

Not So Fast: Ten Challenges to Importing Implicit Attitude Measures to Media Psychology

HART BLANTON

Department of Psychology, University of Connecticut, Storrs, Connecticut, USA

JAMES JACCARD

Silver School of Social Work, New York University, New York, New York, USA

The implicit attitude construct has captured researchers' imaginations and dramatically shifted research priorities in the attitude field. It is likely that these measures will be increasingly utilized by researchers studying media-related phenomena. We review the psychometric foundations of current implicit measures and introduce what we view as 10 pressing psychometric, methodological, and conceptual challenges facing the implicit attitude field today. We call for research that revises current empirical approaches to better address these challenges and provide recommendations for media psychologists who might incorporate these measures into their research programs.

The move from explicit to implicit attitude measurement is one of the most significant changes to occur in the attitude literature in the last 20 years. Although early characterizations of attitudes focused on the implicit and unconscious influences of evaluations on behavior (Hovland, Janis, & Kelley, 1953), and although social psychology has a rich tradition of measuring attitudes indirectly (Webb, Campbell, Schwartz, & Sechrest, 1966) or via methods that might overcome self-report biases (e.g., Jones & Sigall, 1971), newer implicit attitude measures have inspired imaginations in ways that could not have been anticipated in these early efforts. We attribute recent interest to two broad trends within the field, one theoretical and the other methodological. Theoretically, substantial attention has been given to dual process theories

Address correspondence to Hart Blanton, Department of Psychology, Unit 1020, University of Connecticut, 406 Babbidge Road, Storrs, CT 06269-1020, USA. E-mail: hart.blanton@uconn.edu

that contrast automatic, spontaneous and unconscious aspects of human cognition with their more deliberate, controlled and conscious qualities (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Devine, 1989; Fazio, 1995; for reviews relevant to media psychology, see Ewoldsen, Rhodes, & Fazio, 2014). Explicit (self-report) measures of constructs rely on deliberate, controlled, and conscious responding, whereas implicit measures seek to unlock the kinds of automatic and unconscious processes that many psychologists now seek to study. Methodologically, the advent of computer-based inventories has greatly refined the precision with which stimuli can be presented and responses quantified (Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998). As a result, social psychological journals have introduced a multitude of new implicit inventories to researchers (Petty, Fazio, & Briñol, 2009; Wittenbrink & Schwarz, 2007), any one of which might have the potential to reveal psychological states unavailable to researchers who rely on explicit inventories.

Implicit attitude measures also have been important for the impact that social psychology is having on other social science disciplines. Social psychology has become an exporter of implicit attitude measures, and it is common to see these inventories applied to topics ranging from public health and organizational studies to political science and legal research. We expect media psychology will seize on this trend as well. Many of the ways that media influences individuals could go undetected by self-report, and this will likely inspire a move to implicit alternatives. Consider consumer behavior as one example. Advertising that associates positive images with products might cause individuals to internalize positive product associations even if their explicit beliefs about the product seem unaffected (e.g., Strick, van Baaren, Holland, & van Knippenberg, 2009). Or consider prejudice as another example. Stereotypic media portrayals of minorities could promote racial and ethnic biases that individuals fail to consciously appreciate (Dasgupta & Greenwald, 2001). Even in areas of media research where explicit measures have proved useful, researchers might use newer classes of implicit attitude measures to improve their predictive models. For example, explicit partisan political attitudes have been found to predict selective exposure to political news (Stroud, 2011). Implicit partisan attitudes might improve prediction of this outcome over and above these explicit attitudes (see Galdi, Gawronski, Arcuri, & Fries, 2012).

NOT SO FAST

We expect that the coming decade will be marked by a sharp increase in the use of implicit measures in media studies but we also believe that the methodological standards for this research need to be raised relative to past work in social psychology. We state at the outset that the perspective

we take in this article is one that some will find controversial. We draw different conclusions than many implicit attitude researchers because we approach implicit measures from a different starting point. Whereas most implicit attitude researchers come to this topic as laboratory-trained social psychologists with interests in tapping the unconscious, spontaneous, and uncontrollable thoughts that explicit measures might fail to capture, we approach this topic as methodologists interested in exploring and elaborating the psychometric properties of these new inventories. Despite the important work researchers are pursuing on automatic, spontaneous, and unconscious aspects of human cognition, we see nontrivial psychometric limitations in the way current measures have been designed, used, and interpreted.

In this article, we describe what we think are the 10 biggest challenges facing those interested in using valid, reliable, and nonarbitrary implicit attitude measures. For pedagogical purposes, we use an example of a researcher studying media effects on implicit anti-obesity bias or anti-fat bias. We chose this example because there already has been movement toward the use of implicit measures to study the way media depictions of beauty, health, and sex appeal influence outcomes related to health, diet, wellbeing, and prejudice (Ahern, Bennet, & Hetherington, 2008; Berry, McLeod, Pankratow, & Walker, 2013; Bissel & Hays, 2011; Domoff et al., 2012; Gapinski, Schwartz, & Brownell, 2006; Häfner et al., 2008; Huang & Chang, 2005; Skorek & Dunham, 2012). Our analysis, however, is relevant to researchers interested in implicit attitudes measures for any topic in media psychology. We initially focus on challenges facing researchers who use implicit measures to predict other criteria of interest (Challenges 1–8). We then consider challenges facing researchers who use implicit measures as outcome variables (Challenge 9) or as assessment tools (Challenge 10). For each challenge, we present recommendations for addressing it in future research.

TEN CHALLENGES FOR THE MEASUREMENT OF IMPLICIT ATTITUDES

I am pessimistic that these 10 psychometric, conceptual, and methodological issues will be confronted by persons in the media psychology field, because they have not been confronted by the people in social psychology where these measures originated (Reviewer 1).

Before describing the challenges, we wish to comment on the above point that was made by one of the reviewers of an earlier manuscript draft. Although we believe the challenges we identify should be given careful consideration, we agree with the reviewer that mainstream social psychology has given little attention to formal psychometric critiques of implicit measures.

We say this as researchers who have devoted considerable attention the last 10 years to pointing out fundamental psychometric difficulties of the methods that have produced most modern implicit measures (Blanton & Jaccard, 2006a, 2006b; Blanton et al., 2009a, 2009b; Blanton & Jaccard, 2008; Blanton, Jaccard, Gonzales, & Christie, 2006; Blanton & Mitchell, 2011; Buhrmester et al., 2011; Jaccard & Blanton, 2007; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013), only to see these measures continue to be constructed and utilized with little regard for psychometric concerns. As this same reviewer went on to note: “There is an incentive structure that discourages the research the authors recommend: There is more glory in publishing one more [implicit measurement] study than doing the psychometric work needed to address these 10 challenges.” This somewhat cynical view of publication incentives belies that much more the importance of “raising the methodological bar” so that our scientific inferences stand the scrutiny of the highest of methodological standards. It is in this spirit that we review these concerns with an eye toward media applications.

Challenge 1: Be Explicit about Going Implicit

The Challenge. A starting point for any researcher is to articulate the rationale for using an implicit attitude measure in the first place. Most typically, the justification for implicit attitude assessment takes one of two somewhat tenuous and at times contradictory forms. The most common justification for using implicit measures is tied to conceptual considerations, namely that they tap into facets of attitudes or cognitions that are more automatic, spontaneous, or unconscious than attitudes and cognitions that exist in declarative memory. As such, the argument goes, they represent components of attitudes that have been ignored by traditional attitude theorists. It is conceivable, for instance, that media representations of the “thin ideal” can promote feelings of dislike for overweight individuals that respondents can neither observe nor report. If so, then indirect measures are needed.

Interestingly, the pursuit of alternatives to self-report is not new. Attitude research has a long history dating back over 50 years of measuring attitudes and other cognitive states using indirect methods that do not rely on self-reporting. These strategies include physiological assessments, environmental traces, projective methods, and a host of clever techniques designed to circumvent limitations of self-report (see Fishbein & Ajzen, 1975; Webb et al., 1966). Rather, what is new is the decision to give attitudes assessed indirectly a new label—that is, implicit attitudes—and then to assert that, because assessment via implicit measures differs from assessment via self-report, the resulting attitude has different qualities. Earlier developers of indirect attitude measures never conceptualized their inventories as instruments that tapped the more automatic, spontaneous, or unconscious aspects of evaluations. These measures were instead thought to complement what could be assessed

via self-reports, giving a fuller picture of one's (traditionally conceived) attitudes.

This new logic surrounding implicit measures is problematic, as there is no reason to assume that responses on implicit measures cannot be influenced by traditional attitudes; that is, attitudes that can be accessed via self-report (or that might be measured via many of the other methods just noted). Consider the example of a researcher studying anti-fat bias. If one of the study's research participants knowingly embraces an active dislike for people who are overweight, it seems reasonable to predict that this consciously prejudiced individual would have more difficulty quickly associating the concept "fat" with the concept "good" than with the concept "bad," a task that is core to the Implicit Association Test (IAT) for the assessment of implicit attitudes. Simply put, one cannot assume the attitudes detected by an implicit inventory are exclusively automatic, spontaneous, unconscious, or that they possess any other distinct cognitive qualities simply because the method used to quantify scores do not rely on self-reports (Gawronski, LeBel, & Peters, 2007). This first justification, thus, is problematic.

A second common justification is the stated desire to have inventories that are immune to self-presentational biases. Distinctions between the type of attitude assessed via one or another method are no longer relevant. Instead, this justification becomes relevant when one seeks to assess (possibly traditional) attitudes in a manner that will avoid social desirability artifacts. The possibility that self-reports might be distorted by self-presentation biases was a concern of attitude researchers long before the push to developing implicit measures. It has been widely accepted among attitude researchers that people can control what they report on explicit inventories. As such, people might distort their answers on survey questions in order to present themselves in a positive light. Consider as one example the taboo against admitting common prejudices. When asked, many respondents might deny anti-fat attitudes they know they possess as they do not want to project a negative (but accurate) impression to the researcher. The hope with implicit measures is that the measures invoke processes that individuals have limited ability to control and so implicit inventories might reveal attitudes respondents would not willingly report.

A researcher seeking implicit measures based on the logic of the second justification should consider that at times it contradicts the logic of the first justification. If a researcher is worried about eliminating social desirability artifacts that influence self-report, then the researcher also assumes that the attitude of interest is one that is consciously held and controlled by respondents. If respondents seek to obscure their true prejudice against people who are overweight, for instance, then they must be aware of the attitudes they hold. By acknowledging the possibility of positive response bias, the researcher is acknowledging the conscious nature of the attitude. It is, of course, possible that the researcher wishes to assess attitudes that

are more automatic and spontaneous—but not necessarily unconscious. If this is the case, the first and second justifications need not conflict with one another. One might, for instance, want to assess attitudes that are triggered spontaneously by an attitude object in a manner that overrides social desirability constraints.

Setting aside the potential logical inconsistency between the first and second justification, empirical research has offered equivocal evidence that implicit attitude measures are in fact immune to self-presentation. Although there is evidence that the motivation to control prejudice scale (Dunton & Fazio, 1997) moderates implicit–explicit correspondence (e.g., Cameron, Brown-Iannuzzi, & Payne, 2012; Gawronski, Geschke, & Banse, 2003), there is little evidence that the correlation between implicit and explicit attitude measures increases the less people care about socially desirable responding or that it varies across as a function of the magnitude of the social desirability concerns that linked to the criterion (Egloff & Schmukle, 2003; Hofman, Gawronski, Gschwendner, Le, & Schmitt, 2005; see also Hofman, Gschwendner, Nosek, & Schmitt, 2005; Nosek, 2005). Moreover, research conducted on variety of implicit measures suggests that people are able to exert a degree of cognitive control over their responses to implicit measures, suggesting that self-presentational motivations might affect scores on implicit as well as explicit measures (see Gawronski et al., 2007, for a full review). Thus, although it is intuitively compelling that implicit measures might eliminate social desirability concerns, evidence is mixed and suggests at best a more complicated picture.

Recommendations. We have identified two reasons researchers often give for using implicit measures: 1) to measure a different type of attitude that explicit measures cannot assess, and 2) to measure the same type of attitudes but to do so free of impression management concerns. If one's conceptual rationale is the former, then a researcher should do more than simply give scores on an implicit measure a distinct label (such as labeling implicit scores as reflective of automatic or unconscious attitudes). A better strategy is to statistically remove the contaminating effects of explicit attitudes on the implicit measure so that it more purely reflects the unconscious facets of attitude not available to declarative or working memory. Thus, when using implicit measures to predict meaningful criteria, one should also measure explicit attitudes and then statistically hold such attitudes constant (i.e., treat them as covariates) during the modeling process. This approach can ensure that the variation being modeled by the implicit attitude measure is due to a construct that is distinct from what might be assessed via explicit measures. This approach, of course, is not without limitations, as it assumes there are no other sources of measurement error in the implicit measures. Challenges 5 and 6 introduce caveats along these lines. In theory, however, a strategy based on statistically controlling explicit attitude variability in the criterion is sound for many of the questions implicit attitude theorists seek to answer.

If one's conceptual rationale instead is to avoid social desirability responding found in explicit measures, implicit attitude measures might offer some insights but given the state of evidence, we suggest augmenting this approach with strategies for improving the performance of explicit measures. Attitude theorists have evolved a set of standard procedural practices that, notably, are not used in most implicit attitude research but that reduce good impression biases on explicit attitudinal responses (Jaccard & Wan, 2006). These include:

- 1) instructing participants that responding honestly is important to the integrity of the research;
- 2) telling participants that you would rather they not respond to a question at all than to answer it untruthfully;
- 3) enhancing the motivation to respond honestly by informing participants that in order for the research to make a difference, people need to be truthful;
- 4) ensuring respondents of the confidentiality of their responses by assuring them that their names will in no way be connected to the responses they give;
- 5) stressing that data will only be reported in terms of groups of people and that individual responses will not be reported;
- 6) never having respondents report their (explicit) attitudes face-to-face to another individual, instead recording answers in private on self-administered questionnaires or on computers; and
- 7) formally signing an honesty pledge stating they will provide truthful answers. In addition, researchers can administer a social desirability response scale and include this measure as a covariate during the modeling process.

Because most measures of social desirability responding have dubious psychometric properties, we recommend the measures described in Dodaj (2012) for this purpose. Granted, researchers may still feel that an implicit measure is necessary to circumvent social desirability bias. But if explicit attitude measures are to be used in any way in one's research, it is good methodological practice to adopt the above standards when using those measures. Current implicit attitude research does not, so it is not surprising their explicit attitude measures perform badly in theory tests.

Challenge 2: Resolve the Lack of Convergence among Implicit Attitude Measures

The Challenge. Social psychologists excel at using subtle manipulations to generate nonobvious responses (Prentice & Miller, 1992). The rapid development of many new implicit attitude measures is a reflection of this type of

creativity. Researchers have identified a host of experimental manipulations that lead to replicable differences in how individuals categorize stimuli, resolve lexical decisions, and how they respond to a variety of other cognitive tasks. Although the creative use of laboratory tasks has been a feature of social psychology since its inception, the vigorous use of these strategies as a means of quantifying individual differences in attitudes is relatively new.¹ The rapid expansion of new implicit measures in the field reveals some of its creativity, but it also introduces concerns about convergent validity.

A hallmark of a psychometrically sound measure is that it should exhibit convergent validity with other measures that index the same construct, that is, the various measures should sort individuals in roughly the same fashion. This has not been the case with different measures of implicit attitudes. Convergence coefficients between them are typically in the range of 0.00 to 0.30. For example, correlations between the IAT and evaluative priming methods have typically been about 0.05 (Fazio, 1999; Sherman, Rose, Koch, Presson, & Chassin, 2003). The correlation between measures based on evaluative priming and lexical decision making has been about 0.25 (Brauer, Wasel, & Niedenthal, 2000). In a study using the four most common implicit attitude inventories to measure attitudes toward alcohol use in college students, Burrows, Jaccard, and Blanton (2015) found the degree of association between measures ranged from $r = -0.08$ to $r = 0.36$, with a median of $r = 0.13$.² Typically, correlations between two valid measures of the same construct should be above 0.70. These low convergence correlations suggest that the results one will obtain when working with implicit attitudes can depend on the specific measure used.

Recommendations. This issue is not one that any given media psychologist will likely solve in a single study, and we do not task media researchers with solving this problem in the pursuit of other questions they are interested in. But, the field as a whole needs to attend more to the lack of convergence among implicit attitude measures. One frame on the issue is that only one of the measures currently used is valid, so research must determine which measure it is. Another frame is that each measure taps into a different attitudinal facet and that the field needs to figure out which facet corresponds to which measure. Still another frame is that each measure is so strongly dominated by method variance that the influence of actual implicit attitudes on them is weak, at best.

Rather than taking the extant measures as givens and then trying to sort out what each measure reflects conceptually, we believe a healthier perspective is for the field to specify first theoretically meaningful implicit attitude constructs and then derive measures that will tap into those constructs. Let theory drive measurement rather than measurement drive theory. This perspective permits embedding the constructs into a larger nomological network, making it easier to evaluate measure validity by determining

if the measure correlates with constructs it should and does not correlate with constructs it shouldn't. It is beyond the scope of this article to suggest implicit attitudinal facets, but traditional attitude theory can be a guide. For example, distinctions have been made between attitude strength, importance, certainty, complexity, and salience (Petty & Krosnik, 1995). Each of these facets has a potential implicit counterpart that one might assess.

For media psychologists who do not wish to pursue formal research on convergent validity of implicit measures, we suggest a more modest research practice. We suggest that researchers, at a minimum, obtain multiple measures of implicit attitudes in any given study so that the robustness of conclusions across different implicit measures can be assessed (after statistically controlling for explicit attitudes, per Challenge 1). Given the existence of empirically distinct measures that lack clear theoretical distinctions, researchers should also consider adjusting for inflated familywise error rates when evaluating the statistical significance of links between multiple implicit measures and substantive outcomes and they should seek replications of effects that appear on some measures but not others. Rather than glossing over different results for different measures, such results should be embraced so that the field can move forward, however tentatively.

Challenge 3: Applying Scaling Models to Implicit Measures

The Challenge. In 1928, Leon Thurstone published a classic paper, "Attitudes can be measured" in which he articulated a framework for attitude scale construction. This led to the development of a number of scaling approaches that dominated the attitude area for the next 60 years, including Likert scaling, Guttman scaling, several variants of Thurstone scaling, and scaling models based on the semantic differential. These methods typically involve the generation of an initial set of items for possible inclusion on the final scale. A formal model that specifies the presumed relationship between a person's true attitude and the probability of endorsing an item is posited. Data are next collected to evaluate if a given item conforms to this model and items that do not are eliminated from the final attitude scale. Once a final set of items is identified, a scoring algorithm is specified for combining item responses into an overall score that is presumed to reflect the person's attitude. These steps constitute what roughly can be termed a "scaling model." The steps used to convert item-level responses into attitude estimates are empirically grounded and model driven. Subsequent to the application of the scaling model to create the scale, the measure is subjected to additional empirical evaluations of its reliability, validity, and metric properties. For example, research might be pursued that evaluates the measure's reliability using test-retest strategies and its validity using strategies related to construct validity and convergent validity.

One way in which attitude scaling methods differ from one another is the facet of the scaling model that relates true attitudes to the probability of endorsement of an item. The relationship between the true attitude and the probability of endorsement of an item is called the *traceline* or *item operating characteristic* of the item. In classic Likert scaling, the traceline is assumed to be linear—the more positive the true attitude, the greater the probability of endorsing a positive item and the lower the probability of endorsing a negative item. This scaling model leads to the use of item-total correlations as a means for item selection as well as the use of factor analysis to evaluate unidimensionality. By contrast, both Thurstone and Guttman scales assume nonlinear tracelines, so alpha coefficients, factor analyses, and the analysis of item total correlations is not appropriate for them. It is not uncommon for researchers performing theory tests to conduct empirical checks of the properties of the scales they use to ensure that the scales are behaving properly for the target population and empirical context being studied. Researchers typically report alpha coefficients, coefficients of reproduction, and factor analyses of items to affirm the psychometric properties of the scale.

Contrast this approach to that taken with implicit attitude measures. In this domain, researchers create and combine responses absent scaling models and aggregate items without empirically evaluating item properties (Fazio et al., 1995; Greenwald et al., 1998; Nosek & Banaji, 2001; Olson & Fazio, 2014; Payne et al., 2005). To illustrate the norms in this new measurement tradition, imagine the process by which a scale would be created by a researcher interested in adapting the most commonly used implicit attitude measure, the IAT, to assess implicit anti-fat attitudes. The IAT has individuals react to stimuli shown one at a time on a computer screen. The individual looks at a stimulus and then classifies it as quickly as possible into different categories. Some of the stimuli are photos and some are words. Individuals might be shown a total of 120 stimuli, with the time it takes to classify a given stimulus on a given trial recorded. There are four classes of stimulus pictures or words that the researcher requires: fat, thin, good, and bad. The former two classes are represented by pictures of heavy or thin people and the latter two classes by affectively valenced words that are either positively or negatively charged. Individuals are instructed to perform two categorization tasks. We refer to the first task as the *compatible task* (because researchers expect it to be easier for those with anti-fat bias). It requires participants to correctly categorize a shown stimulus into one of two categories, either “a thin person or a positive word” versus “a fat person or a negative word.” We refer to the second task as the *incompatible task* (because researchers expect it to be harder for those with anti-fat bias). It requires participants to correctly categorize each of the same stimuli into the two categories “a thin person or a negative word” versus “a fat person or a positive word.” Each trial latency is analogous to an item on a traditional explicit attitude scale. The items are aggregated by averaging, separately,

the response latencies for compatible trials and the response latencies for the incompatible trials. These two average latencies are then differenced, subtracting the average compatible latency from the average incompatible latency. Note that no item-level analyses are performed on the chosen photos or words to determine if a given stimulus word or stimulus picture is better or worse for purposes of measuring implicit attitudes. The choice of the stimuli is left to the discretion of the researcher with little psychometric guidance.

The developers of implicit attitude measures generally do not make explicit the scaling model they invoke, so one is left to speculate about the operative tracelines. In the case of the IAT, the true implicit attitude is actually a relative attitude that reflects the difference between the implicit attitude toward thin people minus the implicit attitude toward obese people. The more positive this relative attitude, the slower should be the average response latency of the incompatible trials (a linear traceline) and the faster should be the average response latency of the compatible trials (also a linear traceline). This scaling model leads to the prediction that there should be a nontrivial negative correlation between the average response latency for the incompatible trials and the average response latency of the compatible trials, everything else being equal. A formal test of the viability of the scaling model is, thus, possible through examination of this correlation. In our research with the IAT (Blanton et al., 2006), we have rarely observed a negative correlation between the two latencies (even after adjusting for overall processing speed), nor have we found support for other predictions made by the above scaling model. The more general point we make is that unlike traditional scaling methods applied to explicit attitude measures, when researchers use implicit attitude measures, rarely do they formally empirically evaluate the viability of the scaling model that underlies those measures. Instead, validity is assumed.

Recommendations. When using an implicit attitude measure, researchers should attend to the psychometric properties of the measures relative to the underlying scaling model. This is especially true for measures where the choice of stimuli (e.g., photos, words) are left to the discretion of the investigator. Just as it has become routine to report alpha coefficients, examine interitem correlations, and to conduct preliminary analyses of items to ensure unidimensionality for explicit attitude scales, so also should it be routine to examine the psychometric properties of implicit attitude measures as part of a research project. If one views each trial on multitrial measures as an item, it makes sense to perform item analyses to examine if the trial responses conform to the underlying scaling model (e.g., Payne, Hall, Cameron, & Bishara, 2010). A problem with doing so is that the developers of implicit attitude measures rarely make explicit the scaling model they adopt.

If we work within the structure of scaling models that assume linear tracelines, then it is reasonable to conduct item level analyses that examine item-total correlations and item intercorrelations.³ For example, for the IAT, one might report the average within-block item-total correlation for the in-

compatible block of trials and for the incompatible block of trials, respectively. One also can report alpha coefficients or coefficients of reproduction for each block, although these indices inflate with moderate to large numbers of trials so they should be supplemented by characterizations of item-total correlations. Factor or cluster analyses of the individual trial scores within and across blocks can reveal sources of method variance (e.g., factors associated with a specific photo or word) and these also should routinely be examined, sample size permitting. In neurological research that has examined the event-related potentials (ERP) activated when individuals take the IAT designed to measure the relative preference for Whites relative to Blacks, Hilgard, Bartholow, Dickter, and Blanton (in press) found evidence that different control demands are placed on individuals when they categorize Black as opposed to White stimuli (see also Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005). Item-level analyses might help identify photographic stimuli that are less susceptible to such effects. Most implicit attitude measures use difference scores and many, like the IAT, assume a scaling model that predicts a strong negative correlation between the components of the difference scores. Such psychometric evaluations also should be undertaken and reported.

In sum, there is evidence questioning current scoring algorithms used with the most popular of the new implicit measures, the IAT (Blanton & Jaccard, 2015) and there has been no research directly addressing the appropriateness of algorithms for other measures. Researchers should apply the same psychometric standards to implicit attitude measures that we have routinely applied to explicit attitude measures for the past 50 years and this involves routine checks on unidimensionality, interitem correlations, and item-total correlations, as appropriate.

Challenge 4: Using Multiple Single Attitudes Rather Than Forced Composite Relative Attitudes

The Challenge. Most implicit attitude measures are designed to assess relative implicit attitudes rather than an attitude toward a specific attitude object. With an IAT task, for instance, researchers examine the speed with which obese is correctly categorized as contrasted with some opposing concept, like thin. Similarly, it is necessary to contrast a positively valenced word (e.g., good) with a negatively valenced word (e.g., bad). This also is true of most other implicit attitude measures in the field because they embrace procedures that are presumed to reflect contrasting evaluations of contrasting attitude objects. Such approaches constrain the types of theories one can test. Consider again our anti-fat IAT. If this measure accurately represents the construct it is designed to measure, then the true (unmeasured and unobservable) implicit relative attitude (T_{IRA}) that a researcher taps into is a relative preference for thin over fat. This can be represented as the

difference between two distinct implicit attitudes:

$$T_{IRA} = T_{IA-F} - T_{IA-T}, \quad (1)$$

where T_{IA-F} is the true implicit attitude toward the concept fat and T_{IA-T} is the true implicit attitude the concept thin. Suppose we are interested in predicting a criterion, Y , from this relative attitude. A traditional linear model expressing the criterion as a linear function of the relative implicit attitude has the form:

$$Y = \alpha + \beta T_{IRA} + \varepsilon. \quad (2)$$

If we substitute the right hand side of Equation (1) for T_{IRA} in the above equation and expand the products, it can be seen that we actually are modeling:

$$Y = \alpha + \beta T_{IA-F} - \beta T_{IA-T} + \varepsilon. \quad (3)$$

Equation (3) reveals a conceptual model that is restrictive in form. It asserts that the impact of the two implicit attitudes on the criterion is additive and that the effects of each implicit attitude (as reflected by β) are equal in magnitude but opposite in sign. To the extent that this causal structure is not operating, Equation (3) will result in suboptimal prediction of the criterion and can yield inappropriate inferences.

Consider, as an example researchers who seek to relate implicit attitudes to prejudicial actions on the part of employers. Of interest is whether implicit attitudes predict rejection of overweight job applicants. It may be that bias of this form is driven more by implicit disliking of obese individuals (T_{IA-F}) than implicit liking of thin individuals (T_{IA-T}). The true (unmeasured) regression coefficient for T_{IA-F} will be larger than the (true unmeasured) regression coefficient for T_{IA-T} , making Equation (3) a misspecified model. In such a circumstance, the regression coefficient one obtains from regressing prejudicial actions onto an overall index of relative preference, T_{IRA} , will be an intermediate value between the (relatively large) coefficient one would have obtained from a valid measure of T_{IA-F} and the (relatively small) coefficient one would have obtained from a valid measure of T_{IA-T} . As another example, suppose that negative evaluations of the concept fat lead to more prejudicial actions against the obese, the more positively people evaluate the concept thin. This implies an interaction between T_{IA-F} and T_{IA-T} . This causal dynamic cannot be detected with the model in Equation (3).

By conflating the two implicit attitudes, the IAT lacks the needed flexibility to model more nuanced underlying psychological dynamics. Similar issues arise because the measure also simultaneously taps into positive and negative associations. Perhaps employment bias is driven more by negative than positive associations or perhaps these two associations interact dynamically to shape such decisions. If any of these processes occur, the anti-fat IAT

will invoke a misspecified causal model, even if it operates from a correctly specified measurement model.

Recommendations. If one is truly interested in relative implicit attitudes and one can make the strong assumptions implied by Equation (3), then many of the standard implicit attitude measures are appropriate. If researchers are more interested in understanding how multiple attitudes combine to impact or predict criteria without making such assumptions, then they will need an implicit attitude measure that is better suited to doing so, such as the Simple Association test suggested in Blanton et al. (2006).

Challenge 5: Measurement Correspondence

The Challenge. In a now classic paper, Ajzen and Fishbein (1977) presented a penetrating analysis of the commonly observed lack of correlation between (explicit) attitudes and behavior. In one part of their analysis, they argued that a given behavior has different core elements, including the action itself, the setting in which the action takes place, the timing of the action, and the object that the action is directed toward. For an attitude to predict behavior optimally, they argued, the core facets of the measured attitude must be correspondent with the core facets of the behavior. For example, if one is trying to predict if a man will use condoms with his regular partner during the next month, then the measured attitude that should be most predictive is one toward using condoms with one's regular partner in the next month. As the attitude measure deviates from this correspondence (e.g., if one measures the attitude to use condoms in general), then behavioral predictability can be adversely affected. The empirical evidence for this principle of specificity/correspondence is both considerable and impressive.

Most research that uses implicit attitudes to predict behavior does not use behavior-specific attitudes. In many respects, the approach taken for predicting behavior from implicit and explicit attitudinal constructs in implicit attitude studies violates almost all of the well-known attitude-behavior principles that have evolved from the influential work of Ajzen and Fishbein (1977). Consider the implications of correspondence for a researcher studying anti-fat implicit attitudes. Most implicit inventories measure general implicit attitudes toward being fat and being thin. However, anti-fat bias can take many, often independent forms. One individual might hire overweight individuals and be happy to have such individuals as friends but be unwilling to become romantically involved with someone who is overweight. Another individual might be willing to hire an overweight individual but would not be comfortable being connected to that person outside the office. Another individual might be willing to have overweight friends but not employees. Independent of the anti-fat attitudes that might influence

interpersonal reactions, people can differ in the anti-fat attitudes that shape their own self-conceptions and dieting habits. The implicit attitude measures that will be most predictive of these various behaviors are those that take into account the nuances of those behaviors. A single general measure of implicit attitude toward obesity and thinness does not hold much promise, based on everything we know from attitude research over the past 30 years. A focus on general implicit attitudes suggests that criterion prediction will be weak in most instances (see Oswald et al., 2013).

Recommendations. Concerns for measurement correspondence have implications for both the assessment of implicit attitudes and the assessment of explicit attitudes used as covariates in research studies (per Challenge 1). With regards to implicit attitude measurement, future research in the field needs to develop methods that permit investigators to assess implicit attitudes in ways that incorporate the four core facets of behavioral criteria, allowing for correspondence. When using an implicit attitude measure, researchers should consider the four facets of their behavioral criteria and the implications of generalizing across (or ignoring) each facet. Researchers might then adapt an implicit measure to accommodate correspondence considerations. For example the photographic stimuli used in the various classification tasks of the IAT can be selected in ways that take into account all four facets of the behavior to be predicted instead of just the object of the behavior per se. It is not clear if such changes will lead to better behavioral prediction from implicit constructs, but decades of attitude research suggests that it will (see Payne, Burkley & Stokes, 2008, for some evidence of this with implicit measures).

With regards to explicit attitude measurement, it is important that researchers compare the predictive utility of implicit measures relative to explicit attitude measures that incorporate correspondence principles. As noted by Oswald et al. (2013), too often implicit measures are evaluated relative to crude, outdated, general attitude measures (e.g., “feeling thermometers”). It should not be surprising that such crude explicit measures have little predictive utility. Evidence that implicit attitudes predict criteria independent of these crude measures is not particularly informative. A researcher interested in understanding the influence of implicit attitudes on criteria/outcomes should develop explicit measures as if they want these measures to do the best job they can at predicting the criteria/outcome. This means using explicit measures that give full consideration to the issue of measurement correspondence.

Challenge 6: The Problem of Left Out Variable Error

The Challenge. When testing causal models, a well-known source of bias is that of left out variable error, also known as L.O.V.E. (Mauro, 1990). Suppose that in a population, an outcome variable, Y , is a linear function of

two correlated variables, X and Z, such that

$$Y = \alpha + \beta_1 X + \beta_2 Z + \varepsilon.$$

The β coefficients reflect the presumed causal impact of each predictor on Y in the sense that they reflect the number of units Y is predicted to change given a one unit increase in the predictor (holding constant the other predictor). Least squares estimation of these coefficients takes into account the correlations among the predictors, X and Z. Left out variable error occurs when we estimate the effect of a predictor on Y but we omit a key predictor from the estimating equation. For example, we might mistakenly estimate the causal coefficient for X using the model

$$Y = \alpha + \beta_3 X + \varepsilon.$$

In this case, the estimate of β_3 will be a biased estimate of the true causal coefficient β_1 because it does not correctly take into account the effect of Z on Y nor the correlation between X and Z. We have a L.O.V.E. problem.

The vast majority of studies of implicit attitudes that predict meaningful behavioral criteria estimate the effects of the IAT on those criteria without regard to L.O.V.E. (e.g., Rudman & Ashmore, 2007). This is problematic because it can result in biased inferences. Research with implicit attitudes tends to use just one additional predictor, if any, beyond implicit attitudes when estimating causal coefficients, namely a single explicit attitude measure. Returning to the anti-fat example, think of the many constructs that might predict prejudicial reactions to obese job applicants independent of implicit evaluations. These include but are not limited to descriptive and injunctive norms regarding the acceptability of acting on bias, the motivation to control prejudice, attributions for why people become overweight, justice beliefs, and so on. Failure to control for these variables introduces L.O.V.E. into the modeling process. To complicate matters further, implicit attitude researchers rarely control for explicit attitudes in a manner that honors the measurement correspondence that is easily achieved with explicit attitude/specificity (see Challenge 5). Thus, even though an explicit attitude measure is included in the model, it is the wrong attitude, so L.O.V.E. problems exist.

Recommendations. Researchers using implicit attitude measures will benefit from paying greater attention to matters of L.O.V.E. in the context of their modeling. Of course, L.O.V.E. plagues virtually every model researchers work with because it is impossible to specify and measure all of the variables that impinge on a criterion in a way that causes bias in estimates of causal coefficients. This is a particularly important issue in the implicit attitude area, however, because the mere presence of a statistically significant zero-order correlation between implicit measures and criteria is commonly interpreted as evidence of a new class of predictor variables; one that cannot be accessed

via self-report (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Such conclusions are not warranted because many known or plausible explicit causes have not been measured and included in the modeling process. We suggest as a practical solution to the problem of L.O.V.E., that researchers a priori identify what they believe to be the most important predictors to include and then measure and include them as covariates during statistical modeling (giving careful consideration to measurement correspondence with each inventory). For identifying these variables, we suggest drawing on modern attitude-behavior models as a starting point (e.g., Ajzen & Fishbein, 2005). For a discussion of conditions where L.O.V.E. is of most importance, see Mauro (1990).

Challenge 7: Addressing Random Error

The Challenge. Measurement error is a fact of life for explicit attitude measures, just as it is for implicit attitude measures. However, research suggests that implicit measures are more strongly affected by both random error (noise) and systematic error (confounds) than many traditional explicit measures. We consider here random error and then in the next challenge turn to the problem of systematic error.

Random error causes measured values to vary over repeated administration despite the constancy of the attribute or quantity being assessed. Consider as an example a measure of racial attitudes. Racial attitudes presumably develop over a lifetime of socializing influences that ultimately lead to the internalization of (conscious or unconscious) beliefs about members of different races. One would, therefore, not expect a construct like racial prejudice to change much over a short period of time. However, in the most extensive study of the test-retest reliability of the IAT designed to measure anti-black racial implicit attitudes, Cunningham, Preacher, and Banaji (2001) observed a test-retest correlation of 0.32 over four administrations of the IAT, each separated by just two weeks. Their estimate is lower than some estimates for this and other implicit measures, but it highlights the considerable amount of random error that can affect implicit attitude measures.⁴ Although implicit attitudes may remain constant over time, respondent interest, motivation, attention span, and other factors that influence experimental responsiveness can vary considerably, leading to seemingly random shifts in attitudes from one testing session to the next.

There are actually two types of random error that affect multitrial implicit attitude measures. For any given block of trials of the same type, there will be random noise or nonsystematic error that causes a given individual to respond faster or slower to a given trial within that block. Such trial error represents nonsystematic error. Trial error might occur because a particular photo on a given trial is an unusually attractive or unattractive person or because momentary distractions in the testing environment influence laten-

cies on that trial. Perhaps, a loud noise occurs in the room that distracts the person just as the photo appears on the computer screen, thereby slowing his or her response. Perhaps, the person has a momentary distracting thought just when the stimulus appears. The second type of random error is called transitory error because it has the same impact on every trial during one testing session, but a different impact on trials in another testing session conducted at a different point in time. For example, if on one day a person is in a good mood, then performance on latency-based tasks might differ as compared to another day when that person is in a bad mood. Both trial and transitory errors need to be addressed because they create noise in the data.

Recommendations. To reduce trial error, one obvious strategy is to assess implicit attitudes in conditions where there are as few outside distractions as possible. Research often falls short on this recommendation, as researchers commonly conduct implicit attitude studies online in settings where researchers have less control over the testing context. To the extent that this introduces distractions into the testing environment, trial error will be more problematic.⁵ A second strategy is to carefully pretest the stimulus materials (photos, words) used during the assessment task and to eliminate those that have large amounts of trial error associated with them (using the item analysis strategies discussed above). A third potentially useful strategy is to use latent variable modeling in structural equation modeling (SEM) that can formally take measurement error into account when deriving parameter estimates. These approaches typically require multiple indicators of the same construct, which is problematic because of the low correlations between indicators of implicit attitude measures, noted earlier. However, through the creative use of parceling strategies, viable multiple indicators can be generated. For example, one might administer more trials than is typical and then randomly divide trials of a given type into three subsets. Response latencies are averaged within each subset, yielding three indicators of the construct reflecting by each subset. Alternatively, one can use single indicators but adopt a modeling strategy within SEM described by Burrows, Jaccard, and Blanton (2015) based on a priori constraints of error terms.

For transitory error, one strategy to accommodate it is to identify likely sources of such error (e.g., mood states, the time of day, the day of the week, daily life stressors, alertness, hunger, caffeine use), measure them, and then statistically control for them by using the measures as covariates during the modeling enterprise. One may not be able to measure all such sources, in which case researchers should prioritize them and measure the most important ones for purposes of including them as covariates. Transitory error also can be accommodated by making test-retest assessments of implicit attitudes on two occasions (e.g., separated by one week) and then using the two assessments as multiple indicators in a SEM model that has a stable latent variable reflecting the implicit attitude with these two indicators.

Challenge 8: Addressing Systematic Error

The Challenge. In our opinion, one of the most challenging features of implicit attitude measures is the heavy reliance of such measures on reaction time data (Payne et al., 2005). This is because the speed with which people respond to cognitive tasks is impacted by many variables, not just individual differences in implicit attitudes. In the prior section, we noted a host of transitory variables and trial error that can impact response latencies and that, ultimately, can undermine theory tests that rely on them. In this section, we address a second source of contaminating influence based in systematic error.

From data we have collected over the years, we estimate that for any given computer key press designed for use in an implicit attitude inventory, well over 50% of the variance between individuals in the response latency for making that key press is the result of systematic influences regardless of the nature of the item being presented; that is, whether categorizing a photo of an individual who is fat or thin and or hitting a key to indicate recognition of the word “plump” or “slight,” most of the differences in the speed between individuals is due simply to the fact that some people tend to respond faster than others to any cognitive task and that some individuals are more alert and attentive than others. We have referred to this variable as “general processing speed” (GPS; Blanton et al., 2006).

Put more simply, some people are faster to perform mental tasks than others, so the stable influence of GPS on response latency scores can produce reliable differences in how quickly people respond to experimental stimuli irrespective of their implicit attitudes. Implicit attitude researchers are well aware of the problem of GPS and have taken steps to try to eradicate its influence on measures. Most typically, the corrective involves some form of differencing between two averaged response latencies, with the idea that the differencing subtracts out (i.e., controls for) GPS. Recall that the anti-fat IAT is computed by differencing reaction times for two separate tasks, the compatible task and the incompatible task. Virtually all implicit inventories involve this type of contrasting of timed tasks through differencing. We illustrate the underlying logic using the IAT as an example. Let the (average) response latency for the incompatible trials be expressed as a weighted function of the tendency to respond quickly to incompatible trials because of implicit attitudes (T_{IC}) as well as GPS in the following equation:

$$\text{Incompatible Latency (ICL)} = a_1 + b_1 T_{IC} + b_2 \text{GPS} + e_1,$$

where a_1 is an intercept (constant) and e_1 is random measurement error. We can similarly express the (average) response latency for the compatible trials using the same equation form:

$$\text{Compatible Latency (CL)} = a_2 + b_3 T_C + b_4 \text{GPS} + e_2.$$

If we assume that $b_1 = b_3$ and that $b_2 = b_4$, then differencing the two equations yields

$$\text{ICL} - \text{CL} = (a_1 - a_2) + b_1 (T_{\text{IC}} - T_{\text{C}}) + b_2 (\text{GPS} - \text{GPS}) + (e_1 - e_2)$$

and GPS will drop out of the mix. Note that this only occurs, however, if the two tasks are influenced to an equal degree by GPS, that is $b_2 = b_4$. The subtraction method will fail if GPS exerts unequal influence on the two responses that are differenced. Blanton et al. (2006) found evidence against the assumption of equal b_2 and b_4 weights for the IAT and, indeed, Greenwald, Nosek, and Banaji (2003) offered a scoring algorithm, in part, as an explicit acknowledgment of the problem. Unfortunately, the new scoring algorithm adjusts for GPS in ways that introduce a host of new problems (see Blanton, Jaccard, & Burrows, 2015; Blanton, Jaccard, Strauts, Mitchell, & Tetlock, 2015).

The effect of GPS on other implicit attitude measures that rely on response latencies has not been examined in the detail that it has been addressed with the IAT, but due to the heavy reliance on a simple differencing strategy that is known to be insufficient with the IAT, it seems likely that this strategy is insufficient with other reaction time measures as well. Residual contamination of GPS in measures can be problematic when researchers use implicit inventories to test theories, as GPS is correlated (or confounded) with such factors as intelligence, age, alcohol use, a host of cognitive-deficit variables (such as attention deficit disorder), and even attitude constructs like prejudice (e.g., Blanton et al., 2006; Maylor, Rabbitt, James, & Kerr, 1992; Nebes & Brady, 1992; Nebes & Madden, 1988; Verhaeghen et al., 2002; Vernon & Jensen, 1984). As a result, researchers who examine media correlates with unadjusted implicit attitude scores might observe relations that are altered to some unknown degree by these contaminating third variables.

Recommendations. One way to avoid the contaminating effects of GPS and other factors that might be confounded with trial-level responding is to shift away from using reaction time measures and to assess implicit attitudes using other measurable responses.⁶ We believe this is a desirable direction, but even then, researchers must remain vigilant to other cognitive factors that can contaminate scores. Another strategy for accommodating GPS is to obtain an independent assessment of it and then to covary its effects out and adjust for it during the modeling process (see Blanton et al., 2006). This is a superior approach to seeking to address the GPS confound by creating new scoring algorithms that might expunge scores of this confound, as algorithmic creativity can introduce their own artifacts that researchers might need to address (see Blanton, Jaccard, & Burrows, in press).

Challenge 9: Using Implicit Measures as Dependent Variables

The Challenge. To this point, we have considered challenges facing researchers who wish to use implicit measures to predict criteria/outcomes. Quite often, however, implicit attitudes are the criteria researchers wish to predict. Rather than trying to predict prejudicial reactions of individuals based on their anti-fat attitudes, for instance, a researcher might instead be interested in determining if media exposure to the thin ideal leads to an increase in implicit anti-fat attitudes. Research suggests that implicit inventories that ostensibly measure stable attitudes are quite responsive to media-based manipulations (Dasgupta & Greenwald, 2001), so use of implicit attitude inventories as dependent variables might reveal meaningful effects that explicit measures fail to uncover.

There are nontrivial modeling issues that need to be taken into account when using relative implicit attitude measures as outcomes, such as is the case for the IAT. To simplify discussion, consider a single predictor, X , and a criterion defined as $Y_1 - Y_2$ (representing, say, the difference between the attitude toward thin people and the attitude toward obese people). The linear equation is:

$$Y_1 - Y_2 = \alpha + \beta_t X + \varepsilon. \quad (4)$$

Instead of the above equation, one can specify two-component linear equations, one regressing Y_1 onto X and the other regressing Y_2 onto X . These are

$$Y_1 = \alpha + \beta_1 X + \varepsilon \quad (5)$$

$$Y_2 = \alpha + \beta_2 X + \varepsilon. \quad (6)$$

Kenny and Judd (1996) have shown that β_t in Equation (4) is a simple function of β_1 and β_2 in the component equations, namely,

$$\beta_t = (\beta_2 - \beta_1).$$

The coefficient for X when predicting a difference score per Equation (4) is the slope difference between β_1 and β_2 and its test of significance is a test of slope inequality.

If we only have an estimate of β_t , there is an infinite number of possible values of β_1 and β_2 that could produce it. It is only when an analytic strategy is pursued that allows one to estimate β_1 and β_2 vis-à-vis Equations (5) and (6) that the underlying β_1 and β_2 values become evident. In this sense, it is more informative to conduct analyses to isolate estimates of β_1 and β_2 in the spirit of Equations (5) and (6) as opposed to β_t in Equation 4. The analysis of Equation (4) is only useful if one desires to test the hypothesis that the impact of X on Y_1 is the same as the impact of X on Y_2 .

Consider how this analysis translates to the study of media influences on anti-fat implicit attitudes as measured by the IAT. Suppose a researcher exposes half of the participants to thin media images and half to unrelated images and then administers the anti-fat IAT described earlier. The researcher tests for media effects by regressing IAT scores onto the dummy variable for condition (coded 0 and 1).⁷ The IAT outcome represents a difference between two implicit attitudes, the attitude toward thin people (Y_1) minus the attitude toward obese people (Y_2). The regression analysis (or the t test comparing the mean IAT scores) essentially tests if the effect of the manipulation on the implicit attitude toward thin people is the same as the effect of the manipulation on the implicit attitude toward obese people. We doubt that this is what the researcher is truly interested in. Rather, of interest is a) the extent to which the manipulation affects the implicit attitude toward thin people in its own right, and b) the extent to which the manipulation affects the implicit attitude toward obese people in its own right.

Recommendations. The use of differences scores in modeling has a long and controversial tradition in the social sciences, with a great deal of confusion surrounding it. We believe that media researchers will best be served by avoiding the use of relative implicit attitudes as outcomes (unless theory dictates otherwise) and instead use implicit attitude measures that tap into relevant attitudes toward single attitude objects (e.g., the implicit attitude toward thin people; the implicit attitude toward obese people). Measurement strategies using the Simple Association Test, for example, are preferable to those of the IAT because the former allows one to measure single attitudes rather than relative attitudes. In this way, researchers have greater flexibility in modeling the theoretical questions of interest.

Challenge 10: Making Arbitrary Metrics Nonarbitrary

The Challenge. In some cases, researchers use implicit attitude measures—neither as predictors nor as criteria or outcomes—but as metrics that can gauge implicit preferences and the degree of implicit bias. For example, based on the distribution of IAT scores found in most studies, researchers have concluded that the majority of Americans harbor racial, ethnic, sexual biases, and all manner of other prejudicial implicit evaluations. We suspect that media psychologists may use implicit attitude measures for similar purposes, such as documenting the magnitude of anti-fat bias in a population of interest.

Despite the widespread use of implicit measures to estimate mean implicit preferences and prejudices, the practice is problematic because implicit attitude measures have arbitrary metrics (Blanton & Jaccard, 2006a). The term *metric*, as used here, refers to the numbering system that the researcher employs when describing individuals' standings on the construct of interest. With arbitrary metrics, it is not known exactly where a given score locates

an individual on the underlying psychological dimension, nor is it known how a one-unit change on the observed score corresponds to the magnitude of change on the underlying dimension. Imagine for purposes of illustration that one has a perfect, completely unbiased measure of how people treat obese and thin individuals in their everyday interactions. This perfect, error-free measure can place an individual on a single response metric that ranges from -2 (strong behavioral preference for thin people) to $+2$ (strong behavioral preference for obese people), with a zero point that reflects true neutrality (no preference at all). Further suppose a researcher has access to a population of individuals where their scores on this metric are known. We refer to these as “true scores” because they reflect the actual behavioral tendency of the individuals. The researcher now administers the anti-fat IAT to each individual, using the current D score algorithm (Greenwald et al., 2003). This scoring procedure also has a -2 to $+2$ metric, where a score of 0 is said to reflect neutrality.

In the implicit attitude field, it is common to assume that the zero-point derived from computing a difference score for the observed scores is the true dividing line between preferences. With an IAT, for instance, it is common to assume that a value of zero on the observed metric maps onto the value of zero for the true scores. With the anti-fat IAT, for instance, people who have positive IAT scores are assumed to be biased in the favor of thin people, people who have a negative IAT score are assumed to be biased in the favor of obese people, and people who have values near zero are assumed to be mostly neutral. In this example, the researcher has access to the true, unbiased estimate of actual bias, so in this imagined world it would be possible to empirically determine if the IAT zero point maps onto the true zero point. The researcher could compare each individual’s true bias score to their IAT score to determine if the IAT is overestimating bias or underestimating bias in the population. Of course in real life, it is not possible to conduct such tests. But it is possible to conduct research linking IAT scores to other observable forms of bias to see if the preponderance of evidence supports common interpretations of the IAT metric.

This approach was used by Blanton, Jaccard, Stuart, Mitchell, and Tetlock (in press). These researchers reanalyzed data from published studies linking IAT measures of implicit racial and ethnic bias to criteria reflecting differential treatment based on group membership (e.g., their treatment of Black versus White confederates in laboratory settings; Ashburn-Nardo, Knowles & Monteith, 2003; McConnell & Leibold, 2001). Of interest was how the zero point on the IAT lined up with the zero point on behavioral difference scores, for example, for those people who treated Blacks and Whites the same, did they have IAT scores equal to zero on the IAT designed to measure the relative preference for Whites relative to Blacks? Across studies using diverse methods and outcomes, Blanton et al. (2015) found evidence that this was not the case and that the IAT scores were consistently skewed in a way that

characterized people who showed no observable bias as being implicitly prejudiced. By extension and returning to our example, a researcher should be hesitant to infer that an individual who exhibits a positive score on an implicit measure of anti-fat bias will treat obese people any differently than they treat thin people.

Recommendations. Given an arbitrary metric, we do not know how scores on an implicit attitude measure map onto the true scores of underlying bias. It is impossible to equate the observed distribution with the underlying latent distribution without knowledge about how key benchmarks map onto the observed scores. For example, a benchmark of behavioral neutrality toward Blacks and Whites can be used to evaluate if the zero point of the implicit attitude measure reflects behavioral neutrality. A benchmark reflecting the typical percentage of Whites a manager hires over equally qualified Blacks can be mapped onto an implicit attitude score, thereby giving that score meaning, moving us closer toward a nonarbitrary metric. Given earlier discussions about the need for measurement correspondence, researchers should consider using benchmarks tied to the behavioral criteria they seek to predict and understand. The implicit attitude score that represents the dividing line between discriminatory and nondiscriminatory behavior against thin versus overweight employees might be different from the score that is the dividing line between having thin versus overweight romantic partners, and the instrument that is best for assessing the first form of bias might be different than the measure assessing the second.

The work needed to development meaningful metrics is extensive and must take into account the specific criterion, context and population one wishes to characterize. Given the lack of current research tackling this issue, at present, it is best for researchers simply to refrain from making statements of the amount of implicit bias that individuals have and statements about the true distribution of implicit attitudes. Instead, the field needs to start conducting research that will make the metrics of implicit attitude measures nonarbitrary, in the spirit of the work by Blanton, Jaccard, Stuart, et al. (in press) and by mapping important benchmarks onto specific implicit attitude scores. Until that work is done, we suggest researchers exhibit caution and not simply assert by fiat that individuals are biased based on their scores on response metrics. We argue that positive scores on measure of anti-fat implicit attitudes should be interpreted as revealing anti-fat bias only if other meaningful external criteria can be linked to test scores, to show that positive values merit such an interpretation.

CLOSING THOUGHTS

The 10 challenges we have outlined appear daunting and are likely to elicit different reactions from readers. On one extreme, readers might con-

clude that the challenges are overwhelming and that the pursuit of implicit constructs and implicit measurement should be abandoned. On the other extreme, readers might dismiss the concerns and challenges and simply follow current methodological norms. We think neither of these responses is good for our science. The 10 challenges we introduce reveal nontrivial complexities to the study of implicit attitude, but we have also offered constructive suggestions for addressing each of them. To be sure, some of the challenges will require creative thought and substantial research effort to adequately address. But in many respects, the challenges point to difficulties no more dire than the ones that faced explicit attitude research at the start of that measurement movement. In a now classic review of the attitude-behavior relationship, Wicker (1969) concluded that one should expect modest behavioral prediction from explicit attitude measures. This led to what is now commonly referred to as the "attitude crisis" in social psychology, an era in which researchers openly questioned the value of the attitude construct. However, rather than giving up or continuing along the same path, the field responded with constructive efforts to improve attitudinal concepts and measures so that they could do a better job of predicting and understanding behavior (e.g., Fishbein & Ajzen, 1975).

We believe that implicit attitude research is at a similar point. Evidence suggests that implicit attitude measures are not potent predictors of behavior (Greenwald, Banaji & Nosek, *in press*; Oswald et al., 2013, *in press*). Rather than abandoning implicit attitudes entirely or embracing them as they are now, we suggest that whenever research priorities call for their use, investigators approach implicit attitude measures with an eye for sensible ways to improve upon them so that their true potential can be determined. If one follows the recommendations we have made, then research projects that are pursued will look quite different than current studies. The implicit attitude measure will be subjected to extensive preliminary analysis in the context of a scaling model. The psychometric properties of the measure will be elaborated and documented. Measures of explicit attitudes and social desirability response tendencies will routinely be collected and controlled as covariates during statistical modeling. This will be coupled with procedural methods that reduce the likelihood of socially desirable responding. Multiple indicators of constructs will be included so that different sources of random and transitory error can be accommodated when deriving parameter estimates (through latent variable modeling). For implicit attitudes, multiple indicators will be sought, but with the recognition that different implicit measures may tap into different attitudinal facets, thereby resulting in conclusions that are measure dependent. An independent measure of general processing speed will be obtained and statistically controlled along with a host of other methodological and measurement confounds (described above) that can contribute to either inflated or diminished estimates of the effects of implicit attitudes. The research will use implicit measures

that do not dictate a conceptual focus on relative implicit attitudes (i.e., the difference between two implicit attitudes) but instead that allow the researcher to measure each relevant implicit attitude separately. The implicit attitude construct will be embedded within a broader causal theory that includes mediators, moderators, and direct effects that are of conceptual interest. In so doing, the matter of L.O.V.E. will be addressed by including covariates that, if otherwise omitted, are likely to cause bias in the parameter estimates. The explicit constructs included in the causal model will reflect modern day attitude theories (e.g., Fishbein and Ajzen's, 1975, reasoned Action Model) rather than outdated attitude approaches that were rejected as not viable over 40 years ago. For prediction studies, both the implicit and explicit measures will be mapped onto the four facets of the behavioral criterion (action, target, time, and setting) as closely as possible, unless theory dictates otherwise. Given multiple measures of implicit attitudes that are relatively uncorrelated, sensitivity analyses will test if conclusions are robust across the different measures, with adjustments for familywise error rates, as appropriate. The data will be analyzed using SEM in conjunction with sample sizes that support such analyses (which will invariably be larger than sample sizes typical of current implicit attitude research). SEM is applicable to both experimental and nonexperimental designs and represents a general framework that subsumes virtually all of the analytic strategies currently used in implicit attitude research, but more elegantly. Statements about the extent of bias or prejudice, if relevant, will be tied directly to benchmarks that have been empirically documented in prior work designed to address matters of metric arbitrariness.

The research practices we recommend are not common in implicit attitude research in social psychology or the other fields that are turning to these new inventories. We see no better place to start implementing them than research on media psychology.

NOTES

1. The use of laboratory tasks to scale meaningful individual differences can be traced at least back to Bruner's and Goodman's (1947) classic research examining how values and needs influence perceptions and to 1950, when Thistlethwaite (1950) sought to measure racial attitudes by assessing "distortions of reasoning" (p. 442). However, the popularity and enthusiasm for such approaches is a modern-day phenomenon.
2. We used the implicit association task (Greenwald, McGhee, & Schwartz, 1998), the personalized version of this measure (Olson & Fazio, 2003), evaluative priming (Fazio et al., 1995), and affect misattribution (Payne, Cheng, Govorun, & Stewart, 2004).
3. The influential measurement approach of item response theory tends to work with scaling models with nonlinear tracelines, and this is an alternative possibility.
4. Cunningham et al. also found an average test-retest for evaluative priming of $r = 0.25$. Burrows, Blanton, and Jaccard (2015) found two-week test-retest reliability estimates for the beer and water IAT of $r = 0.32$ and for beer AMP of $r = 0.37$. No data on the test-retest

- reliability of the personalized IAT were located. An exhaustive review of the test–retest reliability of all implicit measures is beyond scope of this article, but these findings suggest considerably more error in current implicit measures than standard explicit measures.
5. The new scoring algorithm built into the IAT (Greenwald et al., 2003) has a surprising artifact in that the more a researcher succeeds in reducing trial error, the more extreme respondents' IAT scores will be (see Blanton, Jaccard, & Burrows, 2015). We recommend not using the D score method as a result.
 6. Nonreaction time measures of implicit association do exist but in these measures, researchers again embrace scoring algorithms that focus on differencing (Payne et al., 2005; Pelham, Mirenberg, & Jones, 2002). It is unclear why this strategy has taken such a firm hold on implicit attitude measures but as a general strategy for addressing confounds, stronger methods can be used. In general, we recommend analytic strategies focused on the assessment of third-variable confounds that might influence experimental trials and then measuring these so that they can be statistically controlled.
 7. Inferences do not change if the researcher instead applies a *t* test to compare mean differences.

REFERENCES

- Ahern, A. L., Bennett, K. M., & Hetherington, M. M. (2008). Internalization of the ultra-thin ideal: Positive implicit associations with underweight fashion models are associated with drive for thinness in young women. *Eating Disorders*, 16, 294–307.
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888–918.
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Mahwah, NJ: Erlbaum.
- Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, 21, 61–87.
- Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal Of Personality And Social Psychology*, 62(6), 893–912. doi:10.1037/0022-3514.62.6.893
- Bartholow, B. D., Holgard, J., Dickter, C. L., & Blanton, H. (2015). *Characterizing reactive and proactive cognitive control processes in the implicit association task*. Unpublished manuscript.
- Berry, T. R., McLeod, N. C., Pankratow, M., & Walker, J. (2013). Effects of biggest loser exercise depictions on exercise-related attitudes. *American Journal of Health Behavior*, 37, 96–103.
- Bissell, K., & Hays, H. (2011). Understanding anti-fat bias in children: The role of media and appearance anxiety in third to sixth graders' implicit and explicit attitudes toward obesity. *Mass Communication and Society*, 14, 113–140.
- Blanton, H., Burrows, C. N., & Jaccard, J. (2015). *Implications of left-out variable error and measurement correspondence on estimates of implicit cognition effects on health behavior: The case of alcohol consumption*. Unpublished Manuscript.
- Blanton, H., & Jaccard, J. (2006a). Arbitrary metrics in psychology. *American Psychologist*, 61, 27–41.

- Blanton, H., & Jaccard, J. (2006b). Tests of multiplicative models in psychology: A case study using the unified theory of implicit attitudes, stereotypes, self esteem and self concept. *Psychological Review*, 122, 155–169.
- Blanton, H., & Jaccard, J. (2008). Unconscious racism: A concept in pursuit of a measure. *Annual Review of Sociology*, 34, 277–297.
- Blanton, H., Jaccard, J. & Burrows, C. (2015). Implications of the IAT D-transformation for psychological assessment. *Assessment*. doi:10.1177/1073191114551382
- Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Perspectives on criterion prediction. *Journal of Experimental Social Psychology*, 42, 192–212.
- Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009a). Strong claims and weak evidence: Reassessing the predictive validity of the IAT. *Journal Of Applied Psychology*, 94(3), 567–582. doi:10.1037/a0014665
- Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009b). Transparency should trump trust: Rejoinder to McConnell and Leibold (2009) and Ziegert and Hanges (2009). *Journal Of Applied Psychology*, 94(3), 598–603. doi:10.1037/a0014666
- Blanton, H., Jaccard, J., Strauts, E., Mitchell, G., & Tetlock, P. E. (2015). Toward a meaningful metric of implicit prejudice. *Journal of Applied Psychology*. doi: 10.1037/a0038379
- Blanton, H., & Mitchell, G. (2011). Reassessing the predictive validity of the IAT II: Reanalysis of Heider & Skowronski (2007). *North American Journal of Psychology*, 13(1), 99–106.
- Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology*, 4, 79–101.
- Bruner, J., & Goodman, C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, 42, 33–44.
- Buhrmester, M. D., Blanton, H., & Swann, W. B., Jr. (2011). Implicit self-esteem: Nature, measurement, and a new way forward. *Journal of Personality and Social Psychology*, 100(2), 365–385.
- Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review*, 16, 330–350.
- 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 and Social Psychology*, 89, 469–487.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, 12(2), 163–170. doi:10.1111/1467-9280.00328
- 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 and Social Psychology*, 81, 800–814.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.
- Dodaj, A. (2012). Social desirability and self-reports: Testing a content and response-style model of socially desirable responding. *Europe's Journal of Psychology*, 8, 651–666.

- Domoff, S. E., Hinman, N. G., Koball, A. M., Storfer-Isser, A., Carhart, V. L., Baik, K. D., & Carels, R. A. (2012). The effects of reality television on weight bias: An examination of *The Biggest Loser*. *Obesity*, 20, 993–998.
- Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, 23, 316–326.
- Egloff, B., & Schmukle, S. C. (2003). Does social desirability moderate the relationship between implicit and explicit anxiety measures?. *Personality and Individual Differences*, 35, 1697–1706.
- Ewoldsen, D. R., Rhodes, N., & Fazio, R. H. (2014). The MODE model and its implications for studying the media. *Media Psychology*. Advance online publication. doi:10.1080/15213269.2014.937440
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty, & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Erlbaum.
- Fazio, R. H. (1999). *Correlation between the IAT and the bona fide pipeline*. Unpublished data.
- 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 and Social Psychology*, 69, 1013–1027.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Galdi, S., Gawronski, B., Arcuri, L., & Friesen, M. (2012). Selective exposure in decided and undecided individuals: Differential relations to automatic associations and conscious beliefs. *Personality and Social Psychology Bulletin*, 38, 559–569.
- Gapinski, K. D., Schwartz, M. B., & Brownell, K. D. (2006). Can television change anti-fat attitudes and behavior? *Journal of Applied Biobehavioral Research*, 11, 1–28.
- Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology*, 33, 573–589.
- Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us?: Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, 2, 181–193.
- 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 and Social Psychology*, 108(4), 553–561. doi:10.1037/pspa0000016
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. doi:10.1037/0022-3514.74.6.1464
- 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 and Social Psychology*, 85, 197–216.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97, 17–41.

- Häfner, M., Jagsch, O., Kund, A., Mager, S., Pereira, P., & Zimmermann, A. (2008). "The female may feel male:" Defending against the adverse consequences of exposure to idealized media images. *Journal of Social and Clinical Psychology*, 27, 778–808.
- Hilgard, J., Bartholow, B. D., Dickter, C. L., & Blanton, H. (2015). Characterizing switching and congruency effects in the implicit association test as reactive and proactive cognitive control. *Social Cognitive and Affective Neuroscience*, 10(3), 381–388. doi:10.1093/scan/nsu060
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005a). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measure. *Personality and Social Psychology Bulletin*, 31, 1369–1385.
- Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005b). What moderates implicit-explicit consistency?. *European Review of Social Psychology*, 16, 335–390.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.
- Huang, L., & Chang, C. (2005). Body image and its predictors in Taiwan: Interpersonal pressure and media agreement as mediators. *Chinese Journal of Psychology*, 47, 287–308.
- Jaccard, J., & Blanton, H. (2007). A theory of implicit reasoned action: The role of implicit and explicit attitudes in the prediction of behavior. In I. Ajzen, D. Albarracín, and J. Hornik (Eds.), *Prediction and change of health behavior: Applying the reasoned action approach* (pp. 69–81). Mahwah, NJ: Erlbaum.
- Jaccard, J., & Wan, C. K. (2006). A paradigm for studying the accuracy of self-reports of risk behavior relevant to AIDS: Empirical perspectives on stability, recall bias, and transitory influences. *Journal of Applied Social Psychology*, 25, 1559–1816.
- Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76, 349–364.
- Kenny, D. A., & Judd, C. M. (1996). A general procedure for the estimation of interdependence. *Psychological Bulletin*, 119(1), 138–148. doi:10.1037/0033-2909.119.1.138
- Mauro, R. (1990). Understanding L.O.V.E. (left out variable error): A method for estimating the effects of omitted variables. *Psychological Bulletin*, 108, 314–329.
- Maylor, E. A., Rabbitt, P. M. A., James, G. H., & Kerr, S. A. (1992). Effects of alcohol, practice, and task complexity on reaction time distributions. *Quarterly Journal of Experimental Psychology A*, 44, 119–139.
- 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.
- Nebes, R. D., & Brady, C. B. (1992). Generalized cognitive slowing and severity of dementia in Alzheimer's disease: Implications for the interpretation of response-time data. *Journal of Clinical and Experimental Neuropsychology*, 14, 317–326.
- Nebes, R. D., & Madden, D. J. (1988). Different patterns of cognitive slowing produced by Alzheimer's disease and normal aging. *Psychology and Aging*, 3, 102–104.

- Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, 134, 565–584.
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, 19, 625–666.
- Olson, M. A., & Fazio, R. H. (2014). Reducing the influence of extrapersonal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology*, 86, 653–667.
- Oswald, F., Mitchell, G., Blanton, H., Jaccard, J. & Tetlock, P. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT research. *Journal of Personality and Social Psychology*, 105, 171–192.
- Oswald, F., Mitchell, G., Blanton, H., Jaccard, J. & Tetlock, P. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal importance. *Journal of Personality and Social Psychology*, 108(4), 562–571. doi:10.1037/pspa0000023
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, 94, 16–31.
- Payne, B., Cheng, C., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, 89, 277–293.
- Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). A process model of affect misattribution. *Personality and Social Psychology Bulletin*, 36, 1397–1408.
- Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology*, 82, 469–487.
- Petty, R. E., Fazio, R. H., & Briñol, P. (2009). *Attitudes: Insights from the new implicit measures*. New York, NY: Psychology Press.
- Petty, R. E., & Krosnick, J. A. (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.
- Prentice, D. A., & Miller, D. T. (1993). When small effects are impressive. *Psychological Bulletin*, 112, 160–164.
- Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit association test. *Group Processes and Intergroup Relations*, 10, 359–372.
- Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology*, 22, 13–39. doi:10.1521/jscp.22.1.13.22766
- Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal Of Social And Clinical Psychology*, 22(1), 13–39. doi:10.1521/jscp.22.1.13.22766
- Skorek, M., & Dunham, Y. (2012). Self-enhancement following exposure to idealized body portrayals in ethnically diverse men: A fantasy effect of advertising. *Sex Roles*, 66, 655–667.
- Strick, M., van Baaren, R. B., Holland, R. W., & van Knippenberg, A. (2009). Humor in advertisements enhances product liking by mere association. *Journal of Experimental Psychology: Applied*, 15, 35–45.

- Stroud, N. J. (2011). *Niche news: The politics of news choice*. New York, NY: Oxford University Press.
- Thistlethwaite, D. (1950). Attitude and structure as factors in the distortion of reasoning. *Journal of Abnormal and Social Psychology*, 45, 442–458.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554.
- Verhaeghen, P., Cerella, J., Semenec, S. C., Leo, M. E., Bopp, K. L., & Steitz, D. W. (2002). Cognitive efficiency modes in old age: Performance on sequential and coordinative verbal and visuo-spatial tasks. *Psychology and Aging*, 17, 558–570.
- Vernon, P. A., & Jensen, A. R. (1984). Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences*, 5, 411–423.
- Webb, E., Campbell, D., Schwartz, R., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago, IL: Rand McNally.
- Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25(4), 41–78.
- Wittenbrink, B., & Schwarz, N. (2007). *Implicit measures of attitudes*. New York, NY: The Guilford Press.

Copyright of Media Psychology is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.