 220–227.

Leitner, J. B., Hehman, E., & Snowden, L. R. (in press). States higher in racial bias spend less on disabled Medicaid enrollees. *Social Science and Medicine*.

Maison, D., Greenwald, A. G., & Bruin, R. (2001). The Implicit Association Test as a measure of consumer attitudes. *Polish Psychological Bulletin, 2*, 61–79.

Mann, T. C., & Ferguson, M. J. (2015). Can we undo our first impressions? The role of reinterpretation in reversing implicit evaluations. *Journal of Personality & Social Psychology, 108*, 823–849.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442.

McConnell, A. R., & Leibold, J. M. (2009). Weak criticisms and selective evidence: Reply to Blanton et al. (2009). *Journal of Applied Psychology, 94*, 583–589.

McConnell, A. R., Dunn, E. W., Austin, S. N., & Rawn, C. D. (2011). Blind spots in the search for happiness: Implicit attitudes and nonverbal leakage predict affective forecasting errors. *Journal of Experimental Social Psychology, 47*, 628–634.

McConnell, A. R., & Rydell, R. J. (2014). The Systems of Evaluation Model: A dual-systems approach to attitudes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual process theories of the social mind* (pp. 204–217). New York: Guilford.

McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality & Social Psychology, 94*, 792–807.

Nosek, B. A., & Banaji, M. R. (2001). The go/no go association task. *Social Cognition, 19*, 625–666.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press.

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition & Emotion, 22*, 553–594.

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality & Social Psychology, 86*, 653–667.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality & Social Psychology, 105*, 171–192.

Park, M. (2018, April 17). What the Starbucks incident tells us about implicit bias. *CNN*. Retrieved from www.cnn.com/2018/04/17/health/implicit-bias-philadelphia-starbucks/index.html

Payne, B. K. (2008). What mistakes disclose: A process dissociation approach to automatic and controlled processes in social psychology. *Social & Personality Psychology Compass, 2*, 1073–1092.

Perkins, A., Forehand, M., Greenwald, A. G., & Maison, D. (2008). The influence of implicit social cognition on consumer behavior: Measuring the non-conscious. In C. Haugtvedt, P. Herr, & F. Kardes (Eds.), *Handbook of consumer psychology* (pp. 461–475). Hillsdale, NJ: Erlbaum.
Petty, R. E., & Briñol, P. (2014). The elaboration likelihood and metacognitive models of attitudes: Implications for prejudice, the self, and beyond. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 172–187). New York: Guilford Press.
Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis. W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality & Social Psychology, 90*, 21–41.
Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science, 13*, 79–82.
Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality & Social Psychology, 91*, 995–1008.
Schmukle, S. C., & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance? *European Journal of Personality, 18*, 483–494.
Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25*, 638–655.
Scott, E., Lee, M. J., & Merica, D. (2016, July 8). Clinton says "implicit bias" still exists in U.S. *CNN*. Retrieved from www.cnn.com/2016/07/08/politics/hillary-clinton-joe-biden/index.html
Sherman, J. W. (2009). Controlled influences on implicit measures: Confronting the myth of process-purity and taming the cognitive monster. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new wave of implicit measures* (pp. 391–426). Hillsdale, NJ: Erlbaum.
Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*, 314–335.
Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165–179.
Stieger, S., Goritz, A. S., Hergovich, A., & Voracek, M. (2011). Intentional faking of the Single Category Implicit Association Test and the Implicit Association Test. *Psychological Reports, 109*, 219–230.
Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality & Social Psychology Review, 8*, 220–247.
Wilson, T. D., Lindsey, S., & Schooler, T. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality & Social Psychology, 81*, 815–827.