Can Social Contact Reduce Prejudice and Discrimination?
Evidence from a Field Experiment in Nigeria*

Alexandra Scacco†
Scientist
WZB Berlin Social Science Center
alex.scacco@wzb.eu

Shana S. Warren
Ph.D. Candidate
New York University
shana.warren@nyu.edu

September 2017

Abstract

Can positive social contact between members of antagonistic groups reduce prejudice and discrimination? Despite the voluminous body of research on social contact, existing observational studies are difficult to interpret, due to a potentially severe selection problem: prejudiced people may deliberately avoid contact with out-group members. To overcome this problem, we conducted an education-based, randomized field experiment—the Urban Youth Vocational Training program (UYVT)—with 849 randomly sampled Christian and Muslim young men in the riot-prone city of Kaduna, Nigeria. After 16 weeks of positive social contact, we find no changes in prejudice, but a significant reduction in discriminatory behavior toward members of the religious out-group. We also find evidence of potentially negative consequences of in-group social contact. By focusing on practical skill-building instead of overt peace messaging, our intervention minimizes reporting bias and offers strong experimental evidence that social contact can alter behavior in constructive ways, even in a setting of ongoing violent conflict.

*We are grateful to Kate Baldwin, Chris Blattman, Eric Dickson, Pat Egan, Ryan Enos, Macartan Humphreys, John Jost, Rebecca Littman, Noam Lupu, Jack Snyder and David Stasavage, and participants in the Contemporary African Political Economy Research Seminar (CAPERS) and the NYU Center for Experimental Social Science, and especially to Bernd Beber, for their feedback, advice and support over the course of this project. Abel Adejor, Oluwatosin Akinola, Chom Bagu and Caleb Yanet provided excellent field assistance in Nigeria. Special thanks go to Kyauta Giwa, Community Action for Popular Participation (CAPP), Chima Nnaedzie, and microManna Ltd, our partners in implementing the computer training course intervention. The United States Institute of Peace (USIP) and the New York University Research Challenge Fund (URCF) provided funding for this study. This research was approved via NYU IRB Protocol 14-9985. Our pre-analysis plan is available via the Evidence in Governance and Politics (EGAP) Registry, ID 20150617AA. All errors and omissions are our own.

†Corresponding author
1 Introduction

Can grassroots interventions that increase contact between members of antagonistic groups reduce prejudice, discrimination and conflict? In spite of a vast literature on social contact in psychology and political science, and an explosion of NGO-led contact interventions in conflict settings around the world, basic questions remain about the consequences of intergroup contact in deeply divided societies. Does cooperative contact between individuals from across a deep social cleavage lead to reductions in prejudice and discrimination? How intensive must social contact be to induce positive effects? Does contact affect all groups involved in social interactions equally?

We conducted a field experiment—the Urban Youth Vocational Training (UYVT) project—to test whether sustained contact in an educational setting can improve communal relations in a conflict-prone environment. The UYVT intervention brought together a random sample of Christian and Muslim young men from disadvantaged neighborhoods in Kaduna, Nigeria, a city that has experienced repeated episodes of severe communal violence, for sixteen weeks of computer training. Our experimental design examined whether intergroup social contact can reduce prejudice and discrimination in a context of deep animosity. This study applies field experimental methods to a normatively important goal, the reduction of violent conflict and the promotion of post-conflict stability in deeply-divided societies.

To assess the impact of our intervention, we randomized (1) recruitment into the computer training program, (2) assignment to a religiously homogeneous or heterogeneous classroom, and (3) assignment to a co-religious or non-co-religious learning partner within the classroom. We measured prejudice—beliefs about individuals based solely on their group affiliation—through survey-based assessments of agreement with negative and positive stereotypes. We measured discrimination—differential treatment of individuals based purely on group affiliation—through two behavioral games embedded in our post-treatment survey: a dictator game and a destruction game. We find that though prejudice is resistant to change, intergroup contact can reduce discriminatory behavior. This suggests that contact can change behavior even without attendant changes in entrenched attitudes. Further, we find that part of this reduction can be attributed to a substitution effect—time spent with
members of the out-group reduces time spent with in-group members and thereby reduces opportunities for within-group bonding.

Studying intergroup contact between young men in Kaduna, Nigeria advances our broader understanding of ethnic conflict and peace-building in several ways. First, Nigeria has been a flash-point for Christian-Muslim communal conflict since the 1960s. Ethno-religious pogroms and reprisal attacks in 1966 killed as many as 30,000 civilians and displaced over one million people (McKenna, 1969), Christian-Muslim riots have resulted in an estimated 10,000 deaths in riots in 2000, 2002 and 2011, and recent Boko Haram bombings and reprisals in 2012 and 2014 have killed hundreds. This history of violence has resulted in high levels of ongoing tension typical of ethnic conflicts. Second, religious riots since 2000 have led to extreme residential segregation in Kaduna and other Nigerian cities. This pattern of post-conflict spatial segregation can be found in urban conflict zones around the globe, including cities in Bosnia and Herzegovina, India, Israel, Northern Ireland, and South Africa. Segregation limits contact across the religious divide, a situation that may deepen already prejudiced attitudes developed over years of localized conflict (Glaeser, 2005). Understanding whether grassroots social contact interventions can alter prejudices and decrease discriminatory behavior in such contexts is a key piece in the puzzle of how societies can recover from large-scale communal conflict and manage diversity. From a policy perspective, estimating the extent to which social contact has an independent effect on intergroup relations can help practitioners design more effective peace-building interventions.

This article begins by discussing the potential effects of intergroup social contact on prejudice reduction and peace-building, and questions the social psychology literature’s optimistic assessment of the contact hypothesis. It then reviews relevant features of the Nigerian context, our experimental design, and the type of social contact introduced in the UYVT program. We next present analyses of the effects of our intervention on intergroup attitudes (prejudice) and behavior (discrimination). Our data suggest that while out-group social contact does not affect prejudiced attitudes, it can meaningfully affect discriminatory behavior. The robust and significant effects of assignment to the intervention—a highly valued computer education program—demonstrate that overt peace education components of social contact interventions are not necessary to induce meaningful behavioral change. We con-
clude with implications for the design of social contact interventions and education-centered development programming in divided societies.

2 Social Contact, Prejudice and Discrimination

Scholarship linking intergroup contact and conflict behavior continues to draw heavily from the social contact hypothesis, initially outlined in Gordon Allport’s *The Nature of Prejudice* (1954). The hypothesis posits that interpersonal contact between individuals from hostile groups, if structured within a cooperative and egalitarian framework, should reduce prejudice, promote friendships across a social divide and, as a result, improve intergroup relations. For Allport, the behavioral stakes of prejudice were high. His work suggested that prejudice could lead to a range of pernicious behaviors, including out-group avoidance, discrimination and physical attacks across group lines. The contact hypothesis has been viewed as a promising policy tool to curb prejudice and intergroup hostility for decades, beginning with its application to school desegregation in the United States in the 1950s (Paluck, Green, and Green, 2017). In spite of the staggering volume of empirical research Allport’s hypothesis has inspired (e.g. Amir, 1969; Pettigrew, 1998; Gibson, 2004; Pettigrew and Tropp, 2006; Turner, Hewstone, and Voci, 2007), findings from existing research are difficult to interpret. Many studies link levels of self-reported cross-group interaction in daily life with self-reported levels of prejudice (e.g. Cehajic, Brown, and Castano, 2008; Semyonov and Glikman, 2009; Dixon et al., 2010). While suggestive, these studies face a potentially severe selection problem: prejudiced people may deliberately avoid contact with members of other social groups. Furthermore, participant awareness of the purpose of the study can result in inaccurate self-reporting about prejudice.

Research designs that attempt to eliminate reporting bias often focus on behaviors believed to be related to prejudice, such as electoral support for extremist parties (Kopstein and Wittenberg, 2009), but only experimental methods can fully overcome the selection problem. Most studies that attempt to circumvent both selection and reporting bias problems take place in rarefied contexts—such as higher education (e.g. Van Laar et al., 2005; Shook and Fazio, 2008; Carrell, Hoekstra, and West, 2015) or outdoor education programs (Green
and Wong, 2008) in wealthy democracies—that are far removed from the socially accepted intergroup animosity found in conflict settings where peace-building is needed most. Two recent experimental studies address social contact in contexts in which expressing prejudice is socially acceptable. Burns, Corno, and La Ferrara (2015) find reductions in racial bias among South African university students randomly assigned to non-co-ethnic roommates, while Barnhardt (2009) finds reductions in prejudice among residents of more religiously diverse public housing in India. These studies offer important contributions to the social contact literature, but differ considerably from ours in the populations they address. South African university students represent an elite group whose members likely differ in important respects from our own sample of disadvantaged youth in conflict-prone Nigerian neighborhoods, while Barnhardt (2009)’s sample focuses largely on female heads of household, a group unlikely to participate in communal violence. Reacting to these challenges, two recent reviews of the social contact literature (Paluck and Green, 2009b; Paluck, Green, and Green, 2017) express concern about the dearth of high quality field experimental studies on intergroup contact and prejudice.

The question of how social contact across conflict lines works to reduce prejudice and discrimination has received less attention in the literature (Green and Seher, 2003). A recent meta-analysis of empirical studies suggests three channels (Pettigrew and Tropp, 2008). First, as Allport originally suggested, social contact across cleavage lines may reduce prejudice by increasing knowledge about the out-group and revealing negative stereotypes to be false. Second, it may reduce anxiety about encounters with out-group members. Third, contact may result in increased empathy and perspective-taking. Existing literature implies that the “dosage” of social contact necessary to reduce prejudice may be relatively high, with more intimate contact leading to larger changes in attitudes than less in-depth exposure (Pettigrew, 1997, 1998; Pettigrew and Tropp, 2006; Turner, Hewstone, and Voci, 2007; Gibson and Claassen, 2010; Carrell, Hoekstra, and West, 2015). Our study evaluates these mechanisms and our experimental design leverages varying dosages of out-group social contact.

1Notably, while Van Laar et al. (2005) and Shook and Fazio (2008) rely on random roommate assignment to overcome the problem of selection into contact, in neither case do the authors present data on how roommates were initially paired.
In addition to fundamental questions about internal validity, the external validity of existing studies, and in particular their application to conflict and post-conflict environments, remains an open question. Repeated episodes of intergroup violence may increase the salience of relevant identity cleavages and harden prejudices against the out-group (Fearon and Laitin, 2000; Kaufmann, 1996; Posen, 1993). Ongoing violence produces more rigid boundaries between ethnic groups (De Waal, 2005), and can lead individuals to fear the physical proximity of members of the other group (Fearon and Laitin, 2000) [omitted for review]. Prejudice may not be as responsive to social engineering in the wake of serious conflict, as suggested in a carefully designed study involving prolonged exposure to a reconciliation radio program in Rwanda (Paluck and Green, 2009a). Observational evidence from Jerusalem reveals that when tensions are high, increased intergroup interactions can actually increase the probability of violence (Bhavnani et al., 2014). Similarly, Voigtländer and Voth (2012) demonstrate that anti-Semitic prejudice and cycles of violence have been perpetuated across many generations in Europe. It is reasonable to question whether we should expect short-term interventions to affect deeply held prejudices developed over many years of intergroup conflict and reinforced by ongoing violence.

Social contact may also affect discriminatory behavior, though few studies explicitly make this link. Experimental studies of discrimination in OECD labor and housing markets (e.g. Bertrand and Mullainathan, 2004; Ahmed and Hammarstedt, 2008; Kaas and Manger, 2012), attitudes about granting citizenship to immigrants in Switzerland (Hainmueller and Hangartner, 2013), and variation in taxi fare offers in Ghana (Michelitch, 2015) make clear that discrimination against members of minority groups is widespread in public life. Intergroup discrimination may be both pervasive and more damaging in conflict and post-conflict societies. Behavioral games have been used to identify discriminatory behavior in the form of lesser generosity toward ethnic out-group members in the Balkans (Whitt and Wilson, 2007; Mironova and Whitt, 2014) and Israel (Fershtman and Gneezy, 2001), among non-co-partisans in the United States (Iyengar and Westwood, 2015), and to measure pro-social behavior in conflict and post-conflict environments such as Nepal (Gilligan, Pasquale, and Samii, 2014) and Uganda (Baldassarri and Grossman, 2013).

In one of the few studies to directly test the impact of intergroup social contact on
discrimination, Carrell, Hoekstra, and West (2015) leverage random variation in the racial composition of assigned freshmen study groups at the U.S. Air Force Academy and identify a positive effect of intergroup contact on the probability of a white male cadet choosing a black roommate in his second year. Similarly, Malhotra and Liyanage (2005) find that (non-randomly-assigned) participants in a short-term Sri Lankan peace-building intervention were significantly more generous toward out-group members than non-participants, suggesting that discriminatory behavior may respond to contact interventions.

Members of groups in conflict often live under *de jure* or *de facto* residential and social segregation. By limiting contact with members of the out-group, segregation can intensify existing prejudices (Kunovich and Hodson, 2002; Glaeser, 2005; Enos and Gidron, 2016), limit cross-group trust (Kasara, 2013), undermine cross-group cooperation (Alexander and Christia, 2011), increase political participation due to fear of the out-group (Enos, 2016), and contribute to information failures that perpetuate cycles of conflict (Acemoglu and Wolitzky, 2014). Studies of ethnic violence reach similar conclusions. Conversely, military integration in the wake of ethnic warfare can decrease prejudice (Samii, 2013), and inter-ethnic networks decrease communal violence in India (Varshney, 2003). Social contact interventions may contribute to building such networks. This literature suggests that examining social contact in the context of extreme residential and social segregation—as is the case in our study—is a hard test for the reduction of prejudice and discrimination.

It is important to ask whether we should expect social contact interventions to have similar effects across groups. Studies of support for extremist parties in inter-war Czechoslovakia (Kopstein and Wittenberg, 2009) and the effects of the truth and reconciliation process in South Africa (Gibson and Claassen, 2010) demonstrate that the impact of social contact on prejudice and discrimination can vary widely across ethnic or racial groups. These observational studies suggest that social contact may not affect individuals across a religious or ethnic divide in the same ways, particularly in post-conflict contexts. Experimental studies in South Africa (Burns, Corno, and La Ferrara, 2015) and India (Barnhardt, 2009) identify positive effects of contact among socially dominant groups only. Dixon et al. (2012) similarly point to divergent effects of social contact for dominant and subordinate groups.

Unpacking the relationship between prejudice, contact and conflict has important impli-
cations for policy. Since Allport’s landmark study, practitioners have repeatedly attempted to use forms of positive social contact to improve intergroup relations. The past two decades have witnessed a rapid proliferation of grassroots peace-building initiatives in conflict environments around the world. The goals of these people-to-people interventions are ambitious, aimed at achieving societal transformation through micro-level attitudinal change across both sides of the social divide. Whether through mixed Jewish and Arab tango classes in Israel, gardening projects in the Palestinian territories, inter-ethnic soccer groups in the former Yugoslavia, reconciliation committees in Rwanda, or integrated Christian-Muslim basketball leagues in Nigeria, these interventions are driven by the premise that macro-level peace and stability can be built from the ground up (Maoz, 2000; Griesbeck, 2004; Gasser and Levinsen, 2004; Kuriansky, 2007; Vásquez, 2009; Cárdenas, 2013). The implied causal chain underpinning these projects is a reduction in prejudice and discrimination, leading to an increase in intergroup cooperation and trust, a decrease in conflict and, ultimately, long-term peace and stability.²

Contact-based peace education programs rarely collect systematic data about participant attitudes and behavior before or after the program, making it difficult to assess impact. Most NGO projects rely on convenience sampling and lack a control group. As a result, participants are likely to differ from non-participants in important ways. They may be especially open to new experiences, more cooperative, or less prejudiced against members of the out-group. Further, because most NGO-led projects bundle social contact together with peace education or diversity training, it is impossible to establish a causal relationship between social contact and outcomes. In the next section, we discuss how our intervention helps overcome these challenges to inference.

3 Research Design

Although the core claim of social contact theory—that positive and equal-status social contact with members of the out-group should decrease prejudice—is widely applied in peace-building programs, to our knowledge, the theory has never been directly tested using an

²This logic is not without its critics. For example, even if social contact decreases prejudice at the individual level, it may fail to do so in the aggregate (Forbes, 1997).
empirically rigorous field experiment in an ongoing conflict environment. Our study drew on best-practice sampling techniques to access hard-to-reach populations in conflict-prone areas, a randomized experimental design to assign subjects to treatment, and took multiple steps in survey design and implementation to minimize reporting bias. We conducted a baseline survey in August 2014, the UYVT computer course ran for 16 weeks from September to December 2014, and we conducted an endline survey in January 2015.

3.1 Research Context: Christian-Muslim Relations in Kaduna

Kaduna, a city of more than two million people, is the capital of Kaduna state, and sits at the crossroads of Nigeria’s predominantly Muslim North and predominantly Christian South. Religion is arguably the most salient social cleavage in contemporary Nigeria (Okpanachi, 2010; Lewis, 2007; Lewis and Bratton, 2000; Falola, 1998), and religious identity is certainly the most salient social division in Kaduna state and Kaduna city, where it is reinforced by coinciding ethnic cleavages (Wapwera and Gajere, 2017; Angerbrandt, 2011; Ibrahim, 1991, 1989). Although ethnic and religious demography is controversial in Nigeria, country experts estimate that Muslims comprise a slight majority in both Kaduna state and Kaduna city (Abdu and Umar, 2002; Sani, 2007). While not as poor as Nigeria’s far North, Kaduna state is considerably less prosperous than southern Nigerian states, with higher levels of unemployment, lower average per capita household income, and worse performance on a range of other socioeconomic indicators (Nigerian Annual Bulletin of Statistics, 2001, 2005, 2010).

Survey responses from our study confirm the salience of religion within our sample. Among Muslims, 97% reported going to mosque five times daily in both the baseline and

---

3 According to the 2006 Pew Forum on Religion and Public Life, 76% of Nigerian Christians and 91% of Muslims say that religion is more important to them than their identity as Africans, Nigerians or members of an ethnic group. (Okpanachi, 2010)

4 Within Kaduna, the Muslim population is largely comprised of members of the Hausa ethnic group, while the Christian population consists of members of numerous smaller ethnic groups, including Igalas, Gbagyis, Katafs and many others. (Wapwera and Gajere, 2017). Similarly, within our sample, Hausa is the mother tongue for 81% of all Muslims, but less than 2% of Christians. No single home language accounts for more than 11% of Christians in the sample.

5 No reliable nationwide census has been conducted since 1963. Results from the 1991 census were “corrected” in the wake of violent protests in southwest and northern Nigeria. The most recent 2006 census omits tribal and religious identification.
endline surveys. Similarly, 97% of Christian respondents in both surveys reported going to church at least weekly and over 93% cited a specific denomination. Furthermore, 98% of Muslims in our sample had received Koranic education and over 80% belonged to a Muslim brotherhood, a further commitment to one’s religious identity. These results reflect widely understood norms of religious participation to which young men in Kaduna overwhelmingly adhere. Even if these responses reflect a combination of both social pressure and religiosity, it is clear that, within our sample, religious observance and identity are important to both Muslims and Christians. In addition, residential segregation in Kaduna is based on religion rather than ethnic affiliation. Data from the baseline survey make clear that, within Christian neighborhoods, no single ethnicity represents more than one quarter of residents.

Kaduna state is known internationally as a hotbed of violent inter-religious conflict. Human Rights Watch estimates that as many as 10,000 people have been killed in such violence in the region since 1999 (Tertsakian, 2003, 2005). In February 2000, deadly Christian-Muslim riots shook Kaduna. The fighting began in the wake of public debates over introducing Shari’a law into the Kaduna state criminal code. Although Shari’a provisions had long been incorporated into “personal” or domestic law for Muslims throughout northern Nigeria, the debate raised concerns that Shari’a would be imposed on Christian communities (Abdu and Umar, 2002). The riots began when an anti-Shari’a demonstration passed through Kaduna’s crowded central market, located in a Muslim neighborhood. Rioting lasted for four days, and was only put to rest through military intervention. A state-led judicial commission of inquiry reported 1,295 deaths, although other sources have suggested the numbers may be far higher (Tertsakian, 2003). In addition to the death toll, dozens of churches and mosques, and entire city blocks were burned to the ground. Conservative estimates suggest that at least 125,000 people were temporarily displaced by the conflict (Angerbrandt, 2011).

Kaduna has subsequently experienced smaller-scale Christian-Muslim riots, in 2002 and 2011, and has experienced Boko Haram attacks since 2012, including suicide bombings carried out against churches in Kaduna and the surrounding region in April and June 2012. The bombings killed 57 people, and at least 92 people died in Muslim-Christians clashes that ensued in the aftermath (Gambrell, 2012; Madu and Brock, 2012). In July 2014, just a month before our study began, bombings targeting opposition presidential candidate Buhari and a
large Muslim prayer gathering killed at least 82 people (Garba, 2014). The fatalities and
destruction wrought by these events make them worthy of scholarly focus, and have drawn
in dozens of local and international NGOs with an interest in conflict prevention and youth
programming. These deadly conflicts have resulted in patterns of residential segregation
that have further diminished inter-religious social contact. Kaduna city is physically and
symbolically divided by a river with few crossing points. Muslims live to its north, Christians
to its south (Wapwera and Gajere, 2017). Kaduna therefore offers an ideal research setting
to study the effects of positively structured contact on intergroup relations.

Important socio-economic and political differences continue to divide Christians and Mus-
lims in Kaduna. Our baseline survey data echoes existing evidence that Christians are on
average better-educated and wealthier than Muslims living in Kaduna (Angerbrandt, 2011).
Muslims come from significantly larger households than Christians, with an average house-
hold size of 10.7 versus 5.6. Subjective measures of poverty paint a similar picture, with
Christian respondents significantly less likely to view their households as “poor” relative to
others in Kaduna.\footnote{These differences are significant at the $p < .05$ level.}

In contrast, Muslims have held most senior political posts since the creation of Kaduna
state in 1967. State governors in Nigeria control vast resources, making gubernatorial elec-
tions highly anticipated and often contentious. Only one of Kaduna’s 20 governors has been a
Christian. The composition of the current Kaduna State House of Assembly (elected in April
2015) reveals a similar pattern; Muslims won 24 of the 34 seats and hold all key leadership
positions. Descriptive representation has thus been highly uneven across religious groups.
These socio-economic and political differences highlight the importance of treating Chris-
tians and Muslims as distinct, rather than interchangeable, participants in our experiment.
Throughout our analysis, in addition to presenting aggregate results, we report separate
results for the Christian and Muslim portions of our sample.

### 3.2 Survey and Sampling Design

The goal of our study is to make inferences about the effects of intergroup contact on indi-
viduals in conflict zones, not simply people who volunteer for peace-building programs. We
therefore randomly selected study participants from among the residents of the poorest and most conflict-prone neighborhoods in Kaduna. Since it is typically young men who carry out violence, we restricted our sample to men aged 18 to 25 at the time of the baseline survey.\footnote{See Online Appendix A.1 for further details on the exclusion of women from this study.}

Sampling proceeded as follows. First, we sampled neighborhoods. Since there are no official neighborhood-level data or administrative boundaries, we compiled a list of neighborhoods and their approximate boundaries using data from \cite{omitted for review} and with the help of local NGO staff from our project implementation partner, Community Action for Popular Participation (CAPP). We included all neighborhoods within the Kaduna metropolitan area that are located within an hour’s commute of the centrally-located UYVT course site. We then assessed neighborhoods on two dimensions: We used expert evaluations described in \cite{omitted for review} to identify neighborhoods that had experienced violence in the past, and we used enumerator field assessments to construct a poverty index. Finally, we selected the 16 poorest neighborhoods from among those that had experienced violence in the past.\footnote{We excluded six neighborhoods suspected of harboring active Boko Haram cells. For further details on the neighborhood sampling process, including this exclusion, see Online Appendix A.1.} We focus on poor neighborhoods because violence there is overwhelmingly due to clashes between local residents as opposed to targeting that can occur in wealthier neighborhoods \cite{omitted for review}.

Second, we subdivided these neighborhoods into 46 enumeration areas (EAs), excluding any industrial areas, that could be easily traversed by an enumerator team in a single day. We set each EA’s sample size proportional to its density-weighted area, which we use as a proxy for population size in the absence of census data. We coded an area’s approximate density based on road penetration and the extent of open space, using aerial images from Google Maps and a 2011 government road map of metropolitan Kaduna.

Third, we randomly sampled study participants within enumeration areas. Enumerators followed random walk instructions to select residential plots and used random number lists to select households within plots and subjects within households.\footnote{Enumerators were trained and supervised on-site by the study authors.} The surveys were introduced as studies of the impact of vocational training. At the end of the baseline survey, enumerators collected addresses as well as mobile numbers of friends and family members to help locate
respondents for subsequent rounds of the panel. This contact information was separated in the field from the main body of the survey to maintain respondent privacy.\textsuperscript{10}

Enumerators interviewed respondents at their homes for both the baseline and endline surveys. To minimize problems of reporting bias in asking sensitive questions about prejudice and discrimination, respondents filled in answers to the most sensitive outcome measures themselves without enumerator observation. Our aim was to make it as difficult as possible for anyone other than the survey respondent to learn or guess answers to sensitive questions. While enumerators read questions aloud, respondents filled in simple answer bubble sheets themselves. When finished, they placed their answer sheets into a manila envelope containing other answer sheets (some of which were decoys). Once these answer sheets were separated from the rest of the questionnaire and placed in the envelope, they could only be re-matched to the rest of the questionnaire with a code key held by the study authors.\textsuperscript{11} The survey contained no skip patterns that might help enumerators (or others outside of the survey) infer information about respondent answers.

\subsection*{3.3 Experimental Design}

Our objective was to design an intervention to evaluate the effects of structured contact across the religious divide. With that goal in mind, the Urban Youth Vocational Training (UYVT) program offered 16 weeks of computer training in a small-group setting to Christian and Muslim male youth. To assess the impact of our program on prejudice and discrimination, we introduced randomization at three levels: (1) recruitment into the training program, (2) assignment to a religiously homogeneous or heterogeneous classroom, and (3) assignment to a co-religious or non-co-religious learning partner within heterogeneous classrooms, as shown in Figure 1.

\begin{figure}[h!]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Figure 1 about here}
\end{figure}

We randomly sampled a total of 849 young men between the ages of 18 and 25. Within this sample, 549 randomly selected subjects were invited to join the UYVT program and

\textsuperscript{10}For an in-depth discussion of this method of privacy protection in a panel survey, see [omitted for review].

\textsuperscript{11}An independent post-survey audit confirmed that enumerators followed sensitive question protocols.
300 served as a control group, participating only in the survey components of the study. Approximately one third of the UYVT participants were assigned to religiously homogeneous classrooms, and the remaining two-thirds to heterogeneous classrooms. Within classrooms, UYVT participants were randomly assigned to a partner from their own or the other religious group, with whom they worked in close cooperation on course assignments and custom-designed partner activities for the duration of the program. We also stratified classroom and partner assignment based on three additional demographic measures from the baseline survey to promote a positive social contact experience: prior computer experience, educational attainment, and prior out-group exposure, as measured by the frequency of out-group invitations to one’s home.

This paper makes three primary comparisons, one at each level of treatment assignment outlined in Figure 1. These comparisons include UYVT assignment versus control (Groups A and B), class structure (religiously heterogeneous versus homogeneous class) within course assignment (Groups C and D), and partner type (co-religious or non-co-religious partner) within heterogeneous classrooms (Groups E and F). These three comparisons enable us to make inferences about average program effects, average out-group social contact effects, and high versus low dosage out-group social contact effects. In all our analyses, we estimate intent-to-treat (ITT) effects, which are conservative estimates of the magnitude of the effects of our treatments on those who complied with their treatment assignment.\textsuperscript{12}

Since the main goal of our study is to test whether social contact decreases prejudice and discrimination, our main comparison of interest is the heterogeneous versus homogeneous class assignment (Groups C and D). This comparison varies social contact while controlling for non-intergroup contact induced program effects. We include the Group A vs. Group B analysis for comparability to policy-oriented and other existing research. We summarize our core hypotheses and predictions below:

**Hypothesis 1** UYVT assignment increases generosity (altruism).

There will be a positive program effect on generosity for respondents assigned to any arm of the UYVT program in comparison to the control group (Group A vs. Group B and its

\textsuperscript{12}The complier average causal effect (CACE), or local average treatment effect (LATE), is always larger than the ITT under one-sided non-compliance (Gerber and Green, 2012).
Hypothesis 2 Out-group social contact reduces prejudice and discrimination.

Out-group social contact will lead to a decrease in prejudice and discrimination for respondents assigned to (1) any heterogeneous UYVT treatment arm in comparison to the control group (Group A vs. Group B and all component sub-groups C, D, E and F) and (2) any heterogeneous classroom treatment arm in comparison to homogeneous classroom assignment (Group C vs. Group D and its component sub-groups E and F).

Hypothesis 3 Higher doses of social contact produce larger reductions in prejudice and discrimination.

There will be therefore be an additional decrease in prejudice and discrimination for respondents assigned to heterogeneous pairs vs. homogeneous pairs within heterogeneous classrooms (Group E vs. F).

3.4 Treatment

The UYVT training program structured participant interaction in a basic computer skills class under the supervision of three experienced teachers, one Muslim and two Christian. Homogeneous classes were taught exclusively by a teacher from the same religion as the students. There were 30 course sections: 20 religiously heterogeneous and 10 homogeneous. Each section met twice weekly for a total of four hours per week over 16 weeks. Students remained within the same classroom working with the same partner on a shared laptop for the duration of the course. The curriculum focused heavily on cooperative activities performed jointly by learning partners during each of 29 class sessions. Course topics included basic knowledge of MS Windows™, MS Office™, and introductions to Internet resources such as email, Skype™, and free online educational content. Since over 40% of the sample had never used a computer at all prior to UYVT, and two-thirds had previously used a computer less than once per week, this content was highly valued.

The co-authors personally reviewed the course lecture slides and designed many of the activities. Course materials are available from the authors upon request.
Class sessions were organized to maximize assigned partner interaction through fun, hands-on learning activities. At the beginning of each session, teachers lectured for approximately 30 minutes. The remainder of class time was devoted to partner work, with guidance from teachers. Partners designed flyers that could be used to advertise a future computer course, computed FIFA and UEFA soccer team and country rankings (a topic of great interest to students), researched the West African Ebola crisis, and produced presentations on the country they would most like to visit. To avoid reporting bias, students and instructors were not informed about the main purpose of our study, but instead experienced UYVT as an educational empowerment program targeting disadvantaged communities in Kaduna. By design, no component of the curriculum involved explicit prejudice-reduction or anti-violence messaging.

Our design incorporated incentives for UYVT participation to allow even the poorest students to attend class regularly. First, students acquired highly desirable computer skills necessary for office employment and higher education in Nigeria free of charge. Second, the course met only four hours each week, facilitating participation among students who were also working or in secondary school. Third, students who attended at least two-thirds of all sessions participated in a randomly-drawn raffle to distribute the 25 laptop computers used in the course.\footnote{We conducted the raffle after the endline survey to ensure it did not contaminate results.} Fourth, to keep students engaged with the course material, the curriculum prioritized hands-on learning. Fifth, students received a transport stipend of ₦200 (200 Nigerian Naira, equivalent to one USD) at the end of each class. This was important, as many students lived relatively far from the city center and would not otherwise have been able to afford to reach the centrally-located course site. Sixth, we implemented the course at microManna, the largest computer retailer and training center in Kaduna, which provided office-space and equipment for the project. In a context where corruption and false advertising are common, this affiliation helped assure sampled participants that enumerators were offering a genuine program.

Given the heavy emphasis placed on learning partnerships in the course, it is important to ask whether this aspect of the UYVT treatment worked as planned. Did students ultimately get to know their partners well enough to constitute a meaningful test of the contact
hypothesis? Data from student course evaluations and the endline survey strongly suggest that they did. When asked about their learning partners on end-of-term course evaluations, 69% of UYVT students claimed to have gotten to know them “very well” (with 24% replying they had gotten to know their partners “well,” and only 8% claiming that they had gotten to know their partner “a little,” “not very well” or “not well at all”). By comparison, 37% said they had gotten to know other students in their class very well. Taken together, these responses suggest that subjects took up treatment dosages in accordance with our study design—getting to know their partners best and other members of their class somewhat less well.

Course evaluation results indicate that the vast majority of students had positive experiences with their partners, with 94% responding that they believed working with a partner facilitated learning. Endline survey responses further suggest that UYVT students enjoyed their experience with their partners, with an impressive 92% responding that they had been in touch at least once since the end of the training course, and 88% reporting that they had saved their partner’s mobile number. In both the course evaluation and endline survey questions, responses were nearly identical from students in heterogeneous and homogeneous classes, suggesting that UYVT students across different experimental treatments experienced the training course as an enjoyable, high-quality educational program. Social contact under UYVT was cooperative, egalitarian and positive, providing opportunities to develop knowledge of the out-group, decrease anxiety around out-group encounters, and increase empathy and perspective-taking.

4 Data

Nearly all existing studies of social contact theory limit their hypothesis testing to changes in either attitudes or discriminatory behavior. While the implicit assumption is that prejudice reduction will lead to improved intergroup relations, that hypothesis generally remains untested. To address this knowledge gap, we measured both attitudes and behavior. The

---

15Course evaluations were completed online during class with teachers outside the classroom (N = 359).
16A qualitative researcher who evaluated UYVT classes every day for the duration of the course observed almost exclusively positive interactions between partners.
baseline survey excluded explicit questions about prejudice to avoid priming survey participants to the main purpose of our study. The post-treatment endline survey included prejudice and discrimination outcomes for all study subjects measured four to six weeks after the conclusion of the UYVT course. Online Appendix Table A.2 presents descriptive statistics for all variables analyzed in this paper.

4.1 Prejudice

We follow the literature in defining prejudice as a set of negative beliefs about or attitudes toward an individual based solely on membership in a particular social group. To measure prejudice, we modified Likert scale survey questions that ask whether survey respondents agree with negative stereotypes; the questions were modeled after previous studies in psychology and political science (e.g. Hewstone et al., 2006; Gibson and Claassen, 2010; Paluck, 2010) to fit the Nigerian context. Respondents assessed how well a list of adjectives described non-co-religious individuals in general, using both positive and negative attributes (asked with negative and positive attributes interspersed). We also asked respondents to identify where along a five-point scale they would place members of the out-group, with five being associated with the more positive attribute in each of three adjective pairs. We combined these responses into three indices, as shown in Table 1: a Negative Attributes Index, a Positive Attributes Index and an Out-group Evaluation Index.\footnote{For details on survey instrument development, see Online Appendix A.2.}

[Table 1 about here]

All three of these indices are conceived as measures of prejudice. We do not mean to suggest that they reflect separate underlying conceptual constructs. The negative attributes measures follow the majority of existing social psychology literature on prejudice, while we modeled the out-group evaluation measures on feeling thermometers more commonly used in the political science literature. A key reason why we asked about positive attributes was

\footnote{We conducted an exploratory factor analysis to determine if all items could be combined into a single scale, but found that the scale retained three dimensions corresponding to the three indices. After confirming the uni-dimensionality of each scale following best-practice techniques outlined in (Furr and Bacharach, 2013), we calculated Cronbach’s alpha for each index to confirm that all items are measuring the same construct. Details on the psychometric testing are available in Section A.7 of the Online Appendix.}
to ensure that the survey did not prime respondents, create bias, or promote a negative view of the UYVT intervention or the survey more generally by focusing only on negative descriptions of the out-group. Most importantly, we wanted to allow respondents multiple opportunities to express their explicit prejudices across several widely used survey question types, with the understanding that prejudice can be expressed via disagreement with positive stereotypes as well as agreement with negative ones (Brown, 2011). Distributions of all three indices are included in Online Appendix Figure A.4. To further examine prejudice, the survey also included questions to elicit the extent of knowledge about the out-group, anxiety about contact, empathy toward out-group members and desires for cross-group friendships.

4.2 Discrimination

Discrimination occurs when treatment of individuals from the out-group differs from that of individuals from the in-group based solely on group affiliation. We measured discrimination through two behavioral games embedded in the endline survey: a dictator game and a destruction game. In the former, our measure is based on differential positive behavior (altruism), the latter measures differential negative behavior (destruction). Within behavioral economics, the dictator game has been utilized to measure altruism and norms of fairness (including discrimination) and we suggest that the destruction game is its mirror image.\(^\text{19}\)

In the dictator game, participants chose how to divide ₦100 (about .50 USD) from a common pool with another survey participant. We primed respondents with the first name of this individual, taking advantage of a convenient aspect of Nigerian first names in Kaduna: that they clearly and unambiguously signal religious affiliation. Among those assigned to a UYVT class, we also indicated whether the named individual was a UYVT classmate.\(^\text{20}\) Each respondent played ten rounds of each game, and the order in which a UYVT-assigned respondent was matched with classmates versus other survey respondents was random. Respondents were told that game play was non-reciprocal (e.g., being asked to divide ₦100 with Abdullahi does not necessarily mean Abdullahi will be asked to divide

\(^{19}\)See Camerer (2003) for a comprehensive review of the empirical literature and Whitt and Wilson (2007) for studies that focus on using the dictator game to measure norms of fairness.

\(^{20}\)For example, a prime could be “Abdullahi from your UYVT class” or simply “David,” without further information. For further information on the primes and complete games instructions see Online Appendix Section A.4.
N100 with you). This design mitigated concerns about retribution.\footnote{We did not use deception in the experiment. Further information concerning payouts and the absence of deception is available in Online Appendix Section A.4.}

There was substantial variation in this measure, and the distribution of responses was similar across Christian and Muslim respondents.\footnote{See Online Appendix Figure A.8.} Respondents gave an average of 28\% of their endowments to the recipient in each round of play, in line with the approximately 20\% given in Camerer (2003)’s meta-analysis of dictator game play (56–58), the 23–29\% given by respondents in Whitt and Wilson (2007)’s work in post-conflict Bosnia, and the 29\% given by respondents in Iyengar and Westwood (2015)’s study of partisan discrimination in the United States.

We designed the destruction game to mimic aspects of riot behavior, in which a potential participant obtains a small personal benefit at a larger cost to another person.\footnote{Abbink and Sadrieh (2009) implement a similar game absent any payoffs for destruction; Zizzo and Oswald (2001); Abbink and Herrmann (2011) implement an alternative version in which players absorb a small cost to destroy the other player’s money. The game is well-suited to assess discrimination in the context of intergroup hostility, in particular because religious rioting in Nigeria typically involves destructive rather than appropriative behavior.} As in our prejudice indices we elected to offer both positive and negative measures to elicit responses about discriminatory behavior. In the destruction game, we again prompted respondents with the first name of another survey participant. In each of ten rounds of play the respondent was allocated one or two N50 notes, as were his assigned opponents. The subject could receive an additional N10 for each N50 bill he chose to destroy from the other person’s money. As in the dictator game, there was wide variation in game play with at least some destruction in 66\% of rounds of play.\footnote{This result falls in line with previously observed 39\% destruction rate absent any pecuniary benefit (Abbink and Sadrieh, 2009), 26\% rate with a small cost (Abbink and Herrmann, 2011), and 63\% rate with highly variable initial allocations (Zizzo and Oswald, 2001).} The distribution of responses was again similar across Christian and Muslim respondents.\footnote{See Online Appendix Figure A.9.}

Both behavioral games were administered privately at respondent households during the endline survey. The full text of the instructions read to respondents, sample enumerator questionnaires and sample response sheets are included in Online Appendix A.3 Games Instructions.
4.3 Demographic Covariate Balance

In both the baseline and the endline surveys we collected data on an extensive list of demographic, economic, social and community engagement, and personality-trait covariates. We used these data for three important purposes: to perform a randomization check, to stratify treatment assignment, and to verify that enumerators had located the correct individuals in the endline survey. Our careful randomization process resulted in strong balance between treatment and control groups, as shown in Online Appendix Table A.1, including, for example, with respect to the extent to which respondents reported having been affected by the most recent 2011 riots.\textsuperscript{26} Across 44 pre-treatment covariates, difference-of-means tests were only significant at the $p < .1$ level in two cases, a theoretical question about risk aversion and one of the 16 neighborhoods, Narayi, in which a larger share of respondents were in the control group.\textsuperscript{27}

4.4 Attrition

Completing the second round of a panel survey of disadvantaged young men in a developing country poses serious challenges. Mobile phone numbers frequently change, and residency may change within a five-month period. During the baseline survey (September 2014), our enumerators interviewed 849 respondents and distributed a small survey stipend to compensate them for the time taken on the survey and help avoid panel attrition. Our enumerators successfully located 795 of these young men for the endline survey (January 2015), resulting in a very low overall attrition rate of 6%. Because we were able to update contact information for course participants, the attrition rate was 4% among those assigned to the UYVT treatment and 10% within the control group.\textsuperscript{28}

\begin{footnotesize}
\footnotesubscript{26}For additional information, see Online Appendix Section A.5.
\footnotesubscript{27}We repeat all of our analyses with a control for each imbalanced variable, neighborhood fixed effects and all variables listed in our pre-analysis plan in the Online Appendix.
\footnotesubscript{28}There were no significant differences in attrition between homogeneous and heterogeneous classes, or co-religious and non-co-religious partner assignments.
\end{footnotesize}
4.5 Compliance

At the conclusion of the baseline survey, respondents were assigned to either the UYVT treatment or the control group. If assigned to UYVT, respondents were offered the opportunity to participate in the course. At this time, they had not yet been assigned to any class schedule, class type (heterogeneous versus homogeneous), or partner type (same religion or different religion). Of the 549 participants assigned to treatment, five declined participation at the time of their survey interview. An additional 84 respondents never attended a single UYVT class session, resulting in a compliance rate of 84%. These non-compliers never knew if their class assignment would have been religiously mixed or homogeneous, that the course involved assignment to partners, or the religious identity of their partner. In fact, given that respondents were interviewed by members of their own religious group and the UYVT course site was located in the main commercial area in central Kaduna, respondents were not explicitly aware that they were being invited to a religiously mixed computer course. As such, those that they did not comply with their treatment assignment did not refuse in order to avoid social contact with people of the other religion, or because they found the course experience to be unpleasant.\(^{29}\)

In the endline survey, enumerators were able to interview 71 of the 89 respondents assigned to treatment who never attended the course. 59% of non-attenders interviewed said they were too busy with work, school, travel or family needs to attend. Others cited the distance to the course site, the timing of UYVT classes, illness and other logistical issues as reasons for non-attendance.\(^{30}\) A majority (71%) of the 89 non-compliers were Christian.\(^{31}\)

[Figure 2 about here]

The UYVT course involved 29 class sessions over four months. On average, students who attended at least one class attended 76% (22) of their class sessions, offering ample

\(^{29}\)Non-compliance is not correlated with the class or pair type treatment assignment (\(p = .83\) and \(p = .93\) respectively). For further discussion of balance and a table testing whether non-compliance is predicted by any covariates see Online Appendix Section A.8.

\(^{30}\)Only one respondent indicated that he had come to the course site and ‘did not like what he saw’.

\(^{31}\)To ensure religious balance in the program, we randomly recruited additional Christian participants during the first few days of the course.
opportunities to interact with assigned partners and other classmates. Figure 2 makes clear that only a small minority of students (12%) attended fewer than 10 sessions. There are no statistically significant differences in attendance rates by teacher or class type (heterogeneous versus homogeneous). There is, however, a statistically significant difference in attendance by religion. On average, Christians attended 21 sessions and Muslims attended 23 sessions. Since students of both religious groups attended, on average, over two-thirds of all class sessions, both groups achieved high levels of contact. Furthermore, as shown in Figure 3, aggregate attendance dropped off very little over time after the first few class sessions.\footnote{For session 26, unusually heavy rain kept many students away.}

![Figure 3 about here]

Evidence from the post-intervention survey strongly suggests that the student absences should not be attributed to the quality of course content, teachers, or social contact. Not a single respondent cited anything negative about other students or teachers as a reason for missing UYVT class sessions. Others cited work, school, or family obligations as reasons for missing class. Among students who attended any classes, and therefore were aware of the social dynamics of the UYVT course, compliance with assigned treatment was sufficient to ensure that they received the social contact treatment. Further, grouping all students who attended any sessions as treated biases against finding any effects of the UYVT course and social contact treatments.

5 Findings

We estimate intent-to-treat (ITT) effects with three levels of treatment in a between-subjects design: program assignment (UYVT versus control), class type assignment (homogeneous vs. heterogeneous) within the UYVT course, and course partner type (non-co-religious vs. co-religious) within religiously heterogeneous classes. All estimates are OLS regression results in which the treatment indicator variables represent assignment to the UYVT course (\textit{UYVT}), a heterogeneous classroom (\textit{Heterog. class}), or a non-co-religious course partner (\textit{Heterog.}
We were ex ante agnostic about the ways in which Christians and Muslims might be affected differently by these treatments, and we report results for the full sample as well as for Christian and Muslim subsamples. For all our analyses we specify parsimonious empirical models, relying on our random sampling and treatment assignment to control for potential confounders.

We estimate the following specification for each of the three prejudice indices:

\[ PrejudiceIndex_i = \alpha + \beta_1 Treatment_i + \varepsilon_i \]  

where \( i \) represents a respondent, the Treatment variable represents UYVT assignment (vs. control), heterogeneous (vs. homogeneous) class assignment, and non-co-religious (vs. co-religious) partner assignment within heterogeneous classrooms. All prejudice indices have been coded such that a positive estimated \( \beta_1 \) indicates a reduction in prejudice.

Subjects played ten rounds each of the dictator and destruction games, so we cluster standard errors by respondent, and we include round-of-play fixed effects in these analyses. For each round of play, subjects were randomly assigned to another individual in the study, either an in-group or an out-group member. Games were administered to each individual subject at the time of his survey interview, so subjects did not meet during game play, but were told others’ first names in order to prime religious affiliation. In order to see if the

---

33 This split-sample analysis does not directly show whether there is a statistically significant difference in effects across subsamples. We present results from specifications in which we interact treatments with religious affiliation in the Online Appendix.

34 We discuss and explore the possibility of heterogeneous treatment effects in Section A.15 of the Online Appendix, but do not identify any robustly significant effects.

35 We do not include teacher effects, because they induce collinearity in some class type comparisons. In Online Appendix Tables A.18, A.39, and A.48, we present estimates that control for teacher religion as a robustness check, in instances where it is possible to do so. Our substantive results remain the same.

36 Sections A.10 and A.11 of the Online Appendix include analysis using wild bootstrapped standard errors clustered by class for the prejudice indices and by both class and respondent for the dictator and destruction games, as a robustness check. These alternative clustering methodologies do not change our main results.

37 While these fixed effects absorb round-specific variation for the purposes of our main analysis, we note that first-round play differs somewhat from behavior observed in subsequent rounds. We explore this variation and provide a number of related robustness tests in Online Appendix Section A.12.

38 Random assignments were made within strata, to ensure that all subjects played with individuals of both religious groups.

39 Subjects assigned to UYVT were also primed with whether the named individual was a classmate. A summary of all survey questions and game instructions analyzed in this paper is available in Online Appendix 9.2. Further details are provided as part of our registered pre-analysis plan, available at [link omitted for review].
course treatments affect discrimination, we interact the treatment indicators with whether a subject was randomly assigned to play with a member of the in-group or the out-group via the following specifications:

\[
Action_{i,r} = \alpha + \beta_1 Treatment_i + \beta_2 Treatment_i \times PlayOutGroup_{i,r} + \beta_3 PlayOutGroup_{i,r} + \gamma_r + \varepsilon_{i,r}
\]  

(2)

where \(Action\) represents the number of bills given (dictator game) or destroyed (destruction game), \(i\) represents a respondent and \(r\) represents a round of play, the \(Treatment\) variable again represents assignment at the program, class and pairs level, and \(\gamma_r\) are round-of-play fixed effects.\(^40\) For the dictator game, a negative coefficient estimate for the \(Play\ out-group\) term \(\beta_3\) indicates discrimination in generosity, i.e. fewer bills being given to out-group recipients. For the destruction game, a positive coefficient estimate indicates for \(\beta_3\) indicates discrimination in destructive behavior towards out-group members. In both cases, the interaction term \(\beta_2\) is our coefficient of interest. In the dictator game, positive coefficient estimates for the interaction coefficient \(\beta_2\) indicate a reduction in discrimination due to the treatment. In the destruction game, negative coefficient estimates for the interaction coefficient \(\beta_2\) indicate a reduction in discrimination.

5.1 Prejudice

First, we analyze data about respondents’ attitudes toward members of the out-group. While lab and field experimental evidence from the United States predicts that extended social contact with members of an out-group should reduce prejudice, our findings from highly conflict-prone Kaduna, where negative attitudes toward the out-group are entrenched, are less encouraging. Even in a positive educational setting—an environment with high rates of student satisfaction and one in which students overwhelmingly agreed that they were being treated fairly and equally—we do not see significant reductions in prejudice due to either the

\(^{40}\)We are not able to include class fixed effects for class type comparisons, because of collinearity with the treatment indicator. For pairs type comparisons, results adjusted for class fixed effects are shown in Tables A.17, A.39, and A.48 in the Online Appendix. Estimated effect sizes are nearly identical, and signs and significance identical, to the results from the main analysis absent classroom fixed effects.
UYVT program or social contact with the out-group. Power analysis examining our primary
comparison of interest, homogeneous vs. heterogeneous class assignment, makes clear that
these null results are not due to lack of statistical power.\footnote{Assuming an anticipated effect size of $r = -.21$, the mean effect size identified in Pettigrew and Tropp (2006)’s landmark meta-analytic study of intergroup social contact and prejudice reduction, we have statistical power of over 90% in the full, Muslim and Christian samples for all three prejudice indices. Furthermore, our power analysis is quite conservative; Pettigrew and Tropp (2006) identify mean effect sizes of $r = -.34$ for experimental studies, $r = -.30$ for experimentally manipulated contact, and $r = -.25$ for outcomes scales with Cronbach’s $\alpha > .70$ (as is the case for all three of our Prejudice Indices).}

Results are shown in Tables 2, 3 and 4. The \textit{Negative Attributes Index}, \textit{Positive Attributes Index}, and \textit{Out-group Evaluation Index} all range from 1 to 5 and are coded such that higher values are desirable from the standpoint of intergroup reconciliation: they indicate agreement with positive attributes, disagreement with negative attributes, and positive evaluations of out-group characteristics. Table 2 and Table 3 clearly indicate that none of our experimental treatments reduced prejudice in any meaningful way in either the combined or split samples. At the overall program level, among Christians, we observe two contradictory results: UYVT-assigned Christian respondents were less likely to agree with both negative and positive assessments of Muslims.\footnote{Re-running our analysis on each of the components of the Negative and Positive Attributes indices, we replicate the null findings of the indices with respect to social contact. Assignment to heterogeneous (vs. homogeneous) classes did not significantly affect any of these eleven stereotyped beliefs.}

Analysis of the \textit{Out-group Evaluation Index} (Table 4) yields similar results. At the overall program level, we observe a substantively small, marginally significant ($p < .1$) \textit{negative} effect. The magnitude of the coefficient corresponds to less than one-fifth of a standard
deviation. This finding is driven by Christian respondents, as can be seen when we split our sample by religious group in Columns (2) and (3).

In summary, we find no evidence that the 16-week computer training course reduced prejudice among young men in Kaduna’s poorest and most conflict-prone neighborhoods, and we find no significant effects associated with being assigned to an intergroup social contact treatment (heterogeneous class) within this course or to a non-co-religious partner within a heterogeneous class. Prejudices remain entrenched and largely unaffected by any aspect of the UYVT intervention.\footnote{Out-group assessments in our data are not as negative as one might perhaps expect, in particular among Muslims, as indicated by group-level means shown in Online Appendix Table A.7. A natural concern then is that our null results stem from ceiling effects, but we think it is unlikely that this explains the essentially complete absence of prejudice-related effects. We present related statistics and analysis in Online Appendix Section A.6.1.}

Our endline survey contained a suite of questions designed to test prominent prejudice-reduction mechanisms in the social psychology literature on intergroup contact. Online Appendix Table A.19 presents the full set of questions we used to estimate whether exposure to the UYVT treatments helped increase participants’ knowledge about the out-group, reduce anxiety about out-group encounters, increase participant empathy and perspective-taking or increase interest in cross-group friendships. In each question, respondents were read a statement about the religious out-group. For example: “It is difficult for me to understand Christian (Muslim) customs and ways.” They were then asked whether they strongly agree, agree, disagree or strongly disagree with that statement. Measures for each question range from 1 to 4 and are coded so that higher values reflect normatively desirable effects of intergroup contact—they indicate agreement with positive statements and disagreement with negative statements.\footnote{Full results for prejudice mechanisms tests are available in Online Appendix Tables A.20 through A.30.}

Consistent with the analysis of our main prejudice indices, we find little to no evidence that any of the UYVT treatments had desirable effects along the lines predicted by prejudice-reduction mechanisms in the social contact literature. Beginning with questions about out-group knowledge, we find no increase (or decrease) in the degree to which UYVT participants of any type find it difficult to understand out-group customs. When asked whether they felt they had close friends from the out-group, Muslims in UYVT were slightly more likely to
agree than Muslims in the control group, but Christians were actually less likely to agree that they had close friends from the other religious group. Members of religiously heterogeneous pairs were also less likely to agree with this statement than members of homogeneous pairs.

Turning to mechanisms focused on anxiety about out-group encounters, we find similar results. There is some evidence that UYVT program assignment reduced anxiety for Muslims about spending time with Christians, but the training course had no effect on anxiety among Christians and actually resulted in a decrease in the number of Christians who reported that they would feel comfortable working alongside Muslims. Similarly, the UYVT intervention had no positive impact on perceptions of how rewarding it might be to get to know people of different faiths. In fact, those assigned to an out-group partner were significantly \( p < .05 \) less likely to report that they could imagine having close out-group friends. Finally, we find no evidence to suggest that any of the UYVT treatments led to increased empathy and perspective-taking across group lines. Out of the four empathy measures listed in Table A.19, the only significant finding shows a reduction in the degree to which Christian UYVT participants reported that they could see the good faith and devotion in the way that Muslims pray.

In sum, assignment to the UYVT program, to a heterogeneous UYVT class, or to a non-co-religious learning partner does not appear to have increased knowledge about the out-group, reduced anxiety about out-group interactions, or to have increased empathy across the religious divide. In fact, in the few cases where we find a significant impact, most of these results showed a negative impact of UYVT on prejudiced attitudes. The only statistically significant positive results are found among Muslim respondents. For Christians, the UYVT program at any level of comparison had either no significant effect on prejudice or had a negative impact.\(^{45}\) Prejudiced attitudes in our study context appear resistant to change, but as we demonstrate below, this does not preclude changes in actual behavior.

\( ^{45} \)Further robustness tests adding controls for minor imbalance and those specified in our pre-analysis plan lead to the same conclusions as shown in the Online Appendix.
5.2 Dictator Game

We now turn to our analysis of the behavioral experiments embedded in the endline survey. We show dictator game results in Table 5. We again provide estimates for the sample as a whole and each religious group separately, and we estimate program effects (UYVT assignment), social contact effects (heterogeneous instead of a homogeneous class assignment), and social contact dosage effects (heterogeneous instead of homogeneous pair assignment within a mixed classroom). We now also interact our treatment indicators with a variable identifying rounds of play in which respondents were randomly assigned a member of the religious out-group. This allows us to estimate treatment effects on generosity toward in-group and out-group members as well as discrimination between members of these groups.

Assignment to the UYVT course had a positive and highly significant effect on generosity toward both co-religious and non-co-religious recipients in the dictator game, as shown in columns (1)–(3) of Table 5. Across the full sample, assignment to the training course increased the average transfers by Muslim and Christian respondents to both in-group and out-group members by approximately ₦4.46 As we would expect, subjects generally give less to out-group members, and the UYVT treatment does not significantly change that fact.47 Thus, assignment to the UYVT program alone does not reduce discrimination, as is clear in the insignificant interaction terms in columns (1)–(3).

Independent of UYVT class religious composition, the vocational training course similarly provided a sense of good fortune in having been selected, a positive and personally beneficial experience, close social contact with others, and perhaps an expectation of higher income in the future. This wealth and good fortune effect manifested in increased generosity to both

---

46 This corresponds to an increase from approximately ₦26 to ₦30 for in-group members, and ₦24 to ₦28 for out-group members.

47 Similarly, Whitt and Wilson (2007) observed preferential in-group treatment in lab experimental work in post-conflict Bosnia-Herzegovina and Kosovo, with out-group members being given approximately 23% less money than in-group members in a dictator game. Michellitch (2015) observes discrimination in taxi fares on the order of 20% for non-co-partisans during the 2008 election period in Ghana.
in-group and out-group dictator game recipients. But did this increased generosity increase or decrease discrimination, i.e. the difference in generosity towards members of each group?

Among respondents assigned to the UYVT program, assignment to a heterogeneous class does not lead to an additional significant increase in generosity in the Muslim, Christian, or full samples as shown in columns (4)–(6). But being assigned to a heterogeneous class does have a significant effect on discrimination against the out-group as shown in the Heterog. class × Play out-group coefficient estimates. In fact, we estimate that having been assigned to a heterogeneous class offsets nearly half of the discriminatory play by Muslims (roughly ₦2.5 out of ₦5.8) and offsets discriminatory play by Christians entirely (on the order of about ₦5). This is a striking result: sharing an educational experience with out-group members drastically reduces discriminatory behavior toward the out-group (and eliminates discrimination entirely for a key subgroup), compared to others who enjoyed the same educational experience with members of their own group only. This is particularly remarkable given that subjects do not appear to have (and apparently do not need to have) let go of their prejudices toward the out-group. Subjects’ prejudices towards the out-group may not have changed, but their treatment of its members improves as subjects get to know some of them.

The additional dosage of social contact achieved when in a heterogeneous pair within a heterogeneous class does not reduce discrimination relative to those in co-religious pairs, as shown in the insignificant interaction terms in columns (7)–(9). To some extent, this is to be expected given that assignment to a heterogeneous classroom already substantially reduced discrimination, so much so for Christians that it is not obvious how an additional dosage of social contact could feasibly reduce discrimination further. Instead, our analysis shows that such an assignment is associated with a further increase in generosity of approximately ₦4 (toward both out-group and in-group recipients), a finding driven by the Christian subsample, for whom the increase amounts to about ₦7.\footnote{Mironova and Whitt (2014) found that greater daily contact through residential integration with out-group members was associated with increased generosity toward the out-group relative to those living in more residentially segregated areas in the Balkans.} \footnote{In Online Appendix Tables A.38 and A.45 we present models interacting religious affiliation with the treatment in which none of the triple interaction terms are significant. Thus while our main results for these games differ across religious groups, this difference is not significant at conventional levels.}
But our design allows us to delve deeper. Is the social contact effect driven by a reduction in discrimination in participants assigned to heterogeneous classrooms relative to the control group, a worsening of discrimination in the homogeneous classrooms or a combination of both? Table 6 presents group-level means and differences in generosity towards members of the in-group and the out-group for those assigned to the control group, homogeneous classrooms and heterogeneous classrooms. Looking at the full sample, we observe that the difference in the average number of bills given in Column (3) is nearly identical in the control group (.19) and those assigned to heterogeneous classes (.17), yet this difference—which is discriminatory behavior—is quite substantially larger for those assigned to homogeneous classrooms (.54). This pattern is replicated within the Muslim and and Christian sub-samples.50

[Table 6 about here]

How should we interpret this finding? First, we observe that UYVT assignment to either class type increased generosity to both in-group and out-group members, as shown in Columns (1) and (2), but that among those assigned to homogeneous classrooms, this surplus is given disproportionately to in-group members. Second, we observe that homogeneous classroom treatment was not merely a neutral treatment offering the UYVT program absent social contact. Instead, homogeneous classrooms offered social contact with unfamiliar members of the in-group. If intergroup social contact is challenging for reasons such as communication difficulties, lack of shared norms, anxiety, and lack of empathy, within-group social contact should be considerably less challenging and should facilitate within-group bonding. The result may be greater in-group altruism, without any absolute decline in out-group generosity. Furthermore, homogeneous classrooms may have increased the salience of group identities, again increasing in-group altruism and widening the generosity gap. While we can only speculate, we suggest that intergroup social contact may reduce discrimination through a substitution effect. Social contact with members of the out-group reduces time spent with one’s in-group. We explore this result further in Section 6.

50Online Appendix Table A.31 replicates Table 6 across all possible treatment groups, and yields the same conclusion.
5.3 Destruction Game

The results from the destruction game shown in Table 7 follow a similar pattern. Assignment to the UYVT course leads to less destructive behavior toward both co-religious and non-co-religious recipients, as shown in the first three columns: across the entire sample, the program effect corresponds to a 10% reduction in destruction, off an average of every second bill getting destroyed. This effect is driven by the Christian subsample; the estimated effect is small and not statistically significant for Muslim respondents, as we can see in column (2). As in the dictator game above, we also do not see any program effect on discrimination in columns (1)–(3).

Assignment to a heterogeneous class, i.e. the social contact treatment, on the other hand, reduces discriminatory behavior, just as we observed in the dictator game. Again, this treatment offsets discriminatory play almost entirely, an effect driven by the Christian subsample, as shown in columns (4)–(6) in Table 7. The coefficients are relatively small in absolute terms and are therefore only weakly statistically significant (p < .1), but the estimated effects are large relative to the small estimated effect of being paired to play with an out-group member. That is, the destruction game elicited only limited discrimination against out-group members in the first place, but this discrimination was then offset by having been assigned to a heterogeneous class.

In this game, assignment to a heterogeneous pair within a heterogeneous classroom further reduces discrimination against out-group members and in fact induces discriminatory behavior towards the in-group. Columns (7)–(9) show that this effect is again driven by the Christian subsample, in which subjects assigned to a heterogeneous pair act less destructively toward Muslims than other Christians (in comparison to subjects assigned to a homogeneous pair).

51 Across the 10 plays of the destruction game, subjects have the opportunity to destroy an average of 1.5 bills per play, so an estimated intercept of about .7 implies that subjects destroy nearly half of the bills allocated to others.

52 Subjects played the dictator and destruction games with partners drawn from a sample that included both their UYVT classmates and people they did not know, which raises the question of whether increases in generosity and reductions in discrimination are primarily due to rounds played with their UYVT classmates. The well-established psychological phenomenon of “person positivity bias” suggests that familiar others are treated more favorably than strangers (Sears, 1983). However, when we restrict our dictator game analysis to rounds of play with strangers only, our conclusions remain largely the same. We observe a positive and
6 Discussion

Perhaps the most striking finding to emerge from our study concerns the effect of social contact in homogenous settings. What we term the “in-group bonding” effect was detected through the inclusion of a pure control group (whose subjects received only the survey portions of the study). As shown in Table 6, heterogeneous UYVT class members discriminate at nearly identical levels to pure control group members. In contrast, homogeneous UYVT class members display considerably higher levels of discrimination due to increases in giving to in-group members, a result consistent with Brewer (1999)’s widely-cited finding that prejudice is as much a function of in-group “love” as out-group “hate.”

Although UYVT participation led to generosity gains in dictator game play across treatment groups, contact with previously unfamiliar in-group members produces stronger positive effects than contact with strangers from the out-group. In this regard, intergroup social contact within heterogeneous classes might be more accurately viewed as a check against the potentially adverse effects of exclusively homogeneous social contact.

The ease of bonding within socially salient groups, like religious groups in northern Nigeria, should not be surprising. According to the contact hypothesis, intergroup contact works by creating friendships that bond individuals across cleavage lines. Social contact within groups should work in the same way but should be even more effective. In homogeneous settings, in-group members come to the first day of the contact intervention with a host of advantages, including shared norms of reciprocity, culture, and language—what Habyarimana et al. (2007) refer to as the beneficial “technology” of co-ethnicity. Our results offer new empirical evidence that this shared background serves as a powerful multiplier for the effect of social contact.

Our findings also resonate with the large body of work on systematic school desegrega-

significant program effect on generosity in the overall sample (\( p < .05 \)), and the reduction in discrimination among Christians assigned to heterogeneous classes remains highly significant (\( p < .01 \)). The magnitude of these estimates is attenuated, as we would expect. In the destruction game we again obtain a negative and significant program effect among Christians (\( p < .05 \)), and the reduced destruction effect previously observed among Christians assigned to heterogeneous class and partner treatments is of comparable size. Full results for these analyses are available in Online Appendix Tables A.33 and A.42.
tion in the United States in the past fifty years. Decades of work evaluating the effects of integrated schooling on levels of prejudice and the formation of inter-ethnic friendships has produced generally discouraging results (Stephan, 1978; Schofield and Eurich-Fulcer, 2001; Brown, 2011). Stephan (1978), for example, evaluated 18 studies of the effects of school desegregation on prejudice and found that in half of them, desegregation actually increased white students’ prejudice toward black students.53

The powerful “multiplier effect” of contact we observe within homogeneous UYVT classes also has important implications for the design and implementation of development projects and provision of social services in conflict-prone environments, where programs are often homogeneous by design (precisely to avoid conflict), or due to residential segregation across lines of conflict. Our study suggests that education-based and other development programming may have the unintended consequence of reinforcing preferential in-group treatment. Rather than viewing the inclusion of a socially heterogeneous treatment arm as an extra or “bonus” feature of development interventions, we suggest integrated programming is essential to curb the potentially negative effects of in-group bonding on intergroup relations.54

Another striking finding from our study is that intergroup contact appears to change behavior but not attitudes toward the out-group. While it is possible that this discrepancy is due to the difficulty of accurately measuring prejudice, we suggest several reasons why we believe our intervention genuinely produced behavioral changes in the absence of changes in attitudes.55 Though our study was not designed to assess these arguments empirically, we hope future research will shed light on this important question.

53Schofield and Eurich-Fulcer (2001) reached similar conclusions several decades later. A crucial challenge in real-world settings of school integration is students’ tendency to re-segregate when given the chance (for example, sitting in segregated groups in the school cafeteria or self-segregation along ethnic lines during other unstructured time during the school-day) brown2011.

54The finding arguably raises the question whether the wealth effect, discussed earlier and manifested in increased generosity to both in-group and out-group members in the dictator game, is desirable from the perspective of intergroup relations. The net effect on intergroup relations of a simultaneous increase in both generosity and discrimination towards members of the out-group is unclear, but we suspect it could be detrimental, given that economic inequality and the perceived group threat to which it contributes, can negatively impact intergroup relations (e.g. Quillian, 1995; Kunovich and Hodson, 2002; Humphreys and Weinstein, 2008; Østby, 2008).

55Another possibility is that attitudes towards the out-group as a whole are not affected by intense contact with one individual in a learning pair. We created learning pairs in this way to allow for sufficient depth of interaction to meet the conditions of Allport’s original hypothesis. While learning pairs were situated within classrooms, an intriguing possibility for future research design would be to rotate learning partners over time.
First, several foundational studies in social psychology, (e.g. Bem, 1970, 1972; Fazio, 1987) suggest that attitudes may be slow to change and behavioral changes may actually not only precede attitudinal changes but help to produce them. In these accounts, repeated new behaviors can, through a mechanism of “self-perception,” ultimately lead to changed attitudes and beliefs. Following this logic, it is possible that our findings in the UYVT study on attitudes are in fact intermediate results, i.e. prejudiced attitudes may change at some point in the future, but it is neither unreasonable nor extraordinary that behavior changed first. Particularly for highly sensitive issues, changing one’s behavior may be easier and cognitively less burdensome than articulating a changed belief.

Second, behavior is not simply a mapping of attitudes into actions, but reflects a combination of the effect of attitudes, strategic considerations, and responses to relevant norms of appropriate behavior. In this sense, there may exist more avenues through which an intervention can effect behavioral changes as opposed to changes in attitudes. For example, 16 weeks of participation in the UYVT course could lead to subjects being trained to adhere to norms about appropriate behavior with respect to generosity, fairness, and non-discrimination toward out-group members, as students in heterogeneous UYVT classes observed classmates treating out-group members in a fair and respectful manner. Note that teachers deliberately ensured fairness and non-discrimination in their dealings with students and in interactions among students, but deliberately avoided anti-prejudice programming.\footnote{For a related argument concerning norms, see e.g. Paluck (2009).}

Third, there is empirical precedent for the disjuncture between the effects of an intervention on behavior versus attitudes in intergroup reconciliation in a conflict-prone African setting. In widely cited and highly regarded studies, Paluck (2009) and Paluck and Green (2009a) find that sustained exposure to a radio program aimed at fostering reconciliation between Hutu and Tutsi listeners in Rwanda produced positive changes across multiple measures of behavior toward out-group members, including behavior with material stakes. In contrast, the studies found no evidence of corresponding changes in individual-level attitudes toward or beliefs about out-group members.\footnote{McConnell and Leibold (2001) also find that their primary behavioral measures were uncorrelated with their explicit prejudice measures, suggesting that these measures capture aspects of intergroup relations that do not necessarily move in tandem. Their findings are of particular relevance, because they used similar measures of explicit prejudice and privacy-protection.}
From a peace-building perspective, practitioners care most about how people choose to act and how their behaviors can help or hinder post-conflict peace-building, even if attitudes are undoubtedly of interest. Allport himself emphasized that “... as a rule discrimination has more immediate and serious social consequences than has prejudice” (Allport, 1954, 15). Similarly, in a review of the literature on stereotyping and prejudice, Fiske (2000) points to the paucity of studies of discriminatory behavior in social psychology and looks to behavioral studies as an important agenda for future research within the field, since “... thoughts and feelings do not exclude, oppress, and kill people; behavior does” (312).

7 Conclusion

Does social contact decrease intergroup prejudice and discrimination in urban conflict zones, such as those in Nigeria, Iraq, Israel, and other deeply divided societies? This important question urgently needs a policy response grounded in well-designed research. Our study is motivated by a desire to test the core claims of intergroup contact theory—that positive, egalitarian intergroup contact reduces prejudice and discrimination—in a challenging context, where discrimination, legacies of violent communal conflict, and extreme social segregation are routine parts of daily life.

We find that a grassroots-level intervention that induces contact between members of religious groups in conflict has little effect on intergroup prejudice but leads to increased generosity across treatments and a reduction in discriminatory behavior in heterogeneous classroom settings. These effects are achieved via two channels: a program effect and an intergroup social contact effect. Simply being offered a valuable and appealing program increases generosity to both out-group and in-group recipients. Among those assigned to any UYVT treatment, we observe a statistically significant increase in generosity to others across two types of behavioral games—dictator and destruction games—that subjects played as part of our post-treatment endline survey.

Conditional on assignment to the program, assignment to a heterogeneous class significantly reduces discrimination, i.e. generosity that privileges in-group over out-group members. Social contact with out-group members in the context of a positive, future-oriented
education experience helps close the gap in subjects’ treatment of in-group and out-group individuals. Importantly, this behavioral change does not appear to require a change in attitudes—subjects soften their treatment of the out-group even as they hold on to their prejudices. We also show that more in-depth contact due to having a non-co-religious learning partner can, under certain conditions, amplify these positive effects.

The characteristics of the Kaduna case allow us to generalize to other conflict and post-conflict environments, intergroup contexts in which neither group is clearly socially dominant, settings where residential segregation curtails opportunities for intergroup contact, and to interventions targeting disadvantaged youth. Yet prejudice and ethnic discrimination are major concerns in OECD and developing countries alike. Our findings have several key policy implications. First, program effects are a significant driver of increased generosity toward the out-group. Policy-makers seeking to improve intergroup relations should prioritize program content that is valuable and appealing to draw in participants from disadvantaged backgrounds who might not self-select into peace education programs. Combining educational content and economic development programs with intergroup contact allows donors and governments to address multiple, intertwined challenges at once. Simply offering educational and economic empowerment opportunities not otherwise available to disadvantaged youth may induce goodwill toward out-group members and society at large.

Second, setting goals of behavioral change, rather than prejudice reduction, may be both more realistic and more useful in the long-term. While attitude change is most feasible among adolescents and young adults (Krosnick and Alwin, 1989), we should perhaps not be surprised that prejudice is resistant to change in an environment that has repeatedly experienced violent conflict. Prejudices are formed and reinforced over a lifetime of experience (e.g. Bar-Tal, 1997; Bigler and Liben, 2007). When those experiences include a recent history of violent conflict, prejudice may be particularly difficult to dislodge (Paluck, 2009). Individuals may interpret new social contact experiences in light of pre-established views of the out-group rather than using these new interactions to update their beliefs. In addition, after each session of intergroup contact, participants often return to highly segregated daily lives, with routine exposure to norms-sanctioned prejudice toward the out-group (McCauley, 2002). Increasing out-group generosity and reducing discriminatory behavior in
everyday interactions is crucial in a context of open intergroup hostility and violence and appears achievable regardless of internal prejudices.

Finally, donors and governments should be cautious about perpetuating the cycle of ingroup bonding, particularly in socially segregated contexts. In our experiment, students assigned to homogeneous computer training classrooms exhibited higher levels of discriminatory behavior than members of the control group. Our findings suggest that education and other social services should be provided in integrated and cooperative settings to facilitate contact across cleavage lines.

To advance research on the role of intergroup social contact, we look forward to analyses of the links between prejudice and discriminatory behavior, including the role of changing norms of appropriate behavior in heterogeneous settings. Studies of cooperative intergroup contact in situations characterized by relations of dominance and subordination would also enhance our understanding of the benefits and limits of intergroup contact interventions in conflict mitigation.
References


Lewis, Peter. 2007. *Identity, institutions and democracy in Nigeria*. IDSA.


Figure 1: UYVT Experimental Treatment Arms
<table>
<thead>
<tr>
<th>Negative Attributes</th>
<th>Positive Attributes</th>
<th>Out-group Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrogant</td>
<td>Friendly</td>
<td>Lazy - Hardworking</td>
</tr>
<tr>
<td>Unreasonable</td>
<td>Honest in business dealings</td>
<td>Ignorant - Worldly</td>
</tr>
<tr>
<td>Ungrateful</td>
<td>Responsible</td>
<td>Ungenerous - Charitable</td>
</tr>
<tr>
<td>Fanatical</td>
<td>Good citizens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Peaceful</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dependable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intelligent in school</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2: UYVT Student Attendance
Figure 3: UYVT attendance over time
<table>
<thead>
<tr>
<th>Sample</th>
<th>All</th>
<th>Muslims</th>
<th>Christians</th>
<th>All in UYVT</th>
<th>Muslims in UYVT</th>
<th>Christians in UYVT</th>
<th>All in Heterog. class</th>
<th>Muslims in Heterog. class</th>
<th>Christians in Heterog. class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>716</td>
<td>343</td>
<td>373</td>
<td>474</td>
<td>221</td>
<td>253</td>
<td>277</td>
<td>135</td>
<td>142</td>
</tr>
<tr>
<td>Treatment</td>
<td>480</td>
<td>222</td>
<td>258</td>
<td>322</td>
<td>152</td>
<td>170</td>
<td>122</td>
<td>59</td>
<td>63</td>
</tr>
<tr>
<td>Control</td>
<td>236</td>
<td>121</td>
<td>115</td>
<td>152</td>
<td>69</td>
<td>83</td>
<td>155</td>
<td>76</td>
<td>79</td>
</tr>
</tbody>
</table>

All specifications are OLS regressions in which the treatment indicator variables represent assignment to the UYVT course (UYVT) vs. no course assignment, a heterogeneous classroom (Heterog. class) vs. a homogeneous classroom, or a non-co-religious course partner (Heterog. pair) vs. a co-religious partner within heterogeneous classrooms, respectively. Robust standard errors in parentheses. ** p < 0.01, * p < 0.05, + p < 0.10
Table 3: Prejudice Index, Positive Attributes (scale ranges from 1 to 5, larger values indicate more positive assessment)

<table>
<thead>
<tr>
<th></th>
<th>Program effect</th>
<th>Contact effect</th>
<th>Contact dosage effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>UYVT</td>
<td>-0.11⁺</td>
<td>0.02</td>
<td>-0.19⁺</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Heterog. class</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Heterog. pair</td>
<td></td>
<td>-0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.10)</td>
<td>(0.13)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.14)</td>
</tr>
<tr>
<td>Constant</td>
<td>4.00**</td>
<td>4.21**</td>
<td>3.75**</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.07)</td>
</tr>
<tr>
<td></td>
<td>3.87**</td>
<td>4.21**</td>
<td>3.53**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.09)</td>
<td>(0.09)</td>
</tr>
<tr>
<td></td>
<td>3.96**</td>
<td>4.25**</td>
<td>3.65**</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample</th>
<th>All Muslims Christians All in UYVT Muslims in UYVT Christians in UYVT All in Heterog. class Muslims in Heterog. class Christians in Heterog. class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>780 396 384 509 250 259 301 153 148</td>
</tr>
<tr>
<td>Treatment</td>
<td>515 251 264 346 170 176 134 67 67</td>
</tr>
<tr>
<td>Control</td>
<td>265 145 120 163 80 83 167 86 81</td>
</tr>
</tbody>
</table>

All specifications are OLS regressions in which the treatment indicator variables represent assignment to the UYVT course (UYVT) vs. no course assignment, a heterogeneous classroom (Heterog. class) vs. a homogeneous classroom, or a non-co-religious course partner (Heterog. pair) vs. a co-religious partner within heterogeneous classrooms, respectively. Robust standard errors in parentheses. ** p < 0.01, * p < 0.05, + p < 0.10
Table 4: Prejudice Index, Out-group Evaluation (scale ranges from 1 to 5, larger values indicate more positive assessment)

<table>
<thead>
<tr>
<th>Program effect</th>
<th>Contact effect</th>
<th>Contact dosage effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)  (2)  (3)</td>
<td>(4)  (5)  (6)  (7)</td>
</tr>
<tr>
<td>UYVT</td>
<td>-0.11+</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Heterog. class</td>
<td></td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td>Heterog. pair</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.38**</td>
<td>4.68**</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Sample | All | Muslims | Christians | All in UYVT | Muslims in UYVT | Christians in UYVT | All in Heterog. class | Muslims in Heterog. class | Christians in Heterog. class |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>762</td>
<td>391</td>
<td>371</td>
<td>496</td>
<td>248</td>
<td>248</td>
<td>294</td>
<td>153</td>
<td>141</td>
</tr>
<tr>
<td>Treatment</td>
<td>501</td>
<td>249</td>
<td>252</td>
<td>338</td>
<td>170</td>
<td>168</td>
<td>132</td>
<td>67</td>
<td>65</td>
</tr>
<tr>
<td>Control</td>
<td>261</td>
<td>142</td>
<td>119</td>
<td>158</td>
<td>78</td>
<td>80</td>
<td>162</td>
<td>86</td>
<td>76</td>
</tr>
</tbody>
</table>

All specifications are OLS regressions in which the treatment indicator variables represent assignment to the UYVT course (UYVT) vs. no course assignment, a heterogeneous classroom (Heterog. class) vs. a homogeneous classroom, or a non-co-religious course partner (Heterog. pair) vs. a co-religious partner within heterogeneous classrooms, respectively. Robust standard errors in parentheses. ** p < 0.01, * p < 0.05, + p < 0.10
Table 5: Number of Bills Given in Dictator Game

<table>
<thead>
<tr>
<th></th>
<th>Program effect</th>
<th>Contact effect</th>
<th>Contact dosage effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3)</td>
<td>(4) (5) (6)</td>
<td>(7) (8) (9)</td>
</tr>
<tr>
<td>UYVT</td>
<td>0.47** 0.48** 0.46*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.13) (0.18) (0.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UYVT × Play out-group</td>
<td>-0.09 -0.07 -0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07) (0.11) (0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterog. class</td>
<td>-0.17 -0.11 -0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.16) (0.23) (0.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterog. class × Play out-group</td>
<td>0.39** 0.25+ 0.52**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10) (0.15) (0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterog. pair</td>
<td>0.39+ 0.12 0.67*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.22) (0.29) (0.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterog. pair × Play out-group</td>
<td>-0.01 -0.03 -0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12) (0.17) (0.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play out-group</td>
<td>-0.19** -0.33** -0.02</td>
<td>-0.55** -0.58** -0.51**</td>
<td>-0.18* -0.32* -0.01</td>
</tr>
<tr>
<td></td>
<td>(0.06) (0.08) (0.07)</td>
<td>(0.08) (0.13) (0.09)</td>
<td>(0.08) (0.12) (0.11)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.57** 2.59** 2.55**</td>
<td>3.16** 3.15** 3.17**</td>
<td>2.90** 3.04** 2.74**</td>
</tr>
<tr>
<td></td>
<td>(0.10) (0.14) (0.15)</td>
<td>(0.13) (0.19) (0.18)</td>
<td>(0.14) (0.20) (0.19)</td>
</tr>
</tbody>
</table>

All specifications are OLS regressions in which the treatment indicator variables represent assignment to the UYVT course (UYVT) vs. no course assignment, a heterogeneous classroom (Heterog. class) vs. a homogeneous classroom, or a non-co-religious course partner (Heterog. pair) vs. a co-religious partner within heterogeneous classrooms, respectively. Round-of-play fixed effects included in all specifications. Play out-group indicates rounds of play in which the survey respondent was from a different religion than the recipient. Robust standard errors (in parentheses) clustered by respondent. ** p < 0.01 * p < 0.05, + p < 0.10
Table 6: Mean Number of Bills Given in Dictator Game, by Treatment Assignment

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td></td>
<td>In-group</td>
<td>Out-group</td>
<td>Diff</td>
<td>In-group</td>
<td>Out-group</td>
<td>Diff</td>
<td>In-group</td>
</tr>
<tr>
<td>Control</td>
<td>2.57</td>
<td>2.38</td>
<td>0.19</td>
<td>2.59</td>
<td>2.25</td>
<td>0.33</td>
<td>2.56</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.11)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Homog. Class</td>
<td>3.16</td>
<td>2.62</td>
<td>0.54</td>
<td>3.15</td>
<td>2.57</td>
<td>0.58</td>
<td>3.16</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.09)</td>
<td>(0.14)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Heterog. Class</td>
<td>2.99</td>
<td>2.83</td>
<td>0.17</td>
<td>3.04</td>
<td>2.71</td>
<td>0.34</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>
Table 7: Number of Bills Destroyed in Destruction Game

<table>
<thead>
<tr>
<th></th>
<th>Program effect</th>
<th>Contact effect</th>
<th>Contact dosage effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>UYVT</td>
<td>-0.07*</td>
<td>-0.02</td>
<td>-0.12*</td>
</tr>
<tr>
<td>UYVT × Play out-group</td>
<td>0.02</td>
<td>-0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>Heterog. class</td>
<td>0.04</td>
<td>0.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Heterog. class × Play out-group</td>
<td>-0.05+</td>
<td>-0.02</td>
<td>-0.08+</td>
</tr>
<tr>
<td>Heterog. pair</td>
<td>0.01</td>
<td>-0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Heterog. pair × Play out-group</td>
<td>-0.07+</td>
<td>-0.02</td>
<td>-0.10*</td>
</tr>
<tr>
<td>Play out-group</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.70**</td>
<td>0.65**</td>
<td>0.75**</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
</tbody>
</table>

Sample: All Muslims Christians All Muslims Christians All in Muslims Christians in in Heterog. class Heterog. class Heterog. class

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Muslims</th>
<th>Christians</th>
<th>All in UYVT</th>
<th>Muslims in UYVT</th>
<th>Christians in UYVT</th>
<th>All in Heterog. class</th>
<th>Muslims in Heterog. class</th>
<th>Christians in Heterog. class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>7920</td>
<td>3980</td>
<td>3940</td>
<td>5150</td>
<td>2520</td>
<td>2630</td>
<td>3040</td>
<td>1540</td>
<td>1500</td>
</tr>
<tr>
<td>Treatment</td>
<td>5220</td>
<td>2530</td>
<td>2690</td>
<td>3480</td>
<td>1710</td>
<td>1770</td>
<td>1350</td>
<td>680</td>
<td>670</td>
</tr>
<tr>
<td>Control</td>
<td>2700</td>
<td>1450</td>
<td>1250</td>
<td>1670</td>
<td>810</td>
<td>860</td>
<td>1690</td>
<td>860</td>
<td>830</td>
</tr>
</tbody>
</table>

All specifications are OLS regressions in which the treatment indicator variables represent assignment to the UYVT course (UYVT) vs. no course assignment, a heterogeneous classroom (Heterog. class) vs. a homogeneous classroom, or a non-co-religious course partner (Heterog. pair) vs. a co-religious partner within heterogeneous classrooms, respectively. Round-of-play fixed effects included in all specifications. Play out-group indicates rounds of play in which the survey respondent was from a different religion than the recipient. Robust standard errors (in parentheses) clustered by respondent. ** p < 0.01, * p < 0.05, + p < 0.10