Longitudinal Intergroup Contact Effects on Prejudice Using Self- and Observer-Reports

Kristof Dhont, Alain Van Hiel, Marleen De Bolle, & Arne Roets

Ghent University, Department of Developmental, Personality and Social Psychology

2012

British Journal of Social Psychology

Running head: Longitudinal contact effects with self- and observer-reports

Word count: 6147 (excluding abstract, references, figures, and tables)

Correspondence address:

Kristof Dhont, Department of Developmental, Personality and Social Psychology, Henri Dunantlaan 2, B-9000, Ghent, Belgium, e-mail address: Kristof.Dhont@UGent.be

Acknowledgments:

This research was supported by a PhD research grant from the Ghent University Research Council (BOF) awarded to Kristof Dhont (#01D23607) under supervision of Alain Van Hiel. Arne Roets is a post-doctoral researcher supported by the Research Foundation – Flanders (FWO). Special thanks are due to Miet Craeynest at Lessius University College for her aid in data-collection and to Oliver Christ, Miles Hewstone, and the anonymous reviewers for their thoughtful comments on earlier versions of this article.

Abstract

Longitudinal effects of intergroup contact on prejudice were investigated in a sample of 65 young adults (Sample 1) and a sample of their close friends (Sample 2, N = 172), adopting a full cross-lagged panel design. We first validated the self-report measure of intergroup contact from sample 1 with observer ratings from sample 2 by demonstrating that self-reports and observer ratings of contact were highly correlated. Moreover, we obtained significant cross-lagged effects of intergroup contact on prejudice with both contact measures, thereby providing a second validation for the use of self-reports of intergroup contact. Finally, by the use of latent change modeling we demonstrated that, although no overall significant change in contact and prejudice over time was found, there was meaningful variation in absolute change in the individual levels of intergroup contact and prejudice. In particular, some individuals showed increases while others showed decreases in contact or prejudice across time. Moreover, higher levels of intergroup contact at Time 1 were followed by larger subsequent decreases in prejudice between Time 1 and Time 2, and changes in contact were significantly and negatively related to changes in prejudice. Methodological implications of the findings are discussed.

Key words: prejudice, intergroup contact, cross-lagged design, observer ratings; racism; self-reports

Longitudinal Intergroup Contact Effects on Prejudice Using Self- and Observer-Reports

Over the past several decades, a vast body of research has provided convincing empirical support for the theory that positive intergroup contact is likely to improve intergroup attitudes and reduce prejudice (Allport, 1954; Brown & Hewstone, 2005; Pettigrew, 1998). Analyzing this body of research in a meta-analytic study, Pettigrew and Tropp (2006) confirmed that "intergroup contact typically reduces intergroup prejudice" (p. 766), revealing a moderate mean effect (r = -.21, p < .0001).

Despite the accumulating amount of evidence that supports the contact theory, some important limitations in the research on intergroup contact make the interpretation of many findings difficult. One of these limitations pertains to the extensive use of self-report measures of intergroup contact. A second limitation is the lack of longitudinal studies that have fully explored the cross-lagged relationship between intergroup contact and prejudice. The present study aimed to address these two important methodological issues.

Self- and Observer-Reports of Intergroup Contact

Although past research has employed several different methods and operationalizations to investigate intergroup contact (for a discussion, see Christ & Wagner, in press; Hewstone, Judd, & Sharp, 2011), 81% of the studies included in Pettigrew and Tropp's (2006) meta-analysis relied on self-report measures. Self-reports are inexpensive, relatively quick and simple to administer and interpret, and aid in collecting data efficiently from large samples (Kline, 1993; Paulhus & Vazire, 2007). Moreover, the respondents of self-reports are likely to be motivated to participate and may provide a large amount of information about themselves that they may not usually share with others (Paulhus & Vazire, 2007).

Despite their attractive features, self-reports have been criticized because they are prone to various response biases, such as the tendency to respond in a socially desirable way, to agree with all of the statements regardless of their content, or to maintain consistency in the responses to questions with related content (e.g., John & Robins, 1993, 1994; Kolar, Funder, & Colvin, 1996; Moskowitz, 1986; Paulhus, 1984; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Podsakoff & Organ, 1986).

If self-report measures of intergroup contact have been subject to social desirability biases in previous research, then the registered mean amount and quality of intergroup contact are likely to have been overestimated. Furthermore, participants may also respond in a way that is biased by their prejudice levels, which is even more problematic given that contact research typically investigates the relationship between contact and prejudice. In particular, it is possible that prejudiced people indicate low levels of positive contact because they are biased in how they remember the amount and quality of the intergroup encounters that they have had. Moreover, these people may want to maintain consistency between their experiences with and attitudes toward outgroup members, or they may report frequent negative contact to justify prejudiced attitudes. For the same reasons, non-prejudiced people can be expected to report frequent positive contact. As a result, the strength of the contact-prejudice relationship may be artificially inflated.

Hewstone et al. (2011) recently emphasized the need to validate self-reports of intergroup contact. Moreover, by comparing participants' self-reports with observer-reports of intergroup contact, these authors applied a validation technique that has frequently and successfully been applied in personality and organizational research (e.g., Costa & McCrae, 1988; John & Robins, 1993; Judge & Bono, 2000; Piedmont, McCrae, Riemann, & Angleitner, 2000; Podsakoff et al., 2003; Vazire, 2006; Vazire & Mehl,

2008). In particular, Hewstone et al. (2011) demonstrated a significant agreement between self- and observer-reports of intergroup contact. Moreover, the validity of the relationship between self- and observer-reports was substantiated by demonstrating that this relationship remained significant after controlling for self- and observer-reports of targets' levels of extraversion (a well-observable trait) and outgroup attitudes (of which contact might be inferred). They also showed that the self-reports of contact were only meaningfully related to the observer-reports of contact for the same outgroup but were not related to the observer-reports of contact for different outgroups.

It should, however, be noted that observer-reports are not free of biases. One potential bias pertains to a type of socially desirable responding in which observers present themselves more positively (e.g., having more intergroup contact) than the target that they judge. Another possible bias occurs when observers assume a certain degree of similarity between themselves and the target they judge. This assumed similarity may result in a projection effect, whereby observers attribute characteristics of their own to the person they judge. Although the results of Hewstone et al. (2011) did not suggest a strong influence of social desirability, these authors indeed reported a significant correlation between the reports of participants' own contact levels and the judgments they made about others' contact levels. This latter result may indicate the potential influence of a projection effect. However, as the authors argued, this similarity may also reflect, to a large extent, true similarity among friends, given that people are likely to be friends with people who are similar to themselves (e.g., Byrne, 1997; Deutsch, Sullivan, Sage, & Basile, 1991; Kiesner, Maass, Cadini, & Vallese, 2003). In sum, Hewstone et al. (2011) provided the first empirical evidence for the validity of self-reports of intergroup contact, but they also emphasized that "there is ample room for future research using observer-reports to validate self-reports of contact" (p. 9).

The Causality Issue

The second methodological issue addressed in the present research is whether contact leads to a reduction of prejudice. Despite the explicit causal character of the contact hypothesis, it is noteworthy that our knowledge about intergroup contact relies heavily on findings that were obtained using cross-sectional data (Christ & Wagner, in press; Pettigrew, 2008; Pettigrew & Tropp, 2006), which do not allow for causal inferences about the contact-prejudice relationship. However, several studies have investigated the longitudinal effects of intergroup contact on prejudice, which sheds some light on the causality issue.

The most extensive longitudinal study to date followed a cohort sample of more than 2,000 American students over a period of five years (Levin, van Laar, & Sidanius, 2003). The results indicated that the students who had a greater number of cross-group friends during college were more positively inclined toward outgroup members at the end of their college years. However, equally strong effects were observed for the students' prior levels of outgroup attitudes on the amount of cross-group friendships in college. Other longitudinal studies have also obtained significant effects in both directions (e.g., Binder et al., 2009; Eller & Abrams, 2003, 2004; for an exception see, Brown, Eller, Leeds, & Stace, 2007), whereas Stephan and Rosenfield (1978) demonstrated a significant relationship between the changes in intergroup contact over time and the changes in outgroup attitudes over time. Overall, the available evidence indicates that contact with outgroup members leads to lower prejudice levels, but also that prejudiced people are more likely to avoid intergroup contact.

When considering contact as an intervention strategy to reduce prejudice, the most important path is the one from contact to prejudice, which should be significant when the reverse causal path has been controlled for. However, the majority of previous

longitudinal studies used regression analyses (e.g., Binder et al., 2009; Brown, et al., 2007; Eller & Abrams, 2003, 2004; Levin et al., 2003; see also Christ & Wagner, in press) and tested the effects of contact at Time 1 on prejudice at Time 2 while controlling for prejudice at Time 1 but not for contact at Time 2 (for an exception, see Stephan & Rosenfield, 1978). The reverse causal order in these previous studies has been tested in a similar but separate analysis. One potential drawback of this approach is that the effects of contact at Time 1 on prejudice at Time 2 may have emerged solely because of the association of both variables with contact at Time 2, i.e., due to the stability of contact over time and the cross-sectional association between contact and prejudice at Time 2. Analogously, the effects of prejudice at Time 1 on contact at Time 2 may have emerged because both variables were associated with prejudice at Time 2.

A *full* cross-lagged panel approach allows for the control of these potential confounds. In particular, an effect of contact on prejudice can be demonstrated if contact at Time 1 affects prejudice at Time 2 when controlling for the stability of both of these variables over time, i.e., the autoregressive paths, and the cross-sectional associations between the variables (i.e., including prejudice at Time 1 and contact at Time 2 in the analyses). Hence, such a design allows for the direct comparison of contact effects on prejudice and prejudice effects on contact in the same analysis.

However, one limitation of cross-lagged models is that they do not account for potential individual differences in absolute change across time, but assume that the autoregressive coefficients are the same for all individuals (Christ & Wagner, in press; Hertzog & Nesselroade, 2003). In particular, comparing the scores on a prejudice scale across different points in time may reveal a significant overall decrease or increase in prejudice, but it does not provide information about whether individuals significantly vary in the degree to which they follow this overall trend. Moreover, even when no

significant overall increase or decrease in prejudice is noted, individual prejudice scores may still be subject to an increase or decrease over time.

Evidently, when theorizing about intergroup contact, scholars are likely to be interested in whether inter-individual variations in intergroup contact are associated with inter-individual variations in prejudice *reduction* (i.e., changes in prejudice), which is exactly what the use of latent change modeling (Hertzog & Nesselroade, 2003; McArdle & Nesselroade, 1994) allows to investigate. In particular, instead of calculating difference scores to represent change in these variables, a Latent Change Model (LCM) starts from a two-wave latent factor model and defines for each variable the latent initial level (i.e., latent scores at Time 1) and latent change scores (i.e., reliable scores of change in contact and change in prejudice over time). Because these latent initial level and change scores are represented as factors, the variances of these factors can be estimated as parameters and the latent (change) factors can be modeled as causes or as consequences in a Structural Equation Model (SEM).

Hence, the use of LCM allows to answer the highly relevant question whether more intergroup contact at an initial time point indeed leads to a larger subsequent decrease in prejudice over time. Moreover, with LCM, we can also address whether, besides the initial scores on intergroup contact, also the inter-individual variations in the increase or decrease in intergroup contact can be related to the inter-individual variations in decrease or increase in prejudice over time. Or, to put it otherwise, whether changes in intergroup contact are significantly associated with changes in prejudice.

The Present Study

The goal of the present research was to contribute methodologically to the contact literature in two important ways. First, we aimed to proceed the endeavor to validate self-reports of contact by investigating the agreement between self-reports of intergroup

contact provided by a first sample of participants and observer-reports of intergroup contact provided by a second sample of participants. Second, we aimed to demonstrate longitudinal effects of contact on prejudice with both self- and observer-reports using path-analysis to test a full cross-lagged model, controlling for stability effects and cross-sectional associations.

Using the self-report data from sample 2, we also employed more sophisticated statistical procedures by testing the longitudinal contact effects with Structural Equation Modeling (SEM) using latent variables. Finally, because we wanted to explicitly investigate whether the initial levels of intergroup contact predict *change* in prejudice over time and whether *change* in intergroup contact is significantly related to *change* in prejudice, we also tested an LCM.

Method

Overview

We conducted a longitudinal study in two samples (Samples 1 and 2) of young adults living in the Dutch-speaking part of Belgium. We focused on contact with and prejudice toward immigrants with non-European roots, especially toward people who were from countries with a Muslim majority.

The participants in Sample 1 were invited to the laboratory, where they twice (referred to as Time 1 and Time 2) completed measures of intergroup contact and prejudice on a computer with an interval of approximately two months. Additionally, the participants were requested to distribute up to three questionnaires at Time 1 and Time 2 to three of their closest friends. In this way, the Sample 1 participants recruited the Sample 2 participants. The questionnaires for Sample 2 were enclosed in an envelope with a letter explaining the survey procedure and the participants' rights. Sample 1 participants were instructed that they only needed to ask their friends to participate and to

refer to the accompanying letter and questionnaire for further information about the study.

Sample 2 participants were first asked to rate their friend's levels of intergroup contact, with the explicit instruction not to consult this friend. Next, they completed self-report measures of intergroup contact and prejudice. At Time 1 and Time 2, the questionnaires of Sample 2 participants were returned in closed envelopes within two weeks after Sample 1 respondents completed their questionnaires.

Sample 1

Participants

A total of 65 undergraduate students (89% women, M_{age} = 18.78, SD_{age} = 1.28) participated in the present study in return for course credit. None of the respondents belonged to the target outgroup (all were Belgian and none were Muslim; 63% Christians, 37% atheists, non-religious people, or people who had another religion). A total of 59 participants (91%) also participated at Time 2.

Measures

Intergroup contact Intergroup contact was administered with a self-report measure as well as with observer ratings derived from participants of Sample 2. The self-report measure consisted of seven items (based on previously used items, e.g., Dhont & Van Hiel, 2009) rated on 7-point Likert scales and focused on the number of cross-group friendship experiences and on the quantity of positive intergroup contact. Sample items included "How many immigrant friends do you have?" (1 = none; 7 = many) and "How often do you have contact with immigrants within your circle of friends?" (1 = never; 7 = very often).

The observer ratings of intergroup contact of Sample 1 participants (i.e. the targets) were provided by their friends (Sample 2; i.e., the observers) who completed

seven items analogous to the self-report items, such as "How many immigrant friends does your friend have?" (1 = none; 7 = many). For each participant in Sample 1, an average of 2.65 (Time 1) and 2.05 (Time 2) observer scores were obtained.

The self-report measure and the observer measure of intergroup contact proved to be internally consistent at Time 1 and Time 2. Cronbach's αs ranged from .91 to .95 (average α = .94). To investigate the extent of agreement between the observer ratings at Time 1 and Time 2, we calculated an intraclass correlation (ICC) coefficient, which estimated the homogeneity among observers evaluating the same target (e.g., Judge & Bono, 2000). The ICCs indicated the presence of sufficient agreement between the observers in how the targets were judged; ICCs = .46 and .48 for Time 1 and Time 2, respectively. An ICC above .20 has been used to justify aggregation across observers (Judge & Bono, 2000; Ostroff & Schmitt, 1993). Hence, for further analyses, we averaged the observer scores into a single index.

Prejudice and outrgroup attitudes To assess prejudice and attitudes toward immigrants, participants completed measures of subtle racism, outgroup attitudes, and endorsement of negative stereotypes. The subtle racism scale (Pettigrew & Meertens, 1995; see also Dhont, Roets, & Van Hiel, 2011) was assessed with eight items using 7-point Likert scales (1 = strongly disagree; 7 = strongly agree). A sample item was "I admire the immigrant community who live here under difficult circumstances" (reverse scored).

Outgroup attitudes were measured using a modified version of the 'General Evaluation Scale' (Wright, Aron, McLaughlin, & Ropp, 1997), which asked the participants to describe their feelings about immigrants in general using four 7-point differential scales: cold-warm, positive-negative, hostile-friendly, and contempt-respect. These items were coded with higher scores indicating a more positive attitude.

Lastly, to assess the endorsement of negative stereotypes, participants indicated on 7-point Likert scales (1 = certainly not; 7 = certainly) "whether the following traits represent good descriptions for immigrants in our country": lazy, untrustworthy, arrogant, noisy, and aggressive.

Given that the scores for subtle racism, outgroup attitudes, and stereotyping were highly correlated (see Table 1), we extracted one factor from all items for both measurement occasions, which accounted for 48.54% and 51.67% of the variance in the scores of Time 1 and Time 2, respectively. Mean item loadings on this general factor were |.70| for Time 1 and |.71| for Time 2.

Sample 2: Participants and Measures

A total of 172 participants (62% women, M_{age} = 19.39, SD_{age} = 1.83) were recruited by Sample 1 and completed the questionnaire at Time 1. All respondents belonged to the majority group (95% were Belgian, 5% were Dutch). None of the participants were Muslim (53% Christians and 47% atheists, non-religious people, or people who had another religion). A total of 123 participants (72%) completed the questionnaire at Time 2.

In addition to providing observer ratings for intergroup contact of their Sample 1 friends, respondents in Sample 2 completed the same self-report measures of intergroup contact and subtle racism as the Sample 1 participants.

Results

Preliminary Analyses

Comparing the scores of the Sample 1 respondents who completed the measures at both time points with the scores of those respondents who dropped out before Time 2 revealed no significant differences for any of the variables (all absolute *t*-values < 1.4). Moreover, comparison of the means and covariances of all variables using Little's (1988)

MCAR test revealed that the missing data were missing completely at random, χ^2 (17) = 12.52, p = .77. Therefore, these missing values were estimated using maximum likelihood estimation (Schafer, 1997) with the expectation maximization algorithm. Similar analyses of the scores of Sample 2 revealed no significant differences for any of the variables (all absolute t-values < 1.5). Little's MCAR test indicated that data were missing completely at random, χ^2 (17) = 7.23, p = .98, and therefore missing values were estimated. Tables 1 and 2 present the descriptive statistics and the correlations between the measures.

As shown in Table 1, the targets' self-reports of intergroup contact were strongly correlated with the observer-reports (rs = .71 and .73, for Time 1 and Time 2, respectively), which provides a first indication of the validity of the self-reports.

To obtain an indication of socially desirable responding, we tested whether the observers' ratings of their own levels of intergroup contact were significantly higher than their ratings of the target's levels of intergroup contact. For this purpose, for each target, we averaged the three observers' ratings of their own intergroup contact and compared this score with the averaged score of the observer-reports for the target. These tests revealed no significant differences, F(1,64) = .84, p = .36 and F(1,64) = .33, p = .57, for Time 1 and Time 2, respectively. Hence, the observers did not ascribe significantly higher levels of intergroup contact to themselves than to the target.

Finally, we investigated the (assumed) similarity between the observers' self-ratings of contact and their ratings of the targets' contact levels and found significant correlations between these measures at both time points, rs = .51 and .62, ps < .001 for Time 1 and 2, respectively (see also, Hewstone et al., 2011).

Cross-lagged Analyses with Sample 1 Data

Using Lisrel (Version 8.72), path-analysis with observed variables (instead of latent variables, due to the small sample size) was conducted to test the cross-lagged

relationships between contact and prejudice. In particular, we simultaneously analyzed the longitudinal effects of contact and prejudice at Time 1 on prejudice and contact at Time 2. A first model included the self-report measure of contact (Model 1), whereas a second model included the observer ratings (Model 2). Because initial data screening using Prelis 2.72 indicated that the data showed significant departures from the multivariate normal distribution (i.e., significant skewness) and that several variables departed from the univariate normal distribution, we used a Robust Maximum Likelihood estimation. Because all paths were estimated, these models were saturated (yielding perfect model fit).

Figure 1 depicts the results (i.e., standardized estimates) of both models. Model 1 (Figure 1, panel A) revealed a significant longitudinal effect of contact on prejudice, β = -.19, $p < .01$, whereas no significant longitudinal effect of prejudice on contact was found, β = -.10, *ns*. Similar results were obtained with the observer ratings in Model 2 (Figure 1, panel B). In particular, we obtained a significant longitudinal effect of contact on prejudice, β = -.22, $p < .001$, but no significant longitudinal effect of prejudice on contact emerged, β = -.03, *ns*. In other words, these results from the observer-reports cross-validated the findings obtained using the self-reports.

Cross-lagged Analyses with the Self-Reports of Sample 2

To investigate the cross-lagged relationships between contact and prejudice using the self-report data of Sample 2, we used structural equation modeling (SEM) with latent variables. To smooth measurement error and to maintain an adequate ratio of cases to parameters, we averaged subsets of items into balanced indicator parcels (Little, Cunningham, Shahar, & Widaman, 2002), which were held equal over time. Because the parcels exhibited significant departures from the multivariate normal distribution and several parcels exhibited significant departures from the univariate normal distribution,

we used a Robust Maximum Likelihood estimation. The Satorra-Bentler Scaled chi-square test statistic (χ^2), the comparative fit index (CFI), and the root-mean-square error of approximation (RMSEA) were used to evaluate the goodness-of-fit of the tested models. A satisfactory fit is indicated by a χ^2 lower than double the degrees of freedom, a CFI value greater than .95, and an RMSEA value of less than .06 (Hu & Bentler, 1999).

Measurement Invariance

Before testing the latent longitudinal models of intergroup contact and prejudice, it was necessary to investigate whether the measurement properties of the contact and prejudice measures were sufficiently equal over time (Byrne, Shavelon, & Muthén, 1989; Little, Preacher, Selig, & Card, 2007; Meredith, 1993).[1] Therefore, we needed to establish longitudinal measurement invariance (MI) by comparing a model (Model 1) in which the number of factors and accompanying loadings were specified to be equivalent across time but with freely estimated parameters (i.e., configural invariance) with a second model (Model 2) in which factor loadings of corresponding indicators across time were constrained to be invariant, imposing metric MI (cf. Allemand, Zimprich, & Martin, 2008; Brown, 2006). In a third model (Model 3), an additional constraint of equal intercepts of the manifest indicators across time was tested, implying scalar MI.[2]

As shown in Table 3, Model 1 had an acceptable fit, demonstrating configural invariance across time. Furthermore, the constraints imposed in Model 2 did not result in a significantly worse fit compared with Model 1, confirming metric MI. Finally, Model 3 also achieved a good fit, which was not significantly different compared with the fit of Model 2, while being more parsimonious. Hence, scalar invariance was also supported, allowing a meaningful comparison of the means, covariances, and variances across time.

Having established satisfactory measurement invariance, we tested a full cross-lagged model and an LCM to analyze the structural relationships between the latent

variables. In particular, the full cross-lagged model allowed us to investigate the effects of contact and prejudice at Time 1 on contact and prejudice at Time 2. Additionally, the LCM tested whether the initial levels of contact and prejudice predict changes in contact and prejudice from Time 1 to Time 2 and whether changes in contact are significantly related to changes in prejudice.

In these longitudinal models, the loadings of parallel indicators were constrained to be equal across time and the residual errors of parallel indicators were allowed to correlate in all analyses, reflecting stability in systematic error over time. The first factor loading of each latent variable was set to unity in order to scale the factors.

Full cross-lagged model

The full cross-lagged model included all paths from contact and prejudice at Time 1 to contact and prejudice at Time 2, (i.e., the autoregressive and cross-lagged paths) as well as the within-Time associations. The results (i.e., standardized estimates) of this model test are presented in Figure 2, panel A. The model had a very good fit; $\chi^2(27) = 39.96$, $p = .052$; RMSEA = .053; CFI = .99. In line with the findings in Sample 1, the results revealed that contact at Time 1 significantly and negatively predicted prejudice at Time 2, $\beta = -.19$, $p < .001$, whereas no significant longitudinal effect of prejudice on contact were found, $\beta = .05$, *ns*.

Latent Change Model

Based on a two-wave factor model, an LCM restructures the latent factors to latent level and change factors. In the current LCM, the latent level factors represent the initial levels of contact and prejudice, as defined by the latent scores of contact and prejudice at Time 1. The latent change factors represent the changes in contact and prejudice from Time 1 to Time 2, as defined by the difference between the latent scores at

the two times of measurement (Hertzog & Nessleroade, 2003; McArdle & Nesselroade, 1994).

To allow straightforward interpretations of the change scores (i.e., to know whether 'change' reflects increase or decrease), separate univariate LCMs for contact and prejudice needed to be estimated first. These models included the latent level and change factors and the autoregressive paths between these two latent factors. The means, variances, 95% confidence intervals, and score ranges of both unvariate LCMs are presented in Table 4. Although there was no overall mean change in contact and prejudice, the significant latent variances indicate significant inter-individual variability in the initial levels of contact and prejudice as well as in the individual estimates of true change in these variables. More specifically, as shown in Table 4, both for contact and prejudice, the individual change scores ranged from negative to positive values indicating that individual change can mean an increase as well as a decrease in contact or prejudice

Next, we tested the multivariate LCM comprising the univariate LCMs of contact and prejudice, and the paths between the latent constructs.[3] As can be seen in Figure 2, panel B (presented values are the standardized estimates), a cross-lagged effect was obtained from initial contact to subsequent change in prejudice, $\beta = -.28$, $p < .001$, but not from initial prejudice to subsequent change in contact, $\beta = .09$, *ns*. More specifically, the higher the level of contact at Time 1, the stronger the subsequent decrease in prejudice. Finally, also change in contact was significantly negatively related to change in prejudice, $r = -.23$, $p < .01$, demonstrating that contact and prejudice have opposite developmental patterns. In other words, increasing and decreasing levels of contact are significantly associated with decreasing and increasing levels of prejudice, respectively.

Discussion

The goal of the present study was twofold. First, we wanted to validate a self-report measure of intergroup contact with observer ratings provided by the respondents' close friends. Second, we aimed to demonstrate longitudinal effects of intergroup contact on prejudice with self- and observer-reports of intergroup contact using a full cross-lagged panel design.

Validating Self-Reports with Observer-Reports

Many studies on intergroup contact may be subject to criticism because of their extensive use of self-report measures of intergroup contact, which are potentially biased. However, the present research demonstrates that previous research findings obtained with self-reports are unlikely to merely result from such biases. Indeed, consistent with the findings of Hewstone et al. (2011), we demonstrated considerable agreement between targets' self-reports and observers' ratings of targets' intergroup contact. The reported target-observer agreement was even higher than typically found in research using observer-reports of personality traits (Vazire, 2006; Watson, Hubbard, & Wiese, 2000), which might be explained by the fact that contact with other people is a well-observable feature. As such, both Hewstone et al. (2011) and the present results present evidence for the validity of self-report measures of intergroup contact, thereby reassuring previously reported contact effects based on self-reports.

It should, however, be stressed that observer-reports of intergroup contact cannot be considered a truly objective measure, free of any bias, but rather a useful and complementary source of information that is unlikely to contain the same systematic biases as self-reports (Paulhus & Vazire, 2007; Vazire, 2006; see also Hewstone et al., 2011). Indeed, observer-reports has some limitations as well, like the potential influence of social desirability or assumed similarity biases.

Our results showed that observers did not ascribe significantly higher levels of intergroup contact to themselves than to the target, which makes it unlikely that socially desirable responding had a decisive influence (see also Hewstone et al., 2011). Concerning the influence of assumed similarity, the present study indeed showed a significant correlation between the scores provided by the observers about the targets' and their own intergroup contact scores. While this similarity might be interpreted as a projection bias, it is at least equally likely to reflect true similarity among friends (Hewstone et al., 2011). Moreover, these correlations were smaller than the correlations between these same observer-reports and targets' self-reports. Hence, while a projection effect may have been at work, it can be considered unlikely that this effect has substantively affected our findings.

Overall, because systematic response biases occur within single observers, the use of multiple observers decreases the problem of response biases. Furthermore, multiple observers are also more likely to base their ratings of the target on different situations. Hence, the aggregation of observer-reports provides more valid information than the reports of every single observer (Kolar et al., 1996; Schimmack, 2010). Finally, the observers demonstrated a relatively high level of consensus between each other (i.e., inter-observer agreement) about targets' levels of intergroup contact compared to other studies using observer-reports (Paunonen, 1989; Vazire, 2006), thereby further supporting the validity of our findings.

Nevertheless, we acknowledge that we cannot rule out with absolute certainty the potential influence of biases. Therefore, to further validate self-reports of intergroup contact, future studies may also register participants' contact behavior in a laboratory setting, or may try to develop more objective measures of observed contact outside the laboratory. Such alternative approach can be especially useful to accurately determine the

amount and time participants spent with outgroup members. However, it may be limited in the ability to capture the psychological experiences of the quality of intergroup contact and cross-group friendships.

Intergroup Contact Reduces Prejudice

Pertaining to the second aim of this study, the results provided convincing longitudinal evidence for the prejudice reducing effects of intergroup contact. Indeed, moving beyond the findings of Hewstone et al. (2011), cross-lagged analyses with the observer ratings of contact yielded longitudinal effects on prejudice parallel to the effects obtained with self-reports. Moreover, in these analyses, we simultaneously controlled for the stability effects of contact and prejudice over time and the cross-sectional associations between contact and prejudice within each wave. As such, the present study does not only provide evidence for the predictive validity of the contact measures, but also provides a more rigorous test of longitudinal contact effects on prejudice than the regression analyses traditionally used in contact research (e.g., Brown et al., 2007; Eller & Abrams, 2003, 2004; Levin et al., 2003).

Sample 2 data also showed the longitudinal effects using the statistically superior technique of SEM with latent variables, which controls for measurement error, enabling the initial demonstration of measurement invariance of the constructs over time (Christ & Wagner, in press; Little et al., 2007). Moreover, LCM allowed us to explicitly investigate individual differences in absolute change in intergroup contact and prejudice over time, which had, to the best of our knowledge, not yet been done in other published longitudinal studies on intergroup contact. Unlike latent cross-lagged models, LCM represent changes in the latent variables as factors in the SEM, which makes it possible to estimate the variances of the latent change factors as parameters (Hertzog & Nesselroade, 2003).

In particular, the results demonstrated that, although the overall mean scores of contact and prejudice did not significantly differ between Time 1 and Time 2, significant inter-individual variability in absolute change in these factors emerged. As a result, we were able to show reliable inter-individual differences in the extent to which participants show a decrease or increase in contact and prejudice over time, making the modeling and prediction of these change factors meaningful. In particular, we found a significant effect of initial levels of intergroup contact on change in prejudice over time, indicating that participants with higher levels of intergroup contact at Time 1, exhibit a larger subsequent decrease in prejudice between Time 1 and Time 2. Additionally, a significant negative association emerged between the latent change factors of contact and prejudice, which demonstrates that increases and decreases in intergroup contact over time significantly relate to decreases and increases in prejudice over time, respectively.

In sum, whereas previous cross-sectional and most longitudinal studies have left room for alternative interpretations about the effects of contact on prejudice, the present findings lay a better foundation for inferring the basic tenet of the contact theory that contact leads to a decrease in prejudice (Allport, 1954; Brown & Hewstone, 2005; Pettigrew, 1998).

Somewhat surprisingly, we did not find a significant cross-lagged effect of prejudice on contact, which has repeatedly been reported in previous studies (e.g., Binder et al. 2009; Eller & Abrams, 2003, 2004; Levin et al. 2003). A plausible explanation for the lack of such self-selection effect may be attributed to the research setting and sample characteristics. In particular, (most) participants may not have been able to choose whether they engage in intergroup contact or not. Indeed, Sample 1 participants and probably most Sample 2 participants were first-year college students who followed obligatory theoretical and practical courses together with their fellow immigrant students

(the latter group constituted more than 10% of the student population). Moreover, during the interval between Time 1 and 2, students (including the participants) were assigned to small work teams for some courses, which might have influenced individual variations in interethnic contact, irrespective of initial levels of prejudice.

Potential Limitations

Before closing, some limitations of the present study should be acknowledged. A first limitation pertains to the fact that Sample 1 consisted mostly of female undergraduate students. As such, the generalizability of our findings concerning the validation of self-reports with observer-reports may be restricted. To address this issue, future studies should use more heterogeneous samples.

A second issue concerns the present procedure to gather the observer-reports. In particular, we instructed the participants of both samples to avoid any communication between targets and observers about the survey. However, some targets and observers may have discussed their surveys, thereby inflating the correlations between the self- and observer-reports of intergroup contact. Future research can avoid this limitation by simultaneously administering the self- and observer-reports, which can be achieved in a laboratory setting (e.g. Hewstone et al., 2011), or with online registration. Nonetheless, the use of longitudinal designs does not permit to exclude communications between participants, as they can talk about the content of the surveys at any time between the two measurements occasions.

Conclusion

The present study contributes to the contact literature by providing a double validation for the use of self-reports of intergroup contact. First, we demonstrated a significant agreement between self- and observer-reports of intergroup contact. Second, we showed significant longitudinal effects of intergroup contact on prejudice with both

measures using a full cross-lagged panel design. The use of latent change modeling further substantiated the contact theory. In particular, we obtained significant inter-individual variability in absolute change in participants' levels of intergroup contact and prejudice over time, further showing that individual change could mean an increase as well as a decrease in contact and prejudice. Finally, we demonstrated that more intergroup contact at Time 1 predicted stronger reductions in prejudice across time and that changes in contact and changes in prejudice were significantly and negatively associated. In sum, by addressing these important methodological issues that characterizes the contact literature, we can place greater confidence in the validity of many findings previously obtained with self-reports of intergroup contact.

Notes

1. Preliminary tests of the separate measurement models at Time 1 and Time 2, yielded very good fits, $\chi^2(4) = 5.85$, $p = .21$; RMSEA = .052; CFI = 1.00 for Time 1 and $\chi^2(4) = 3.04$, $p = .55$; RMSEA = .00; CFI = 1.00 for Time 2.

2. In order to scale the latent variables, variances were fixed to 1 and factor means to 0. These identification constraints were relaxed in conjunction with more restrictive models of measurement invariance.

3. Given that the multivariate LCM was constructed with exactly the same indicators and with an equal number of parameters as the full cross-lagged model, the fit of both models was identical, $\chi^2(27) = 39.96$, $p = .052$; RMSEA = .053; CFI = .99.

References

Allemand, M., Zimprich, D., & Martin, M. (2008). Long-term correlated change in personality traits in old age. *Psychology and Aging, 23*, 545-557.

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., et al. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *Journal of Personality and Social Psychology, 96*, 843-856.

Brown, T. A. (2006). *Confirmatory Factor Analysis for applied research*. New York: Guilford Press.

Brown, R., Eller, A., Leeds, S., & Stace, K. (2007). Intergroup contact and intergroup attitudes: A longitudinal study. *European Journal of Social Psychology, 37*, 692-703.

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology, 37,* 255-343.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships, 14*, 417-431.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456-466.

Christ, O., & Wagner, U. (in press). Methodological issues in the study of intergroup contact: Towards a new wave of research. In G. Hodson & M. Hewstone (eds.), *Advances in Intergroup Contact*. London: Psychology Press.

Costa, P. T. Jr., & McCrae, R. R. (1988). Personality in adulthood: A 6-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*, 853-863.

Deutsch, F. M., Sullivan, L., Sage, C., & Basile, N. (1991). The relations among talking, liking, and similarity between friends. *Personality and Social Psychology Bulletin, 17,* 406-411.

Dhont, K., & Van Hiel, A. (2009). We must not be enemies: Interracial contact and the reduction of prejudice among authoritarians. *Personality and Individual Differences 46,* 172-177.

Dhont, K., Roets, A., & Van Hiel, A. (2011). Opening closed minds: The combined effects of intergroup contact and Need for Closure on prejudice. *Personality and Social Psychology Bulletin, 37,* 514-528.

Eller, A., & Abrams, D. (2003). 'Gringos' in Mexico: Cross-sectional and longitudinal effects of language school-promoted contact on intergroup bias. *Group Processes and Intergroup Relations, 6*, 55–75.

Eller, A., & Abrams, D. (2004). Come together: Longitudinal comparisons of Pettigrew's reformulated intergroup contact model and the Common Ingroup Model in Anglo-French and Mexican-American contexts. *European. Journal of Social Psychology, 34*, 229–256.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18*, 639-657.

Hewstone, M., Judd, C. M., & Sharp, M. (2011). Do observer ratings validate self-reports of intergroup contact? A round-robin analysis. *Journal of Experimental Social Psychology, 47,* 599-609.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61,* 521−551.

Judge, T. A. & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology, 85*, 751-765.

Kiesner, J., Maass, A., Cadinu, M., & Vallese, I. (2003). Risk factors for ethnic prejudice during early adolescence. *Social Development, 12*, 288−308.

Kline, P. (1993). *Personality: The psychometric view*. London: Routledge.

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality, 64*, 311−337.

Levin, S., van Laar, C., & Sidanius, J. (2003). The effects of ingroup and outgroup friendship on ethnic attitudes in college: A longitudinal study. *Group Processes & Intergroup Relations, 6*, 76-92.

Little, R. J. A. (1988). Missing data in large surveys. *Journal of Business and Economic Statistics, 6*, 287–301.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002) To parcel or not to parcel: Exploring the question and weighing the merits. *Structural Equation Modeling, 9*, 151-173.

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *Journal of Behavioral Development, 31*, 357-365

McArdle, J. J., & Nesselroade, J. R. (1994). Structuring data to study development and change. In S. H. Cohen, & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological innovations (pp. 223–268)*. Hillsdalle, NJ: Erlbaum.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543

Moskowitz, D. S. (1986). Comparison of self-reports, reports by knowledgeable informants, and behavioral observation data. *Journal of Personality, 54*, 294-317.

Ostroff, C., & Schmitt, N. (1993). Configurations of organizational effectiveness and efficiency. *Academy of Management Journal, 36,* 1345-1361.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598−609.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology (pp. 224−239)*. New York, NY: Guilford Press.

Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology, 56*, 823–833.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49*, 65-85.

Pettigrew, T. F. (2008). Future directions for intergroup contact theory and research. *International Journal of Intercultural Relations, 32,* 187-199.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of Social Psychology, 25*, 57-75.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*, 751-783.

Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology, 78*, 582−593.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12,* 69–82.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* London, Chapman & Hall.

Schimmack, U. (2010). What multi-method data tell us about construct validity. *European Journal of Personality, 24*, 241–257.

Stephan, W, & Rosenfield, D. (1978). The effects of desegregation on racial attitudes. *Journal of Personality and Social. Psychology, 36,* 795-804

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality, 40*, 472−481.

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology, 95,* 1202−1216.

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*, 546–558.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology, 73,* 73-90.

Table 1. Descriptive statistics and correlations between the variables of Sample 1 at Time 1 (T1) and Time2 (T2)

				Intergroup Contact				Prejudice Indicators							
				Self reported		Observer ratings		Subtle racism		Outgroup attitudes		Stereotypes		General index	
		Mean (*SD*)	α	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
Intergroup contact															
Self reported	T1	2.85 (1.42)	.95		.86***	.71***	.67***	-.40***	-.49***	.42***	.42***	.03	-.24*	-.32*	-.45***
	T2	2.76 (1.19)	.94			.70***	.73***	-.45***	-.53***	.46***	.42***	-.02	-.34**	-.37**	-.49***
Observer ratings	T1	2.66 (1.00)	.93				.88***	-.32**	-.40***	.29*	.42***	.06	-.28*	-.23†	-.40***
	T2	2.61 (.99)	.94					-.31**	-.34**	.26*	.33***	.01	-.26*	-.23†	-.34***
Prejudice Indicators															
Subtle racism	T1	4.04 (1.04)	.88						.86***	-.79***	-.48***	.61***	.71***	.94***	.84***
	T2	4.09 (1.12)	.90							-.75***	-.62***	.45***	.71***	.82***	.93***
Positive putgroup attitudes	T1	4.26 (.98)	.85								.65***	-.52***	-.71***	-.87***	-.81***
	T2	4.02 (.99)	.84									-.27*	-.63***	-.53***	-.80***
Stereotypes	T1	3.69 (1.22)	.89										.72***	.79***	.56***
	T2	3.81 (1.19)	.90											.81***	.88***
General Index	T1	0 (1.00)													.85***
	T2	0 (1.00)													

Note. . † $p < .10$; * $p = .05$; ** $p < .01$; *** $p < .001$

Table 2. Descriptive statistics and correlations between the variables in Sample 2 at Time 1 (T1) and Time 2 (T2)

				Contact		Prejudice	
		Mean (*SD*)	α	T1	T2	T1	T2
Contact	T1	2.75 (1.27)	.93		.84***	-.31***	-.39***
	T2	2.66 (1.16)	.94			-.21**	-.38***
Prejudice	T1	4.45 (1.20)	.86				.78***
	T2	4.56 (1.04)	.87				

Note. ** $p < .01$; *** $p < .001$

Table 3. Testing longitudinal measurement invariance of the self-report measures of intergroup contact and prejudice with Sample 2 data.

	$\chi 2$	df	*p*	RMSEA	CFI	$\Delta \chi 2$	Δdf	ΔModels
Model 1	8.68	8	.37	0.022	.999			
Model 2	13.13	13	.44	0.008	1.00	4.46	5	2–1
Model 3	16.62	16	.41	0.015	.999	3.49	3	3–2

Table 4. Results of univariate Latent Change Models using the self-report data of Sample 2

Latent variable	Mean	Variance	95% Confidence interval		Range	
			Lower	Upper	Min	Max
Contact Level	2.89***	1.75***	0.24	5.54	0.71	7.23
Contact Change	-0.09	0.43***	-1.40	1.22	-2.26	1.87
Prejudice Level	4.06***	1.12***	1.94	6.18	1.11	6.15
Prejudice Change	0.10	0.38***	-1.13	1.33	-1.66	4.78

Note. *** $p < .001$.

Figure Captions

Figure 1. Cross-lagged model testing the longitudinal effects of intergroup contact on prejudice in sample 1 with self-report (A) and observed (B) levels of intergroup contact. Presented values are standardized coefficients, $^{\dagger}p < .10$ $^{*}p = .05$; $^{**}p < .01$; $^{***}p < .001$.

Figure 2. Latent cross-lagged model (A) and Latent Change Model (B) demonstrating longitudinal effects of intergroup contact on prejudice in sample 2. Presented values are standardized coefficients, $^{*}p = .05$; $^{**}p < .01$; $^{***}p < .001$.

Figure 1.

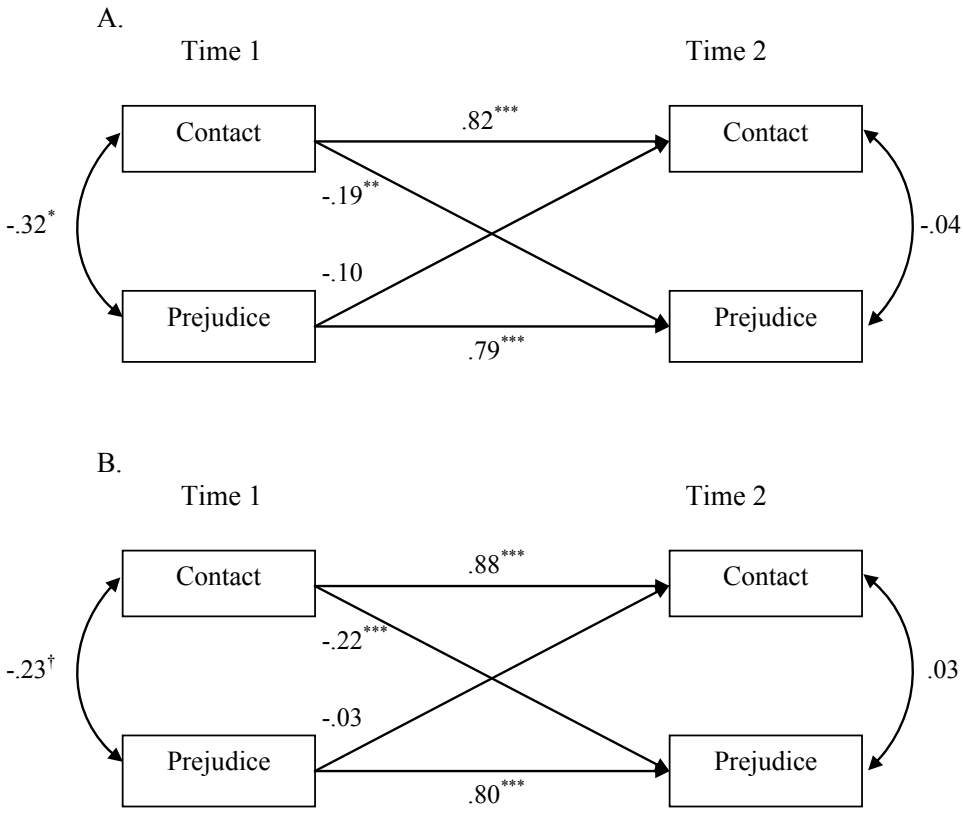
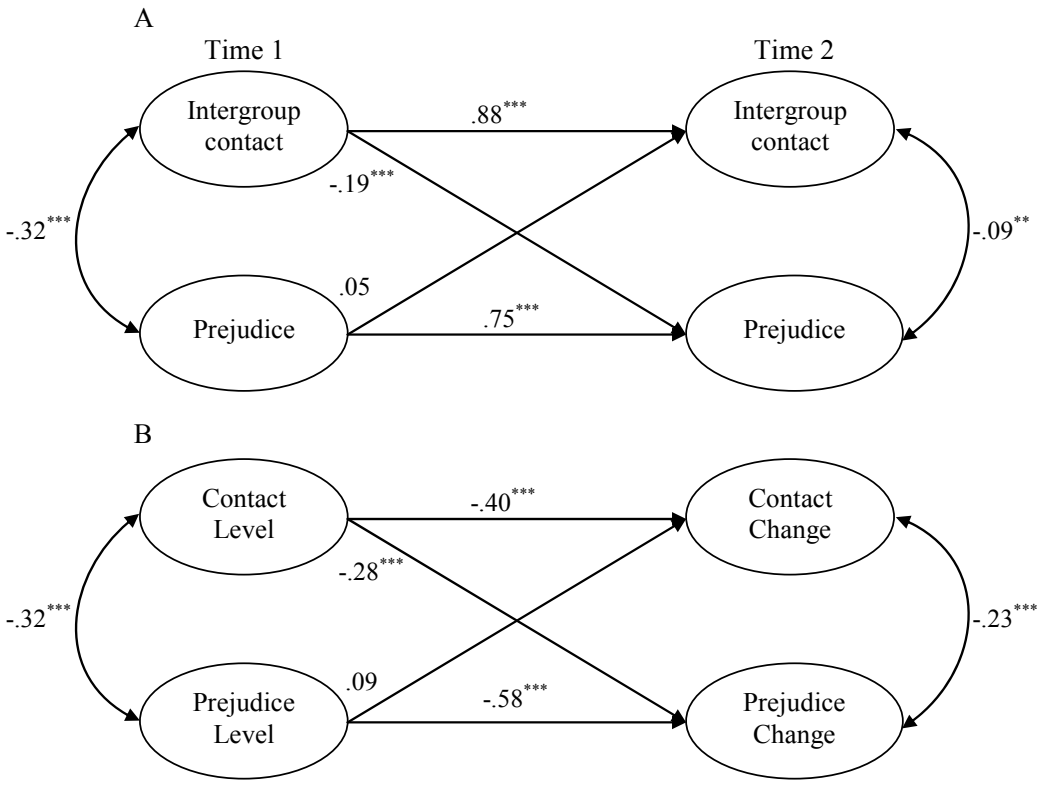

A.
Time 1
Time 2
Contact
Contact
Prejudice
Prejudice
.82***
-.19**
-.10
.79***
-.32*
-.04
B.
Time 1
Time 2
Contact
Contact
Prejudice
Prejudice
.88***
-.22***
-.03
.80***
-.23†
.03

Figure 2


A
Time 1
Time 2
Intergroup contact
Intergroup contact
.88***
-.19***
-.32***
-.09**
.05
.75***
Prejudice
Prejudice
B
Contact Level
Contact Change
-.40***
-.28***
-.32***
-.23***
.09
-.58***
Prejudice Level
Prejudice Change