2018 ■ Vol. 14 ■ no. 3

# A Primer on the benefits of differential treatment analysis when predicting discriminatory behavior

Rickard Carlsson [a], Jens Agerström [a], Donald Williams [b] & Gary N. Burns [c, ✉]

[a]Linnaeus University
[b]University of California, Davis
[c]Linnaeus University and Florida Institute of Technology

**Abstract** ■ A central question in social psychology is to what extent individual differences in attitudes, prejudices, and stereotypes can predict discriminatory behavior. This is often studied by simply regressing a measure of behavior toward a single group (e.g., behavior toward Black people only) onto the predictors (e.g., attitude measures). In the present paper, we remind researchers that an analysis focusing on predicting the differential treatment (e.g., behavior towards Black people vs. White people) has a higher conceptual validity and will result in more informative effect sizes. The paper is concluded with a list of suggestions for future research on the link between attitudes, prejudices, stereotypes and discrimination.

**Keywords** ■ Discrimination; attitudes; stereotypes; prejudice; methodology. **Tools** ■ R.

**Acting Editor** ■ Roland Pfister (Universität Würzburg)

**Reviewers** ■ One anonymous reviewer

✉ gburns@fit.edu

*RC*: 0000-0002-6456-5735; *JA*: 0000-0001-6134-0058; *DW*: 0000-0001-6735-8785; *GNB*: 0000-0001-7484-567X

10.20982/tqmp.14.3.p193

## Introduction

A central question in social psychology is to what extent attitudes, stereotypes and prejudice can predict discriminatory behavior. The literature has recently been synthesized in two systematical meta-analytical reviews (Talaska, Fiske, & Chaiken, 2008; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013). Two types of statistical analysis are roughly equally commonly used in the literature in order to examine this question. We will henceforth refer to them as the single group analysis and the differential treatment analysis.

Single group analysis means regressing a measure of behavior toward members of a single group (e.g., Black people) onto the study predictors (e.g., Amodio & Devine, 2006; Dovidio & Gaertner, 2000; Stepanikova, Triplett, & Simpson, 2011). Differential treatment analysis focus on within-individual differences in behavior towards different groups (e.g., Black people – White people), by regressing the difference scores onto the predictors (e.g., Dovidio, Kawakami, & Gaertner, 2002), or through multi-level modeling (e.g., Ziegert & Hanges, 2005).

In the present paper, we will argue that only the differential treatment analysis is a valid and potentially reliable method of predicting discriminatory behavior from measures of attitudes, stereotypes or prejudice. We present our argument based on the construct of discrimination, difficulties in interpreting effect sizes from single group analysis, and complications when testing for the presence of individual differences in discrimination. In order to illustrate the practical difference between the two approaches, we will have a closer look at a published study (Heider & Skowronski, 2007) where the conclusions between the two types of analysis differed to the extent that a re-analysis by Blanton and Mitchell (2011) lead to a formal correction by the original authors. Following this, we will apply the ideas of the present paper by designing and analyzing a fictive study where we compare the conclusions that can be drawn from the two types of analysis. The paper is concluded with a few methodological recommendations for future research.

### *The Single Group Analysis Lacks Conceptual Validity*

Discrimination is almost exclusively operationalized as differential treatment between groups, both in psychological research (see e.g., Ahmed, 2010; Milgram, Mann, & Harter, 1965; Word, Zanna, & Cooper, 1974) and in economics (see e.g., Bertrand & Mullainathan, 2004; Riach & Rich, 2002). This consensus among discrimination researchers is not the least surprising, since discrimination is inherently a relative concept. To discriminate is, per definition, to distinguish between something and something else. For example, in the case of ethnic discrimination, it is to make a difference between people based on their ethnicity. In simplistic terms, discriminatory behavior is the difference in behavior towards two groups ($Y_1 - Y_2$).

The conceptual nature of discrimination is not necessarily intentionally ignored by single group analysis, but this type of analysis fails to take the relative nature into account and thus changes the operationalization of discrimination. Regressing a measure of behavior towards a single focal group entails operationalizing discrimination in an absolute, rather than a relative, sense. Any operationalization of discrimination that does not capture the inherent relative nature of discrimination has poor conceptual validity as a discrimination measure. It is, at best, a measure of a correlated proxy ($Y_1$) to the real discriminatory behavior ($Y_1 - Y_2$).

### *Difficulties in Interpreting Effect Sizes from Single-Group Designs*

Beyond lacking conceptual validity, measuring $Y_1$ as a proxy of $Y_1 - Y_2$ introduces a number of difficulties in interpreting effect sizes. We are not the first to acknowledge this problem, yet researchers in the field have generally not elaborated much on this issue. This is illustrated by Talaska et al. (2008) in a meta-analysis where the issue of measuring behavior toward a single group is briefly mentioned in a footnote:

> ...the size of the correlations in the studies that measured behavior toward outgroup targets only, without reference to behavior toward ingroup targets, may be inflated by this phenomenon [ignoring the relation of the attitude measure to ingroup behavior]. This finding questions the meaning of studies that measure behavior toward outgroup members only. Perhaps some attitude measures simply predict who will be more or less aggressive or conformist, rather than who will behave in a specifically prejudiced manner. (p. 274)

In order to illustrate the difference in interpreting the effect sizes (e.g., the correlation) derived from a single-group design and a differential treatment design when predicting discrimination, we will now make an analogy with consumer products. Let us assume that we have a measure of attitudes toward Coca-Cola. Naturally, this attitude should predict one very important behavioral outcome: drinking Coca-Cola. Now, suppose we are specifically interested in whether people with positive attitudes toward Coca-Cola prefer drinking Coca-Cola to Pepsi-Cola when offered a choice. That is, if they discriminate between Coca-Cola and Pepsi-Cola. Clearly, this question of whether people prefer drinking Coca-Cola to Pepsi-Cola is not the same as whether they like drinking Coca-Cola. It is easy to imagine that positive attitudes toward Coca-Cola are positively related to drinking both Coca-Cola and Pepsi-Cola. This would be the case if the attitude measure also captures a general preference for soft drinks. Hence, the attitude measure that correlates most strongly with the tendency to drink Coca-Cola is not necessarily the attitude measure that correlates most strongly with the tendency to prefer drinking Coca-Cola instead of Pepsi-Cola.

As illustrated, a measure's ability to predict behavior towards a single object is not the same as its ability to predict differential behavior toward that object and another object. A measure's ability to predict behavior toward Black people is thus not always the measure that most strongly predicts differential behavior toward Black versus White people (i.e., discrimination). Indeed, similar to the imaginary Coca-Cola attitude measure that had a component related to soft drinks more generally, the attitude measure toward Black people might have a component that is related to behavior toward people in general, regardless of race. Suppose our behavioral outcome is how friendly a person is toward Black people. As noted by Talaska et al. (2008), an attitude measure that captures a latent construct of friendliness toward people in general, regardless of race, would spuriously increase the correlation with friendly behavior toward black people. However, it would also increase the correlation, in the same direction, with friendly behavior toward White people. In terms of predicting discrimination, this measure may end up just as good as, or even worse than, a measure that did not capture that general component.

Another aspect that has not received attention in the literature is that by correlating an attitude measure with behavior toward single-group only, we may get an underestimation of the measure's predictive validity. Imagine a study that captures behavior toward Black individuals ($Y_1$) and White individuals ($Y_2$). The measure may be poorly predictive of behavior toward Black individuals per se ($Y_1$), yet be a good predictor of differential behavior ($Y_1 - Y_2$). This will happen if part of the discrimination variance is due to variation in how White individuals are treated ($Y_2$).

In the single group case, then the discrimination would be simply missed due operationalizing discrimination from a single group design.

***Difficulties in Testing for Individual Differences in Discrimination***

A key assumption of the single-group operationalization is that some of the variance in the behavior toward the group in question (e.g., Black people) is due to individual differences in discrimination. If this assumption is true, then a measure of attitudes, prejudices or stereotypes would indeed predict this discrimination variance. Yet, how can we know how much, if any, of the variance in the behavior towards this single group is due to discrimination? Of course, without a comparison group, we cannot. Hence, predicting behavior toward a single group puts the measure to the task of predicting all variance in this behavior, without knowing if any of the variance is actually due to discrimination. Indeed, the only evidence that the variance in the outcome was, indeed, due to discrimination is the correlation with the attitude, prejudice or stereotype measure itself. Thus, a single-group operationalization results in circular logic: the evidence of the measure's ability to predict discrimination is based on its correlation with a behavior, whose only evidence of having measured discrimination is the correlation with the measure.

Of course, knowing what variance is due to discrimination and what is error variance is not easy. However, by adopting a differential treatment design, the researcher is in a much better position to discern between discrimination variance and error variance since this approach allows them to calculate a main discrimination effect and then assess whether this effect is significantly moderated by the predictor measure. That is, to show that people in general discriminate to a certain degree, but that some people discriminate less, and some people discriminate more. In this regard, the main effect can be thought of as a manipulation check as it shows that discrimination actually occurred in the experiment.

***Spurious Effects in Single Group Analysis: A Case Example***

In this part, we will take a closer look at a study by Heider and Skowronski (2007) that had been re-analyzed and disputed by Blanton and Mitchell (2011). We choose this study because it is rare that the results are available both in the form of absolute scores toward Black partners and White partners, and as difference scores. Hence, it is possible to directly compare single-group and differential treatment operationalizations within the same study.

Heider and Skowronski (2007) assessed White university students' implicit and explicit racial attitudes and to what extent such attitudes predict cooperation with Black and White partners in a Prisoner's Dilemma game. Although the nature of their study design allows for a differential treatment operationalization of discrimination, Heider and Skowronski (2007) conducted a single group analysis where they regressed the co-operation scores for the Black partner onto the attitude measures. They concluded that the attitude measures "significantly predicted cooperation scores" (p. 62).

In their re-analysis, Blanton and Mitchell (2011) show that the effect disappears if the difference score is used as the dependent variable. Indeed, the IAT correlations with the level of cooperation with Black and White partners are very similar and in the same direction ($r = -.21$ and $-.18$ respectively). Hence, the IAT does not predict discrimination, but a single-group operationalization misleads the authors to draw that conclusion. Rather, it appears as if people with higher implicit race bias scores on the IAT are less co-operative regardless if their partner is Black or White, and that this does not lead to any differential treatment.

**Comparing Single Group and Differential Treatment Analyses: A Fictive Study**

We will now proceed by illustrating with a concrete example how to design a study, and analyze the data, in order to take full advantage of a differential treatment operationalization. To this end, we will analyze data from fictitious datasets that we have generated for this purpose using R, and the R-packages `lme4` (Bates, Maechler, Bolker, & Walker, 2015) and `psych` (Revelle, 2018). The code can be found in the supplementary data on the OSF (https://doi.org/10.17605/OSF.IO/SN4QZ) and it allows for re-running the reported simulations with different sets of parameter values. We will return to the details of these simulations later.

For the following section, we first analyze the results in accordance with a single-group operationalization and then in accordance with a differential treatment operationalization.

***The Fictive Study***

In this example study, psychology students ($N = 200$, all Swedish) completed a measure of their attitudes toward Syrian refugees at the start of the semester. The attitude measure was embedded within a larger political survey to avoid raising suspicion. Two months into the semester, the students participated in an experiment where they collaborated with two partners to solve a series of tasks presented online. They were able to interact with the partners in a chat room. At the start of the session, the two partners introduced themselves in the text chat. One partner introduced himself/herself as a native Swedish student. The other partner introduced himself/herself as a student who had arrived as a refugee to Sweden from Syria three years ago. This first part constituted the manipulation of ethnicity (Swedish vs. Syrian) and was done through standardized text generated by a chat bot. Gender of the partner was determined randomly.

Immediately after this presentation, real partners took control of the chat and the collaboration task. The partners were trained confederates who were blind to the introduction in the chat before they took control. The confederates were further not aware that the research question concerned discrimination. The actual collaboration task was moderately difficult and took 15 minutes to complete. The participants collaborated with both partners simultaneously.

After the collaboration task was over, the participant was told that the computer would randomly select one of the three team members to be the leader who freely gets to decide how much compensation each team member receives. The participant was always selected to be the leader. He/she was informed that the two other team members were unaware of the size of the leader's budget (€25), and how much money the leader chose to transfer to him/herself and the other team member. Hence, this study uses a variant of the well-known dictator game (Kahneman, Knetsch, & Thaler, 1986).

**Results - Single Group Analysis.** The results show that the participants allocated a mean of €3.95 (SD = 1.10) to the Syrian partner. Furthermore, the Syrian refugee attitude measure was found to be positively correlated with money distributed to the Syrian partner, $r(198) = .36$, 95% CI[.23, .47], $p < .001$. Hence, those with more negative attitudes toward Syrian refugees gave less money to the Syrian partner, suggesting that the attitudes predict discrimination of Syrians.

**Results - Differential Treatment Analysis.** The results show that the participants allocated a mean of €3.95 (SD = 1.10) to the Syrian partner, and M = 4.9 (SD = 1.02) to the Swedish partner and kept a mean of €16.14 (SD = 1.87) for themselves. Clearly, participants acted selfishly as they kept most money for themselves. Further, they allocated more money to the Swedish partner than the Syrian partner: mean difference = 0.95, SD = 0.99, 95% CI 0.81, 1.08], $t(199) = 13.585, p < .001$. Thus, the participants discriminated against Syrians.

The attitude scale was virtually uncorrelated with the difference score, $r(198) = .00$, 95% CI [-0.14, 0.14], $p = .992$. The confidence interval allows us to reject the presence of anything but very weak effects. Hence, the results suggest that the attitude scale did not predict discriminatory behavior at any meaningful level. However, it was positively and moderately correlated with both allocation toward Syrian, $r(198) = .36$, 95% CI[.23, .47], $p < .001$, and Swedish partners, $r(198) = .39$, 95% CI [.26, .50], $p < .001$. Thus, it seems as if the attitude scale simply predicted generosity. Indeed, the more positive attitudes toward Syrian refugees, the less the participants kept for themselves, $r = -0.42$, 95% CI [-.53, -.30].

In sum, there is no indication that the Syrian refugee attitude scale predicted discrimination, rather it seems to have predicted selfish behavior in general. Another possibility is that both the attitude scale and the distribution task share variance that is due to self-presentation. That is, the same people who tried to appear non-prejudiced also tried to appear unselfish in the experiment by sharing more money with their partners. However, in doing so they did in fact end up discriminating between the Syrian and Swedish partner.

### *A comment on the Simulations*

The beauty of simulating data is that we can get any data that we want. We chose this example because it is illustrative of the difference between the approaches. We reached this value by simulating data with a true positive discrimination effect of the partner's ethnicity, a true positive main effect of attitudes on sharing money and zero interaction effect. In other words, the simple effects for Syrian and Swedish partners were set to be identical. The overestimation can be even bigger if simple effects are larger for the Swedish partners and thus the correlation with the difference score would be in the opposite direction. If the simple effect for the Swedish partner is negative, then the prediction of discriminatory behavior will instead be underestimated. Indeed, it is possible that the attitude measure is a strong predictor of discriminatory behavior despite showing a correlation of zero in the single group analysis. The two approaches will only be identical when the simple effect for the Swedish partners is zero. The code of the simulation is documented, to allow for exploring different scenarios by setting relevant parameters.

## Discussion

In the present paper, we have argued that researchers aiming to predict discrimination by means of attitudes, prejudices or stereotypes, should direct their attention away from the quite common approach of predicting behavior towards single groups (e.g., behavior toward Black people), and instead focus on predicting individual differences in differential treatment. This approach is a better fit conceptually, since discrimination is inherently relative and cannot be operationalized in absolute terms. It further reduces the risk of inflated (i.e., due to shared irrelevant variance) or attenuated (i.e., due the effect being driven by behavior towards the comparison group) correlations that can arise from a one-sided focus of predicting behavior towards a single group.

Although we have chosen to specifically focus on the role of the comparison group in discrimination outcomes, we think that it is important to put this into perspective by also discussing other issues when predicting discrimination. For example, whether the outcome reflects discriminatory behavior or merely self-reported behavioral intentions (see Baumeister, Vohs, & Funder, 2007), or whether the stimuli need to be treated as random effects (Judd, Westfall, & Kenny, 2012), are certainly pressing concerns for discrimination researchers, as is the call for an increase in statistical power and transparent reporting (Schimmack, 2012). Furthermore, if participants are not naïve to the research questions (as in field experiments, e.g., Bertrand & Mullainathan, 2004), perhaps the outcomes will be distorted due to motivation to respond without prejudice. We are not suggesting that researchers should lose track of these important issues either. Rather, we think that attention to these aspects will converge in good designs. Treating stimuli as random practically necessitates several observations per participants, which also greatly improves the estimation of individual differences in discrimination. Similarly, relying on more ecologically valid behavioral outcomes may improve the diagnostic value of a certain behavior and thus leave little room open for variance that is not due to individual difference in discrimination. For example, the choice to sit next to a person may be a trivial choice and participants may select this at random. To identify individual differences in discrimination of this kind, many trials may be needed. In contrast, an experiment where the participant makes high-stake decisions to invite people to job interviews, the choice to invite, say, two white candidates and not to invite an equally qualified black candidate, can be highly diagnostic, since it was indeed possible to invite all three of them. And if the White candidates are rarely randomly left out from the invitations, a non-invite of a Black candidate is strongly diagnostic of discrimination.

The increased attention to estimating the level of discrimination on the individual level naturally coincides with more high-powered designs. Indeed, increasing statistical power is not only about larger samples. Outcomes with less random noise (differential treatment instead of single-group and more diagnostic outcomes) and multiple observations per participant will greatly increase statistical power for the same number of participants.

This paper focused on a statistical issue in predicting discriminatory behavior. As indicated in our analogy with consumer products (Coca-Cola vs. Pepsi-Cola) this issue is not limited to discrimination due to for example, ethnicity and race, but to all constructs that are inherently differential. Thus, the points made in this article is relevant for studies looking at other kinds of discrimination (e.g., visual or auditory), or preferential behavior (e.g., choosing between two schools). Reviewing this is beyond the scope of the present article, but we simply note that to the extent that researchers in any area are interested in regressing measures of relative concepts onto a set of predictors, they will benefit from using a differential "treatment" analysis.

So far, our focus has been on how to design and analyze future studies. An equally important question is what to do with the abundance of "discrimination" studies relying on a single-group approach. Our proposal is simple: ignore them as outcomes of discrimination. Today, meta-analysis as a tool is becoming increasingly popular. With it comes the realization that mashing together studies of low and high quality has its problems. That one spoiled apple will spoil the barrel may be an exaggeration, but when the literature is clotted with studies using invalid operationalizations, the problem becomes dire. Here meta-analysts have a huge responsibility. Meta-analysts should not draw conclusions based on aggregations of those studies that have valid and reliable operationalizations of discrimination and those studies that do not.

We do realize that meta-analysts have to make the best of the available literature. Setting the bar too high will likely exclude almost every study and make the meta-analysis meaningless. Yet, doing so will send a strong signal that more research of higher quality is needed. Selecting studies based on methodological quality for inclusion in meta-analyses will likely help to improve the predictive validity of attitudes, prejudices, and stereotypes.

## References


Ahmed, A. (2010). Muslim discrimination: Evidence from two lost letter experiments. *Journal of Applied Social Psychology*, *40*, 888–898. doi:10.1111/j.1559-1816.2010.00602.x

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*(4), 652–661. doi:10.1037/0022-3514.91.4.652

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. doi:10.1111/j.1745-6916.2007.00051.x

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*, 991–1013. doi:10.1257/0002828042002561

Blanton, H., & Mitchell, G. (2011). Reassessing the predictive validity of the IAT II: Reanalysis of Heider & Skowronski (2007). *North American Journal of Psychology*, *12*, 99–106.

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, *11*(4), 315–319. doi:10.1111/1467-9280.00262

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*(1), 62–68. doi:10.1037/0022-3514.82.1.62

Heider, J. D., & Skowronski, J. J. (2007). Improving the predictive validity of the implicit association test. *North American Journal of Psychology*, *9*, 53–76.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54–69. doi:10.1037/a0028347

Kahneman, D., Knetsch, J., & Thaler, R. (1986). Fairness and the assumptions of economics. *The Journal of Business*, *59*, 285–300.

Milgram, S., Mann, L., & Harter, S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly*, *29*, 437–438.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of iat criterion studies. *Journal of Personality and Social Psychology*, *105*, 171–192. doi:10.1037/a0032734

Revelle, W. (2018). *Psych: Procedures for personality and psychological research*. Evanston, Illinois, USA: Northwestern University.

Riach, P., & Rich, J. (2002). Field experiments of discrimination in the market place. *The Economic Journal*, *112*, 480–518. doi:10.1111/1468-0297.00080

Schimmack, U. (2012). The ironic effect of significant results on the credibility of multiple-study articles. *Psychological Methods*, *17*(4), 551–566. doi:10.1037/a0029487

Stepanikova, I., Triplett, J., & Simpson, B. (2011). Implicit racial bias and prosocial behavior. *Social Science Research*, *40*, 1186–1195. doi:10.1016/j.ssresearch.2011.02.004

Talaska, C. A., Fiske, S. T., & Chaiken, S. (2008). Legitimating racial discrimination: Emotions, not beliefs, best predict discrimination in a meta-analysis. *Social Justice Research*, *21*, 263–296. doi:10.1007

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109–120. doi:10.1016/0022-1031(74)90059-6

Ziegert, J. C., & Hanges, P. J. (2005). Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. *Journal of Applied Psychology*, *90*, 554–562. doi:10.1037/0021-9010.90.3.553

### Open practices

The *Open Material* badge was earned because supplementary material(s) are available on the journal's web site.

### Citation

Carlsson, R., Agerström, J., Williams, D., & Burns, G. N. (2018). A primer on the benefits of differential treatment analysis when predicting discriminatory behavior. *The Quantitative Methods for Psychology*, *14*(3), 193–205. doi:10.20982/tqmp.14.3.p193



Received: 06/10/2017 ~ Accepted: 04/08/2018