Imagined vs. Actual "Others"
An Experiment on Interethnic Giving in Afghanistan[*]

Luke N. Condra[·]
Mohammad Isaqzadeh[†]
Sera Linardi[‡]

This version: July 27, 2016

Abstract

Does willingness to aid "others" change when in their physical presence? We argue that studies cueing non-coethnics through names and photos may underestimate discrimination resulting from actual interethnic interaction. In an experiment in Kabul, Afghanistan, Dari-speaking day-laborers contribute their earnings to a hospital under one of three randomly-assigned experimental conditions. In *In-group*, the hospital is in a Dari-speaking province; in *Out-group-Abstract* and *Out-group-Real*, it is in a Pashto-speaking (Pashtun) province. While subjects in *In-group* and *Out-group-Abstract* wait for the experiment with only Dari-speakers present, subjects in *Out-group-Real* wait among both Dari-speakers and Pashto-speakers. When Pashtuns are absent, the findings accord with other experiments that find little to no out-group discrimination. However, the physical presence of Pashtuns (*Out-group-Real)* decreases contributions by 25%. Consistent with the threat hypothesis, contributions decrease the longer Dari-speakers wait with Pashtuns, though subjects' youth and ability to speak Pashto mediate this effect.

[*] For their helpful suggestions and feedback we thank Claire Adida, James Fearon, Uri Gneezy, Malte Lierl, Amanda Robinson, Rick Wilson, and participants at the Midwest Political Science Association Annual Meeting and the Behavioral Models of Politics Conference. We thank Katherine Yoon for her valuable research assistance. We gratefully acknowledge the Hewlett International Grant Program, the Global Studies Center, and the Central Research Development Fund, all of the University of Pittsburgh, for financial support. This research was approved by the University of Pittsburgh's Institutional Review Board (PRO14050489).

[·] Graduate School of Public and International Affairs, University of Pittsburgh, lcondra@pitt.edu.
[†] Department of Political Science and Public Administration, American University of Afghanistan, mohammad.isaqzadeh@gmail.com.
[‡] Graduate School of Public and International Affairs, University of Pittsburgh, linardi@pitt.edu.

Is there taste-based discrimination against ethnic-out groups? The empirical literature examining this question has returned mixed results. On the one hand, experiments that cue the out-group in lab-style economic games in an 'abstract' way to subjects through photographs (e.g., Habyarimana et al. 2007), ethnically identifiable names (e.g., Fershtman and Gneezy 2001), or outright disclosure of ethnic identity (e.g., Chuah et al. 2014) find little to no evidence of discrimination toward the out-group. But experiments that physically expose subjects to the out-group find much stronger evidence of discriminatory behavior or prejudicial attitudes resulting from interaction (e.g., Adida, Laitin, and Valfort forthcoming, Enos 2014). What accounts for this discrepancy? Are there important triggers of ethnic identity that are present in the flesh but not in the abstract?

To address this question, we conducted a field experiment in which we raise contributions for a real public good (a hospital), among day laborers in Kabul, Afghanistan. In the experiment we employ an abstract cue of ethnic identity separately from a real cue of the out-group via physical exposure to compare the effects of each. We take advantage of the relatively clean ethno-linguistic cleavage in society where the Pashtuns, who remained politically dominant after the bloody interethnic Afghan civil war of the 1990s,[1] speak Pashto as their first language while most non-Pashtun minorities speak Dari. We focus on whether Dari-speakers give (anonymously) less to a hospital in the Pashto-speaking Helmand province (*Out-group* treatment) than they do to the branch in the Dari-speaking Panjshir province (*In-group* treatment). We manipulated the salience of ethnic cues in the *Out-group* treatment by having the Dari-speaker wait for the experiment amongst only other Dari speakers (*Abstract*) or also with some Pashto speakers present (*Real*).

[1] Between 1992 and 2000, scholars estimate at least 27,500 battle deaths in this civil war (Lacina and Gleditsch 2005).

The nature of the physical interaction is not determined by the experimenters, but left entirely to the discretion and choice of the subjects, as it would be in real life.

We find that subjects do *not* discriminate against out-group Pashtuns when out-group members are not present (*Out-group-Abstract*). But discrimination is significant when Pashtuns are physically present (*Out-group-Real*). The decrease in contributions is stark – about 25% – and shows that an absence of discrimination against the out-group in the abstract does not equate to no discrimination in the flesh.

To explore possible mechanisms driving this behavior, we make use of the fact that the length of time that subjects were physically exposed to the Pashtun out-group before the contribution decision is exogenously determined because subjects were called in random order to be enumerated. We find that contributions decrease the longer subjects waited alongside the Pashto-speakers. This suggests that due to Pashtuns' historical dominance in Afghanistan, the perception of threat, or the salience of ethnic identity, increased for the Dari-speakers as the time of exposure to spoken Pashto increased. If that is the case, then an observable implication is that those subjects whose personal characteristics would increase the perceived threat from hearing Pashto – namely those who are not bilingual in Pashto and those who were born before the carnage of the 1990s civil war – should be more negatively affected by spoken Pashto. Indeed, we find this to be true.

Our contributions are several. First, our study shows that finding no discrimination in response to an abstract cue of the out-group does not mean there will not be discrimination when the out-group is physically present. Our experimental design enables us to establish this finding in an environment where the existence and length of physical exposure is randomly assigned, avoiding the problem in observational studies where subjects may have selected to live or remain in places where they are physically exposed to the out-group.

Second, several features of our experiment contribute to the external validity of our results. While lab experiments have contributed greatly to our theories of intergroup contact, it can be difficult to generalize their findings to complex and volatile environments such as the one we study here. Subjects are drawn from the same population – low income Afghans – whose behavior we care about understanding. We embed the experiment in the context of a real and relevant charitable organization, so far from playing a 'game', subjects are confronted with an opportunity to contribute to a public good.[2] Moreover, the treatment is naturalistic. Subjects are left to wait with others, as they would while waiting to vote or to receive services. This emulates situations where strangers mix together for a random period of time, and in which power dynamics that are missing in lab interactions emerge and may affect behavior.

Third, our results contribute to an emerging and interesting set of results on how duration of exposure to the out-group affects attitudes and behavior toward the out-group in everyday life. We show that increasing duration of physical exposure to the out-group – which ranges from a few minutes to about five hours – negatively affects altruism toward the out-group. Very few studies have examined this question,[3] but putting our results together with the few that have suggests an

[2] Charitable giving is a normal and familiar activity in Islam (Lambarraa and Riener 2012).

[3] For example, Enos (2014) finds a statistically significant increase in discriminatory attitudes towards Mexicans among whites after three consecutive days of brief physical exposure (<30 minutes) to Mexicans; however, this increase is no longer significant after ten consecutive days of exposure. Alexander and Christia (2011) finds that students who attended an ethnically segregated school gave less in a public goods game if the players are multi-ethnic. This is not true of students from ethnically mixed schools.

inverse-U-shaped relationship: discrimination initially increases with physical exposure to the out-group but then decreases when exposure is repeated over the long term (e.g., weeks).

Fourth, we speculate on the nature of factors that might mediate these negative effects of out-group exposure through analysis of heterogeneous treatment effects in our sample. Our reasoning here is *post hoc* and thus should be considered exploratory, but makes an important point that should be tested in future research. We find that physical exposure to the out-group does not cause discrimination among young and bilingual subjects. We reason that language acts as a vehicle for integration or a sign that someone already has integrated with the out-group, enabling higher quality interaction. Similarly, relative youth may represent little to no memory of the civil war of the 1990s that was fought largely along ethnic lines. While such factors are context-dependent, these findings suggest that the process of integration will differ depending on personal characteristics that enable meaningful interethnic communication and decrease the impact of historical antagonisms between groups.

Theoretical Motivation

Many observational studies on the effect of ethnic diversity on attitudes and behavior toward the out-group find evidence in favor of the contact hypothesis,[4] but these results may suffer from selection problems, whereby prejudiced people might avoid intergroup contact, biasing results (Hopkins 2010, 49, Pettigrew 1998). As such, scholars have turned to lab and field experiments to

[4] The evidence is particularly strong when certain conditions from Allport (1954) are present: equal status between the groups in the situation; common goals exit for groups; the situation encourages intergroup cooperation; and authorities, law, or custom support the interaction (Pettigrew and Tropp 2006).

isolate the independent effect of ethnic mixing on prosocial behavior (usually in standard economic games), exposing subjects to the out-group. We review this literature in terms of the kind of cue provided, in the 'abstract' and through 'real' physical exposure.

Subjects have been exposed to the out-group in the 'abstract' in three main ways. One is through the use of photographs that reveal others' race to subjects (e.g., Burns 2004, Fong and Luttmer 2009, Habyarimana et al. 2007). Though the geographic settings of these experiments differ widely, the consistent finding is that partner's race does not affect the size of offers in these dictator-style games, or even attitudes toward the out-group (Havekes, Coenders, and Van der Lippe 2013). Similarly, experiments that have used names commonly associated with a particular ethnic or religious group to 'abstractly' cue trust/dictator game partner's ethnic identity have found no significant differences in average offers made to the out-group (Bouckaert and Dhaene 2004, Chuah, Fahoum, and Hoffmann 2013, Fershtman and Gneezy 2001). This is also the case when ethnic identity is cued by simply stating or revealing a game partner's ethnicity outright (e.g., Chuah et al. 2014, Waring and Bell 2013).

In contrast, studies with real (rather than abstract) exposure to the out-group suggest that there indeed is discrimination toward the out-group. For example, Enos (2014) finds an increase in anti-Mexican attitudes in surveys when white commuters waited alongside Mexicans for trains, while Alexander and Christia (2011) find smaller contributions in public goods games with abstract out-group ethnic cues among students who attended an ethnically segregated school.

However, various design features in these studies make it difficult to compare the effect of abstract vs. real exposure on behavior. For example, some studies feature physical exposure to the out-group, but not an abstract cue (e.g., Enos 2014), while others feature an abstract cue but not a real cue of ethnic identity within the experiment itself (e.g., Alexander and Christia 2011). Others

investigate how physical exposure affects attitudes, but not behavior (Enos 2014). Still others focus on third party sanctioning as a mechanism for inducing prosocial behavior across ethnic lines (e.g., Bernhard, Fischbacher, and Fehr 2006, Jeon, Johnson, and Robinson forthcoming), an institutional addition that we do not pursue here due to our focus on pure taste-based discrimination.

Even for studies that are most similar to ours in their design, abstract and real cues are both included, making it difficult to separate their effects on contributions to the public good. For example, though coethnics trust each other more than non-coethnics while playing trust games together in Malawian market areas (real cue), the partner's ethnic identity is communicated by revealing her ethnically-homogenous home village (abstract cue), so the effects of the two cue types are not investigated separately (Robinson forthcoming). Similarly, increasing the number of Senegalese Muslims in a dictator game session decreases French donors' offers to Senegalese Muslim partners (real cue), but session diversity is communicated through the number of Muslim names revealed on a screen (abstract cue) (Adida, Laitin, and Valfort forthcoming). In contrast, we want to isolate the effects of abstract and real cueing of the out-group in the same experiment in order to test whether there is a difference in how they affect behavior.

These literatures on the effects of abstract and real cues of the out-group on prosocial behavior have developed largely in parallel to each other. Our synthesis of these two literatures suggests that behavior toward the out-group is affected differently by abstract and real cues of the out-group. To investigate this question, we must first include both types of exposure to the out-group in the same experiment but also be able to estimate their effects separately. Second, we should test an argument in the literature, whether physical exposure increases out-group discrimination by

randomly varying the length of physical exposure to the out-group,[5] providing a continuous measure of physical exposure as a proxy for the salience of ethnic identity. This should provide an exogenous increase in the salience of ethnic identity among those who had physical exposure to the out-group. We should expect more discrimination from those for whom ethnic identity was made more salient. Third, to tie our study to the existing literature that focuses on prosocial behavior exhibited through economic games, our experiment examines the relevant and familiar activity of charitable giving, but does so in the context of contributing to a real hospital that is well-known to our subjects. We embed the experiment in Afghanistan, a weak state that struggles to provide public goods, since we care most about understanding how private citizens respond to solicitations to contribute to public goods provision in similar contexts.

Experimental Design

Our experiment takes place in Afghanistan, a country where ethnic identity has long featured prominently in politics. Among ethnic groups defined by descent-based attributes (Chandra 2006), Pashtuns are the largest, making up some 40% of the population (Barfield 2010). Though they are split into different tribes, Pashtuns "have a powerful feeling of ethnic identity, based on a tradition of common descent; a characteristic Indo-Iranian language, Pashto; and a communal code known as Pashtunwali" (Riphenburg 2005, 37). Tajiks are the second largest ethnic group in the country and speak Dari, instead of Pashto. Other non-Pashtun minority ethnic groups, most of whom

[5] Specifically, "the more we are brought into physical proximity with people of another race or ethnic background, the more we stick 'to our own' and the less we trust the 'other'" (Putnam 2007, 142).

speak Dari, include Hazaras, Uzbeks, and Nuristani, among others, each of which has a strong ethnic identity.[6]

The modern Afghan state was founded by a Pashtun, Abdur Rahman Khan, in the 1880s through repression and subjugation of non-Pashtun minorities. Indeed, according to one scholar, the "sociology of Pashtun dominance over the other ethnic communities in the country forms the very substance of political developments and state building in Afghanistan" (Shahrani 1986, 25; quoted in Simonsen 2004, 710). With the exception of a brief period in 1929, Pashtuns controlled the Afghan state and dominated Afghan politics until the civil war in 1992 (Barfield 2010), and state "policies favoring Pashtuns...caused resentment among other ethnic groups" (Simonsen 2004, 709). The relationship between Pashtuns and non-Pashtuns turned even more hostile during the civil war as the minorities challenged the Pashtuns' claim to power and ethnic-based militias struggled to capture the state. Over three years, at least 25,000 civilians were killed (Barfield 2010, Lacina and Gleditsch 2005). Almost all factions engaged in arbitrary killing, torture, and sexual violence against civilian members of other ethnic groups (Oxfam 2009), and although the Pashtun-dominated Taliban promised to bring peace to the country, they used brute force against non-Pashtuns, massacring Hazaras and expelling Tajiks from their villages in the northern plains of Kabul (Barfield 2010, HRW 1998, Nordland 2012, Oxfam 2009).

[6] Our focus on ethnicity is not to say that other bases of social identity in Afghanistan do not matter, but Afghan society has long been divided mainly along language: the Pashto-speaking dominant group vs. Dari-speaking ethnic minorities. Furthermore, recent experimental evidence suggests that the emphasis that journalists and scholars tend to place on the political importance of non-ethnic identities – *qawm* as 'solidarity group' being most prominent among them – may be empirically unfounded (Jochem, Murtazashvili, and Murtazashvili 2015).

The post-Taliban political system has tried to promote national unity and bridge ethnic divisions. The 2004 Constitution recognizes all ethnic identities, renders all ethnic groups equal before the law, and even obliges the state to promote material equality through balanced development of all areas of the country. Formal recruitment quotas for ethnic groups were considered in the formation of the Afghanistan National Army and Afghanistan National Police (Giustozzi and Isaqzadeh 2012).

There is some evidence that not only ethnic relations have improved since 2001, they may be less politically divisive. At the time of our experiment, the country was in the midst of a presidential election, deciding between an ethnic Pashtun (Ghani) and a Tajik-Pashtun mixed candidate (Abdullah), who most considered Tajik. But issues that dominated the political landscape in the lead up to the 2014 election – such as persistent violence and corruption – crossed ethnic lines. A pre-election survey of over 2,000 people across multiple provinces revealed peace and security to be the issue most respondents wanted parliament to prioritize (Callen and Long 2015). On the eve of the election, one young man remarked, "[h]istorically, the Pashtun people have ruled this country. But in spite of this, I don't care who wins the election. Just put an end to the violence" (Craig 2014). Another poll taken just prior to the election found that an astounding 83% of respondents would find it easy or somewhat easy to accept a non-coethnic as president (ACSOR 2014).

Our experiment takes place in Kabul, so it is worth noting that these changing attitudes may be particularly apparent in the capital. It is an ethnically diverse city, due to increasing urban migration, and in this way is different from the more ethnically homogeneous rural areas of the country. Kabul's neighborhoods tend to be ethnically segregated, though not homogeneous. In terms of ethnic mixing in the workplace, most residents of these areas intermingle with other

ethnic groups since they are very likely to work outside these neighborhoods, particularly in the city center, where you can find people representing all ethnic groups.

Within this context, we conducted our field experiment in Kabul in June 2014. Given the historical rivalry between Pashtuns and non-Pashtun ethnic minorities, our primary interest is the willingness of non-Pashtun minorities to aid Pashtuns. Since the ethno-linguistic divide is quite stark, this experiment will focus on comparing the contribution of Dari-speakers to public goods that would mainly benefit other Dari-speakers, with contributions to public goods that would benefit Pashto-speakers.[7]

Our partner organization for the experiment is EMERGENCY Hospital, a well-known international nonprofit that provides medical care to Afghans (Mogelson 2012). Two of Emergency's main hospitals in Afghanistan are located in two different provinces, each dominated by a particular ethnic group. The first is located in Anabah, in the Panjshir Valley of northern Afghanistan, a region heavily populated by Dari-speaking minority ethnic groups. The second hospital is located in Lashkar-gah, Helmand, in southern Afghanistan, a region dominated by

[7] In our subject pool, 89% of Pashtuns stated Pashto as their first language while 96% of non-Pashtuns stated Dari as their first language. 40% of Pashtuns speak Dari as a second language and about 37% of non-Pashtuns speak Pashto. Despite this bilingualism, the ethno-linguistic divide is quite clean: only 2 out of 46 subjects who stated their ethnicity as Pashtuns requested to be enumerated in Dari and 4 out of our 213 non Pashtun subjects requested to be enumerated in Pashto. This indicates subjects are more comfortable speaking the language associated with their ethnicity.

Pashtuns.[8] By alternating the beneficiary of the contribution between these two hospitals that are run by the same nonprofit, we vary the ethnic identity of those who would benefit from the public good while holding constant the attributes of the public good itself (e.g., type of services, reputation, etc.).

Each morning over several days, we recruited male day-laborers waiting to be hired for the day in market areas around the city.[9] This is important as subjects are drawn from the very population about which we wish to make inferences: typical Afghan men in that they earn a low income, are sporadically employed in a low-skill occupation, and mirror the ethnic diversity of the country. However, while we note below how the demographic characteristics of our sample (e.g., income and education) are in line with available data from national surveys, our results are still based on a convenience sample, as we have no way to ensure that our subjects are representative of the broader Afghan population.

---

[8] These regions are highly ethnically homogeneous (http://www.icr.ethz.ch/data/geoepr). Each ethnic group's demographic dominance in the two provinces is key because it allows us to be confident in assuming that our subjects infer that contributions to the Helmand hospital will generally benefit Pashtuns and contributions to the Panjshir hospital will generally benefit non-Pashtuns.

[9] We recruited men only for two reasons. First, due to cultural norms surrounding male-female interaction, we would have had to conduct a separate experiment for women only, and resources and time did not permit this. Second, and more substantively, we are more interested in male behavior with regard to out-group discrimination, as others have found discrimination effects to be driven almost entirely by men (Fershtman and Gneezy 2001).

We chose different market locations each morning to avoid duplicate recruitment and to maximize the geographic breadth of our sample. Due to the security situation in the country, it was imperative that the experiment was completed in as few days as possible. Subjects in the same treatment were therefore all recruited on the same day. This strategy may have affected balance of demographics across treatment conditions, which we will control for in our regressions. While the ideal would have been to take the day's recruits and place them randomly into one of our three different treatment conditions upon arrival at our enumeration location, we could not do so for logistical reasons. Most important, we would have needed two separate and identical waiting areas for our subjects at the enumeration location, an ethnically mixed one (for Dari-speakers in *Out-group-Real*) and an ethnically homogeneous one (for Dari-speakers in *In-group* and *Out-group-Abstract*), so as to retain the integrity of the treatments, and we could not do this at our location.

Our personnel approached potential subjects in the market area and asked if they would participate in a survey about the labor force in Kabul. In exchange, they would be paid 500 AFN (about $8.60 at the time, a little over a day's wage). Subjects who provided oral consent were transported to the enumeration location where they waited in the courtyard to be enumerated. Enumeration of the first group of subjects began around 7 AM and the last group of subjects had finished and exited the location by about noon, a span of about 5 hours. As illiteracy is high in Afghanistan, we paired each subject with an enumerator to orally administer the survey and record subjects' answers. Subjects were called in from the courtyard in groups of 5-7 to be enumerated.

We did not force or direct interaction among waiting subjects in any of the treatment conditions (and hence in spite of physical proximity, social interaction could be positive, negative, both, or non-existent). Our goal was to allow participants in our study to interact as they saw fit (Enos 2014), thus approximating the way ethnic mixing occurs in everyday life in Kabul. Crucially,

subjects were aware of the ethnic identity of those around them both before and during the interview through two features of the experiment. We estimate that the walled courtyard in which subjects waited to be enumerated is roughly 1000 square meters, small enough to enable subjects to hear the language spoken by others waiting around them, but large enough that subjects had choice in who to sit next to and interact with while they waited. Also, as subjects were called in from the courtyard to be enumerated, interviewers confirmed the language used for enumeration with his subject pairing out loud, so subjects are aware of the linguistic mix in the groups.

To maintain privacy, each subject-enumerator pair sat behind a privacy partition. Once all subjects in the session had finished the survey, each received payment in an envelope (six 50 AFN bills, five 20 AFN bills, and ten 10 AFN bills). Then, a solicitor told subjects that they had the (unexpected) opportunity to contribute to a nonprofit hospital, Emergency, by reading the text below:[10]

Thank you for completing this survey. To thank you for your participation, we would like to compensate you with 500 AFN for your time. On the table in front of you there is an envelope with 500 AFN inside. Today you also have the opportunity to donate part of your compensation to EMERGENCY in Helmand [Panjshir] Province, in the south [north]. As you know, EMERGENCY is committed to caring for the victims of war. The organization provides treatment to our Pashtun [Tajik] brothers and sisters who are victims of war in the south [north]. We invite you to contribute to the treatment of victims of war in the region. If you would like to donate to EMERGENCY, please leave the money you wish to donate in the envelope. Remove the money you wish to keep for yourself from the envelope. If you decide not to make any donations, you can leave the envelope empty. Before you leave the room, please drop the envelope in the box by the door.

[10] This paper is part of a broader experiment, reported in Condra, Isaqzadeh, and Linardi (2016). Here we restrict analysis to sessions in which the solicitor was dressed as a civilian. The solicitor is Pashtun but is physically of indeterminate ethnicity and completely fluent in Dari to the extent that even the enumerators were not able to guess his ethnic background.

The text is read aloud in Dari[11] and then the solicitor and enumerators leave the room to allow subjects privacy in their contribution decision. Subjects made their decision and then leave the enumeration location with the money they wish to keep.[12] Note that the solicitation mentions not only the location of the hospital, but also the ethnicity of the recipients of the donations, and does so several times, to strengthen the (abstract) cue of ethnic identity.

We run three treatments: *In-group*, *Out-group-Abstract*, and *Out-group-Real*. Recall that we are focusing on the behavior of the Dari-speakers; hence the beneficiary hospital in the In-group treatment is in a Dari-speaking province (Panjshir), while in both the Out-group treatments it is in the Pashto-speaking province (Helmand). In the *In-group* and *Out-group-Abstract* treatments, we recruited only Dari-speakers, so subjects were waiting only with other Dari-speakers before making their contribution. This is not the case in the *Out-group-Real* treatment, where we recruited Pashto-speakers alongside Dari-speakers at a ratio of 1:2 and had both groups wait in the same courtyard for the experiment. The linguistic mix of around 1 Pashto speaker to 2 Dari speakers is also maintained in each of the 5-7 person enumeration groups. It is important to note that because the solicitor for the charitable contributions is the same person in all of the sessions, any effects that we observe are due to the experimental treatments.

An important aspect of the design is that the order in which subjects were called is randomly determined. We populated the sessions by selecting the nth person from the day's registration list of recruits, where n differed each day. Authors checked in with field personnel every few rounds to

[11] Since Dari is used for official communication, the solicitor read the script in Dari and his words were privately translated to Pashto by the Pashto enumerators, as necessary.

[12] Envelopes were unobtrusively numbered so contributions could be matched with the survey.

make sure that the protocol was being followed correctly. This means that subjects' wait times for enumeration are randomly varying between several minutes to several hours.

Empirical Results

As our focus is on the willingness of non-Pashtun ethnic minorities (Dari-speakers) to contribute to the dominant Pashtun ethnic group, we drop the (few) Pashtun subjects that we have from the analysis and refer to non-Pashtun subjects simply as subjects.[13] Summary statistics of key variables for our subjects are presented in Table 1 (Panel A). The average age of our subjects is 33. Only half are educated at all, and those who have been average 6.8 years of education, which is in line with the best available demographic survey evidence on the average level of education across the entire population (Hopkins 2014). 42.5% of subjects speak more than one language; the average number of languages spoken is 1.6, with Pashto being the most common second language. Subjects earn about 4346 AFN ($75) per month, about average for the broader population (Hopkins 2014), which provides them with $2.50 a day for their family. They have worked as day-laborers for about 10.4 years, on average, but most are able to find work only 2.6 days out of a week. These individual-level attributes allow us in our regression analysis to control for factors that might influence contributory behavior.

Table 1 Panels B, C, and D, show summary statistics across the *In-Group*, *Out-group-Abstract*, and *Out-group-Real* sessions, respectively, and Appendix Table A1 shows the balance of demographic variables across our treatment conditions. While subjects are mostly quite similar across the sessions, we note that there is evidence of some imbalance across treatment conditions

[13] Summary statistics on Pashtun subjects' contributions in *Out-group-Real* are shown in Appendix Table A4 and depicted in Appendix Figure A4.

on several characteristics. Subjects in *Out-group-Abstract* worked fewer days and earn less than subjects in the other two conditions, while subjects in *Out-group-Real* are younger and are less likely to feel poor than subjects in the other two. This may be due to differences in demographics of day-laborers across market locations. However, controlling for these characteristics in our regression models does not substantially affect results.

Average contribution in *In-Group* is 20 AFN while contribution in Out-group is 21.4 AFN when Pashtuns are not physically present (*Abstract*) and 16 AFN when Pashtuns are present (*Real*). Figure 1 shows the kernel density plots of contributions in the three treatment conditions. The figure indicates that the contributions in *In-group* and *Out-group-Abstract* are similar, but the density of contributions in *Out-group-Real* appears quite different, particularly in how many more subjects are giving nothing relative to the other conditions. The figure also clearly shows the non-normality of the variables' distributions, with outliers visible in the upper tails of the data in *Out-group-Abstract* and *Out-group-Real*. Indeed, Shapiro-Wilk tests reject the null hypothesis that the contributions variable is normally distributed in any of the three treatment conditions ($p<0.01$), indicating the need to account for this in tests for treatment effects.

We therefore use Wilcoxon rank-sum (non-parametric) tests to test for differences in contribution levels across treatment conditions. The results confirm that while abstract exposure has no effect on giving relative to the in-group condition, physical exposure does. Specifically, these tests cannot reject the null hypothesis that contributions in *In-group* are the same as those in *Out-group-Abstract* ($p=0.96$), but do reject the null for contributions in *In-group* vs. *Out-group-Real*

(p=0.04) and in *Out-group-Abstract* vs. *Out-group-Real* (p=0.04), all of which is suggested by Figure 1.[14]

As a robustness check, Appendix Table A2 reports results from Wilcoxon rank-sum tests for differences in contributions using sessions, rather than individuals, as the unit of analysis. Our earlier findings are confirmed even with these rather conservative tests: the difference in contributions between *In-group* and *Out-group-Abstract* is again insignificant (p=0.65), but the decreases in giving between *Out-group-Abstract* and *Out-group-Real*, as well as between *In-group* and *Out-group-Real*, are significant ($p<0.05$, p=0.08, respectively).

Finally, we show estimates of main treatment effects from standard regression analysis (OLS) in Table 2.[15] The base model in column 1 regresses an individual's contributed amount on the treatment dummy variables of interest (*Out-group-Abstract* and *Out-group-Real*). Only two controls are included: enumerator fixed effects and a 'session' variable that indicates wait time for enumeration. The latter variable is a rough proxy of the opportunity cost of participating in the experiment, since the longer subjects wait in the courtyard, the less chance they have of finding employment opportunities upon leaving our site. Others have found that contributions in experimental games decrease in later rounds (Alexander and Christia 2011, Castro 2008, Chuah et al. 2014, Waring and Bell 2013), and consistent with those studies, the 'session' variable here is

[14] Kolmogorov-Smirnov tests return similar results. The test fails to reject the null hypothesis that contributions in *In-group* are drawn from the same distribution as those in *Out-group-Abstract* (p=1.00), but the null is rejected for *In-group* vs. *Out-group-Real* (p=0.02) and for *Out-group-Abstract* vs. *Out-group-Real* (p=0.03).

[15] Full results of the models are shown in Appendix Table A3.

negative and significant and so we include it in all models. Robust standard errors are clustered at the session-level to account for within-session correlation.

In column 1, we see that there is no statistically significant difference in the amount contributed between *In-group* and *Out-group-Abstract*, but there is a substantively and statistically significant drop in contribution size in *Out-group-Real* of 8.40 AFN ($p<0.05$). Even after adding extensive demographic controls to the regression model (column 2), these results still hold, with an estimated decrease in contribution size of 6.66 AFN for *Out-group-Real* that is still marginally statistically significant.[16]

Result 1: There is no out-group discrimination in the abstract: contribution is the same in In-group *and* Out-group-Abstract*.*

Result 2: There is out-group discrimination when the out-group is "real" (physically present): contribution in Out-group-Real *is lower than contribution in* In-group*.*

Next, we test whether discrimination toward the out-group is increasing in the salience of ethnic identity. Recall that one of the advantages of our experimental design is that the wait for enumeration creates an exogenous increase in the length of physical exposure to a particular social environment in all our treatment conditions. To estimate the effect of time until enumeration on behavior, we create a dummy variable that splits the enumeration sessions from a given day into two groups, Early and Late. Subjects in the 'Late' category had to wait until the later half of the morning for enumeration (i.e., between about 2.5 and 5 hours).

Table 2 (columns 3-4) shows regression results from the main model (column 2), but estimates a separate model for Early and Late sessions. We see that there is no discrimination in the Early

[16] Results of a *F*-test show that the demographic controls in column 2 have little explanatory power ($p=0.54$), but we include them in subsequent models, though the results on those variables are omitted in the tables.

sessions (column 3). However, by the Late sessions (column 4), subjects who have been waiting alongside Pashtuns (*Out-group-Real*) are now giving 12.78 AFN less ($p<0.05$) to the Helmand hospital than subjects who did not wait alongside Pashtuns. Figure 2 illustrates this by plotting the mean contributions and 95% confidence intervals in the three treatments. While Panel A pools all the sessions together, Panel B splits them by Early and Late. Panel B shows that despite starting at a similar level to the other two treatments at the first half of the day, contributions in *Out-group-Real* has decreased significantly relative to the other two by the later half.

Result 3: Discrimination against the out-group intensifies with physical exposure to the out-group: contributions in Out-group-Real *are the same as that in* In-group *in Early sessions, but by Late sessions, contributions have decreased by half.*

This result is noteworthy not only for what it reveals about how prolonged exposure to the out-group affects behavior within this experiment, but it is worth remarking on how it may fit into an emerging set of experimental findings on the effect of time spent with the out-group. One can think of this experiment as showing the effects of initial exposure to the out-group (including those who might normally self-select into interaction with the out-group in everyday life and those who would not), lasting up to about 5 hours. This length of exposure negatively affects prosocial behavior toward the out-group. But from this experiment alone, we do not know what would happen if, say, we lengthened that time of interaction, or if we exposed non-Pashtuns to the Pashtun out-group repeatedly over multiple days, or if we institutionalized interaction in a more formal way.

Fortunately, others' experimental results provide some idea of what might happen under these conditions, and taken together, these results fill out our understanding of exposure's effects more generally. Enos (2014) exposes whites to Spanish-speakers for a period of minutes on a daily basis while waiting for their commuter train, either for three days or ten days in succession. He finds that this exposure negatively affects whites' views of the out-group (as measured through

survey opinions on immigration policy), but the size of that negative effect diminishes for subjects exposed to the out-group over the 10-day span, relative to 3-days. This suggests that "repeated exposure to an outgroup can mitigate initial negative reactions" (Enos 2014, 3702). Results in Alexander and Christia (2011) dovetail with this observation about the effects of repeated exposure over a long period of time. They show that only Bosnian students who previously had been (randomly) placed in an ethnically segregated school decreased their contributions in a public goods game in response to increasing ethno-religious diversity of experimental sessions; students who were attending an ethnically mixed school were not affected by the ethnic diversity of the session. That investigation is based on a much longer period of subjects' exposure to the out-group, suggesting a further decrease in negative affect toward the out-group with repeated exposure over multiple weeks, consistent with the contact hypothesis. Taken together, these three sets of results suggest the possibility of an inverted-U-shaped curve, with discrimination toward the out-group increasing in initial physical exposure to the out-group, but then gradually decreasing after some peak.

Exploring the Mechanism

The fact that we observe discrimination in the *Out-group-Real* condition but not in *Out-group-Abstract*, and that this discrimination increases with length of exposure to the out-group, is consistent with the argument that behavior toward the out-group differs markedly depending on the salience of ethnicity. In this section we lay out and test a further observable implication of this mechanism, though we caution that our reasoning and analysis is *post hoc* and should be considered speculative. Nevertheless, the argument is supported by the data, and the evidence is consistent with other studies of similar behavior in countries with a history of interethnic conflict.

We end the section by ruling out several alternatives to the argument that would challenge our interpretation of the main results.

We begin an investigation of the mechanism driving the observed behavior by noting that "[f]actors that curb contact's ability to reduce prejudice are now the most problematic theoretically, yet the least understood", and "the effects of these factors are likely to be moderated by the degree to which group membership is salient during contact" (Pettigrew and Tropp 2006, 767). Indeed, evidence from other experiments shows that subjects' reaction to cueing the out-group is moderated by characteristics of both the individual and the setting (Glaeser, Laibson, and Scheinkman 2000, Haile, Sadrieh, and Verbon 2008), which informs our examination of subject characteristics in our sample.

The political context of Afghanistan discussed above suggests two demographic factors in particular that we might expect to moderate any effects of Dari-speaking minorities' exposure to the Pashto-speaking dominant group on contributions to the public good. The first is age. Anecdotal evidence points to the youth of Afghanistan being less tied to ethnic identity than older Afghans, perhaps because of faint to no memory of the brutal ethnic rivalries of the 1990s.[17] This is supported by evidence on interethnic altruism from other contexts that have a similar history of interethnic violent conflict. For example, in dictator games designed to measure interethnic altruism and carried out in Syria, Boznia-Herzegovina, and Kosovo, Whitt, Wilson, and Mironova (2016) find that the more time that has passed since civil conflict, the more altruistic are their subjects toward the out-group.

[17] A Kabul University professor observes that "the young Tajiks, the young Uzbeks, the young Pashtuns, they are all friends. The ethnic card doesn't work anymore" (Craig 2014).

The second factor is language. One of the prerequisites for quality interaction between groups would be the ability to speak to each other. Those subjects who also can speak Pashto are better able to interact with the out-group, possibly facilitating more positive attitudes and less discrimination. Bilingualism here may also be a signal of selection for Dari-speaking subjects who have interacted significantly with Pashtuns in the past.

If physical exposure increases the salience of ethnic identity, then we would expect that in *Out-group-Real*, those most likely to hold out-group prejudices – the old and those who do not speak Pashto – should decrease their contributions to the Pashtun out-group. By contrast, the young and bilingual – who are less likely to hold such prejudice – should not be as affected by physical exposure. We do not have an *a priori* theoretical rationale for predicting which of these factors – age or language – should mediate the effect of our treatments to a greater degree.

We therefore look for age and linguistic affinity as mediating variables through regression analysis in which the treatment variables (*Out-group-Abstract* and *Out-group-Real*) are interacted with dummy variables indicating an age under the median of 30 (*Young*) and an ability to speak Pashto (*Speaks Pashto*). Results of these models (which include the full set of demographic controls) are reported in Table 3 and discussed below. In columns 2-4, we investigate the mediating effect of language by including interactions of the treatment dummy with *Speak Pashto*. Following our findings that discrimination emerges over time (see Table 2, columns 3-4), in columns 3-4 we split the sample to Early and Late. In columns 5-7, we follow the same strategy with *Young*.

We first investigate the effect of demographics on the abstract cue. First, the treatment dummy *Out-Abstract* is never significant, suggesting that the abstract cue does not have a significant negative effect on non-Pashtun minorities who do not also speak Pashto in addition to their

primary language (columns 2-4) or on those who are older (columns 5-7). Next, the coefficient on *Out-Abstract x Speak Pashto* is significantly positive in column 2 (All, 13.89) and column 4 (Late, 18.04) but never in *Out-Abstract x Young* (columns 5-7), suggesting that the significant difference in reaction to *Out-Abstract* lies not across age groups but across language abilities. Looking at the linear combinations of coefficients (*Out-Abstract + Out-Abstract x demographic dummy*) in the bottom of Table 3, we can see that the abstract cue has a positive effect on subjects that are bilingual in Pashto (column 2, 11.42) and marginally, on younger subjects especially in the early session (column 6, 8.438).

These results are a surprise to us. The abstract Pashtun cue did not cause a significant negative reaction among any subgroups of our subjects, who as non-Pashtun minorities are likely to perceive Pashtuns as a threat especially if they do not speak Pashto or are old enough to have experienced the civil war. However, this cue is strong enough to cause an increase in giving among those subjects who also speak Pashto in addition to their primary language, Dari. Interestingly, this phenomenon of giving more to the out-group is also observed in Alexander and Christia (2011), where Bosnian students attending ethnically mixed schools actually contributed *more* in public goods games in more ethnically diverse sessions. It may be reflective of how minorities show

deference toward a more dominant out-group in some circumstances,[18] though we cannot be sure. This is the fourth result.[19]

Result 4: Neither those who Do not speak Pashto nor Older negatively discriminates in the Out-group-Abstract. *Non-Pashtun minorities who also speak Pashto appear to favor Pashtuns in their contributions, giving 13.89 AFN ($p<0.05$) more in* Out-group-Abstract *than in* In-group.

We now examine how demographics mediate the effect of the physical out-group cue (*Out-group-Real*). Unlike *Out-group-Abstract*, the coefficient on the treatment dummy *Out-group-Real* is significantly negative across columns 2-7, especially in the Late sessions. This indicates that non-Pashtun minorities who are unable to speak Pashto (column 4, -18.30) or are older (column 7, -15.96) become significantly more unwilling to give to Pashtuns as their exposure to Pashto-speakers enters its second or third hour. On the other hand, the interaction terms *Out-Real x Speak Pashto* and *Out-Real x Young* are always positive and relatively large in magnitude, though rarely significant. When the linear combination of *Out-group-Real+Out-group-Real x demographic dummy* is computed, we can see that non-Pashtun minorities who also speak Pashto in addition to their primary language (Dari) or are younger are relatively unaffected by the physical cue, even in the case of prolonged exposure that continues into the later half of the day (column 4, -6.087 ($p>0.10$) and column 7, -5.629 ($p>0.10$)).

---

[18] This idea has been suggested by survey research. Minority Egyptian Christian women (particularly the poor and less educated) deemphasized their religious identity in the presence of Muslim enumerators (Blaydes and Gillum 2013). Similarly, African Americans are more accommodating of whites when interviewed by whites (Davis 1997).

[19] Appendix Figures A1 and A2 graphically illustrate contributions among non-Pashtun minorities across treatment conditions, split out by the demographic conditions investigated here: ability to speak Pashto (Figure A1) and age (Figure A2).

This is our fifth result. The findings are consistent with what we would expect: for those most likely to hold out-group prejudices against the Pashtuns – the old and those who do not speak Pashto – waiting alongside Pashto-speakers may have activated a vague sense of threat that may have accumulated over time and decreased contributions. On the other hand, the young and bilingual – who are less likely to hold anti-Pashtun prejudice – appears unaffected by the physical exposure and increased salience of ethnic identity.

Result 5: Those who do not speak the out-group's language or are old enough to have experienced past conflicts with the out-group discriminate when the out-group is physically present. Minorities who do not speak Pashto give 13.11 ($p<0.05$) less in Out-group-Real *than they do in the* In-group *treatment; older minorities give 9.42 ($p<0.05$) less.*

In concluding our discussion of these results as indicative of out-group discrimination, we address several possible alternative explanations and their observable implications in light of the data. The first is that any treatment effects can be explained by the expectation of (monetary) reciprocal generosity, as is the case in some other experiments that use trust games. Our design, however, precludes any personal monetary return to our subjects' contributions, and hence behavior can only be altruistic, not strategic. The second explanation we can rule out is reputation or social sanctioning. The complete anonymity attached to the contribution decision eliminates sanctioning (either within the experiment or after it concluded) as explanations for behavior. The third possibility is that observed behavior is simply an artifact of opportunity cost. While other studies report decreased contributions in later rounds of experiments for subjects who play iteratively, the exogenous time of exposure to the out-group allows us to isolate the part of time's effect on contributions that derives simply from opportunity cost of waiting from the part of the effect that is due to physical exposure to the out-group (*Out-group-Real*). The difference across conditions over time is consistent with discrimination emerging from increasing salience of ethnicity. The fourth alternative explanation is that Dari-speakers might assume that Pashtuns will

"take care of their own" in the *Out-group-Real* condition, prompting Dari-speakers to reduce their contributions without fear that the Pashtun-area hospital would receive less funding. But note that if such anticipation were guiding Dari-speakers' behavior (and not discrimination), we would expect that not only would Dari-speakers' contributions to the Pashtun-area hospital drop in Late sessions in the *Out-group-Real* condition, but contributions also would drop noticeably in Early sessions of that condition. In other words, wait time should not affect behavior by the logic of this alternate explanation, but we see clearly that it does.

Conclusion

In areas where the government is either unwilling or unable to provide public goods, citizens turn to social networks of their own in an effort to fill the gap (Humphreys and Weinstein 2009). The question of how ethnic diversity affects willingness to contribute to such efforts is thus particularly important in these settings. In a lab-in-the-field experiment conducted in the ethnically diverse city of Kabul, Afghanistan, we examine how both abstract and real exposure to a dominant, Pashto-speaking ethnic out-group affects Dari-speaking ethnic minorities' willingness to sacrificially contribute to enhance the welfare of strangers from that out-group, without any expectation of a return on that investment. The fact that our subjects are drawn directly from the population whose behavior we care about, and the fact that contribution decisions are made in the context of a real and relevant public good, enables a more meaningful interpretation of the empirical results.

These results reveal the complicated nature of ethnic diversity's effects on cooperative and altruistic behavior. While subjects were just as willing to contribute to help their own in-group as the out-group Pashtuns in the abstract, physical exposure to the out-group showed the limits of this generosity, as contributions decreased markedly. The difference across these two 'cues' of the out-

group is evidence of discrimination, an argument supported by the fact that the longer subjects were exposed to the out-group, the less they contributed. While the unequal status of Pashtuns and non-Pashtuns in Afghanistan makes it difficult to generalize to contexts in which groups are more equally situated politically and culturally, Afghanistan is far from an outlier case. Many weak countries transitioning out of conflict feature politically and socially unequal ethnic groups.

We reason that physical exposure to the out-group increases the salience of ethnic differences, but found that this has non-uniform effects on prosocial behavior depending on age and ability to speak the language of the out-group, which we speculate are measures of one's affinity toward the out-group. In the Afghan context, young, bilingual Afghans show evidence of having moved beyond ethnicity as a reason to discriminate against others. These findings are in line with classic theories which posit that assimilation to the dominant group in society occurs by gaining access to that group's social structures (Gordon 1964), and language and youth should help accomplish this. In contrast, those in society least able to bridge the ethnic divide – the older and those who do not speak Pashto – are likely to inhibit progress in this area.

References

ACSOR. 2014. Afghanistan's Ethnic, Regional Divisions Produce a Dead Heat in its Presidential Race. Langer Research Associates.

Adida, Claire L., David D. Laitin, and Marie-Anne Valfort. forthcoming. "'One Muslim Is Enough!' Evidence from a Field Experiment in France." *Annals of Economics and Statistics*.

Alexander, Marcus, and Fotini Christia. 2011. "Context Modularity of Human Altruism." *Science* 334 (6061):1392-1394.

Allport, Gordon. 1954. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley Publishing Company.

Barfield, Thomas J. 2010. *Afghanistan: A Cultural and Political History*. Princeton, NJ: Princeton University Press.

Bernhard, Helen, Urs Fischbacher, and Ernst Fehr. 2006. "Parochial Atruism in Humans." *Nature* 442 (7105):912-15.

Blaydes, Lisa, and Rachel M. Gillum. 2013. "Religiosity-of-Interviewer Effects: Assessing the Impact of Veiled Enumerators on Survey Response in Egypt." *Politics and Religion* 6 (3):459-482.

Bouckaert, Jan, and Geert Dhaene. 2004. "Inter-Ethnic Trust and Reciprocity: Results of an Experiment with Small Businessmen." *European Journal of Political Economy* 20:869-886.

Burns, Justine. 2004. "Race and Trust in Post-Apartheid South Africa." *CSSR Working Paper 78*.

Callen, Michael, and James D. Long. 2015. "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review* 105 (1):354-381.

Castro, Massimo Finocchiaro. 2008. "Where are you from? Cultural differences in public goods experiments." *Journal of Socio-Economics* 37:2319-2329.

Chandra, Kanchan. 2006. "What is Ethnic Identity and Does it Matter?" *Annual Review of Political Science* 9:397-424.

Chuah, Swee-Hoon, Reema Fahoum, and Robert Hoffmann. 2013. "Fractionalization and Trust in India: A Field-Experiment." *Economics Letters* 119 (2):191-194.

Chuah, Swee-Hoon, Robert Hoffmann, Bala Ramasamy, and Jonathan H. W. Tan. 2014. "Religion, Ethnicity and Cooperation: An Experimental Study." *Journal of Economic Psychology* 45:33-43.

Condra, Luke N., Mohammad Isaqzadeh, and Sera Linardi. 2016. "Clerics and Scriptures: Experimentally Disentangling the Influence of Religious Authority in Afghanistan." Working paper, University of Pittsburgh and American University of Afghanistan, July 8.

Craig, Tim. 2014. "Abdullah Abdullah, front-runner in Afghan presidential race, seeks to quell ethnic fears." *Washington Post*, May 18. http://wpo.st/P9ZH1.

Davis, Darren W. 1997. "The Direction of Race of Interviewer Effects among African-Americans: Donning the Black Mask." *American Journal of Political Science* 41 (1):309-322.

Enos, Ryan. 2014. "Causal Effect of Intergroup Contact on Exclusionary Attitudes." *Proceedings of the National Academy of Sciences of the United States of America* 111 (10):3699-3704.

Fershtman, Chaim, and Uri Gneezy. 2001. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics* 116 (1):351-377.

Fong, Christina M., and Erzo F. P. Luttmer. 2009. "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty." *American Economic Journal: Applied Economics* 1 (2):64-87.
Giustozzi, Antonio, and Mohammad Isaqzadeh. 2012. *Policing Afghanistan*. New York: Columbia University Press.
Glaeser, Edward L., David I. Laibson, and Jose A. Scheinkman. 2000. "Measuring Trust." *Quarterly Journal of Economics* 115 (3):811-846.
Gordon, Milton Myron. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101 (4):709-725.
Haile, Daniel, Abdolkarim Sadrieh, and Harrie A. A. Verbon. 2008. "Cross-Racial Envy and Underinvestment in South African Partnerships." *Cambridge Journal of Economics* 32 (5):703-724.
Havekes, Esther, Marcel Coenders, and Tanja Van der Lippe. 2013. "Positive or negative ethnic encounters in urban neighborhoods? A photo experiment on the net impact of ethnicity and neighbourhood context on attitudes towards minority and majority residents." *Social Science Research* 42:1077-1091.
Hopkins, Daniel J. 2010. "Politicized Places: Explaining Where and When Immigrants Provoke Local Opposition." *American Political Science Review* 104 (1):40-60.
Hopkins, Nancy, ed. 2014. "Afghanistan in 2014: A Survey of the Afghan People." *The Asia Foundation*, Available from: http://asiafoundation.org/resources/pdfs/Afghanistanin2014final.pdf.
HRW. 1998. "Afghanistan: The Massacre in Mazar-i Sharif." Human Rights Watch Report, Available from: https://www.hrw.org/legacy/reports98/afghan/Afrepor0.htm.
Humphreys, Macartan, and Jeremy M. Weinstein. 2009. "Field Experiments and the Political Economy of Development." *Annual Review of Political Science* 12:367-78.
Jeon, Sangick, Tim Johnson, and Amanda Lea Robinson. forthcoming. "Nationalism and Social Sanctioning across Ethnic Lines: Experimental Evidence from the Kenya-Tanzania Border." *Journal of Experimental Political Science*.
Jochem, Torsten, Ilia Murtazashvili, and Jennifer Murtazashvili. 2015. "Social Identity and Voting in Afghanistan: Evidence from a Survey Experiment." *Journal of Experimental Political Science* 2:47-62.
Lacina, Bethany, and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21 (2-3):145-166.
Lambarraa, Fatima, and Gerhard Riener. 2012. "On the Norms of Charitable Giving in Islam: A Field Experiment." *Courant Research Centre Discussion Papers*, Available from: https://ideas.repec.org/p/ags/iaae12/126795.html.
Mogelson, Luke. 2012. "Pacifists in the Cross-Fire: The Kabul Hospital that Treats All Sides." *New York Times*, May 18.
Nordland, Rod. 2012. "Top Afghans Tied to '90s Carnage, Researchers Say." *New York Times*, July 22.
Oxfam. 2009. "The Cost of War: Afghan Experiences of Conflict, 1978-2009." Report. Available from: https://www.oxfam.org/sites/www.oxfam.org/files/afghanistan-the-cost-of-war.pdf.

Pettigrew, Thomas F. 1998. "Intergroup Contact Theory." *Annual Review of Psychology* 49 (1):65-85.
Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90 (5):751-783.
Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture." *Scandinavian Political Studies* 30:137-174.
Riphenburg, Carol J. 2005. "Ethnicity and Civil Society in Contemporary Afghanistan." *The Middle East Journal* 59 (1):31-51.
Robinson, Amanda Lea. forthcoming. "Nationalism and Ethnic-Based Trust: Evidence from an African Border Region." *Comparative Political Studies*.
Shahrani, M. Nazif. 1986. "State Building and Social Fragmentation in Afghanistan: A Historical Perspective." In *The State, Religion, and Ethnic Politics: Afghanistan, Iran, and Pakistan*, edited by Ali Banuazizi and Myron Weiner, 23-74. Syracuse, NY: Syracuse University Press.
Simonsen, Sven Gunnar. 2004. "Ethnicising Afghanistan? Inclusion and Exclusion in Post-Bonn Institution Building." *Third World Quarterly* 25 (4):707-729.
Waring, Timothy M., and Adrian V. Bell. 2013. "Ethnic Dominance Damages Cooperation More than Ethnic Diversity: Results from Multi-Ethnic Field Experiments in India." *Evolution and Human Behavior* 34:398-404.
Whitt, Sam, Rick K. Wilson, and Vera Mironova. 2016. "Evidence for Out-Group Altruism in the Aftermath of Violence." Working paper.

Table 1. Summary Statistics on Demographics for Subjects across All Treatments

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A: All Non-Pashtun Subjects					
Amount Contributed (AFN)	214	19.1	20.4	0	100
Age	214	33	12.1	18	79
Languages Spoken	214	1.61	0.721	1	4
Married	213	0.793	0.406	0	1
No Education	214	0.500	0.501	0	1
Years of Education (if educated)	96	6.833	3.131	2	14
Monthly Income (AFN)	214	4346	2212	88	12000
Do not feel poor	214	0.210	0.408	0	1
Days of work last week	214	2.57	1.26	0	7
Time in occupation (years)	214	10.4	8.69	0	50
Panel B: Non-Pashtuns contributing to Non-Pashtuns(In-group)					
Amount Contributed (AFN)	67	20.0	17.7	0	50
Age	67	36	13.3	18	79
Languages Spoken	67	1.58	0.781	1	4
Married	66	0.788	0.412	0	1
No Education	67	0.478	0.503	0	1
Years of Education (if educated)	31	7.032	3.440	2	14
Monthly Income (AFN)	67	4427	2385	88	10000
Do not feel poor	67	0.194	0.398	0	1
Days of work last week	67	2.60	1.26	0	7
Time in occupation (years)	67	9.51	8.81	0	50
Panel C: Non-Pashtuns contributing to Pashtuns (Out-group-Abstract)					
Amount Contributed (AFN)	74	21.4	21.7	0	100
Age	74	34	11.5	18	64
Languages Spoken	74	1.51	0.646	1	3
Married	74	0.784	0.414	0	1
No Education	74	0.486	0.503	0	1
Years of Education	33	6.576	3.042	2	13
Monthly Income (AFN)	74	3893	1787	88	10000
Do not feel poor	74	0.135	0.344	0	1
Days of work last week	74	2.18	1.11	0	5
Time in occupation (years)	74	11.8	9.79	0	50
Panel D: Non-Pashtuns contributing to Pashtuns (Out-group-Real)					
Amount Contributed (AFN)	73	16	21.1	0	100
Age	73	29	10.6	18	62
Languages Spoken	73	1.74	0.727	1	4
Married	73	0.808	0.396	0	1
No Education	73	0.534	0.502	0	1
Years of Education (if educated)	32	6.906	2.988	2	12
Monthly Income (AFN)	73	4732	2379	88	12000
Do not feel poor	73	0.301	0.462	0	1
Days of work last week	73	2.96	1.30	0	7
Time in occupation (years)	72	9.72	7.19	0	30

Table 2. Effects of Abstract vs. Real Exposure to Out-group on Contributions

	All	All	Early	Late
	(1)	(2)	(3)	(4)
VARIABLES	Amount	Amount	Amount	Amount
Out-Abstract	1.54	2.44	4.52*	0.38
	(2.36)	(2.46)	(2.29)	(3.20)
Out-Real	-8.40**	-6.66*	-5.34	-12.78**
	(3.41)	(3.93)	(6.14)	(5.67)
Constant	34.47***	25.23*	25.34	29.10
	(5.28)	(13.63)	(20.39)	(18.82)
Demographic Controls	No	Yes	Yes	Yes
Observations	214	212	99	113
R-squared	0.08	0.12	0.20	0.18

Notes: All models control for session and include interviewer fixed effects. Robust standard errors clustered on session. See Table A3 for complete results. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Effects of Abstract vs. Real Exposure to Out-group on Contributions over Time (by Speaks Pashto and Young)

		Speaks Pashto			*Young*		
	All	All	Early	Late	All	Early	Late
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Out-Abstract	1.714	-2.475	2.963	-6.087	2.389	1.139	1.940
	(2.424)	(3.139)	(3.908)	(4.746)	(3.320)	(4.841)	(5.100)
Out-Real	-8.924**	-13.11**	-9.226	-18.30***	-9.421**	-5.452	-15.96***
	(4.405)	(4.963)	(8.489)	(4.572)	(4.286)	(6.858)	(4.670)
Later half of morning	2.071	1.157			2.314		
	(4.501)	(4.611)			(4.230)		
Speaks Pashto	-1.730	-10.77**	-9.902	-7.711	-5.637	-7.814	-1.509
	(3.258)	(4.563)	(5.993)	(6.388)	(3.889)	(5.214)	(5.649)
Out-Abstract * Speaks Pashto		13.89**	4.112	18.04*			
		(6.703)	(8.502)	(10.42)			
Out-Real * Speaks Pashto		12.10**	10.65	12.38			
		(5.605)	(8.613)	(7.556)			
Young	6.642	6.436	4.161	4.676	3.887	0.285	1.063
	(4.314)	(4.296)	(5.139)	(7.713)	(4.936)	(7.916)	(4.916)
Out-Abstract * Young					-1.199	7.299	-5.792
					(6.356)	(8.542)	(9.261)
Out-Real * Young					7.919	5.102	10.33
					(5.934)	(9.273)	(7.189)
Constant	20.28**	22.92**	17.48	27.37	22.45**	19.21	34.48*
	(8.534)	(8.607)	(11.88)	(16.20)	(8.429)	(11.25)	(18.39)
Observations	212	212	99	113	212	99	113
R-squared	0.134	0.152	0.222	0.213	0.143	0.220	0.207
		Effect of treatment if subject also speaks Pashto			Effect of treatment on youth		
Out-Abstract + Out-Abstract* demog. dummy		11.42**	7.075	11.96	1.190	8.438*	-3.852
		(5.376)	(5.494)	(8.188)	(4.507)	(4.390)	(6.195)
Out-Real + Out-Real* demog. dummy		-1.104	1.420	-5.915	-1.502	-0.350	-5.629
		(5.576)	(6.649)	(8.641)	(6.468)	(9.445)	(7.721)

All models control for session ordering, enumerator fixed effects, and demographics. Robust standard errors clustered on session. *** p<0.01, ** p<0.05, * p<0.1

Figure 1. Kernel Density Plots of non-Pashtun minorities' Contributions to EMERGENCY Hospital across all treatments

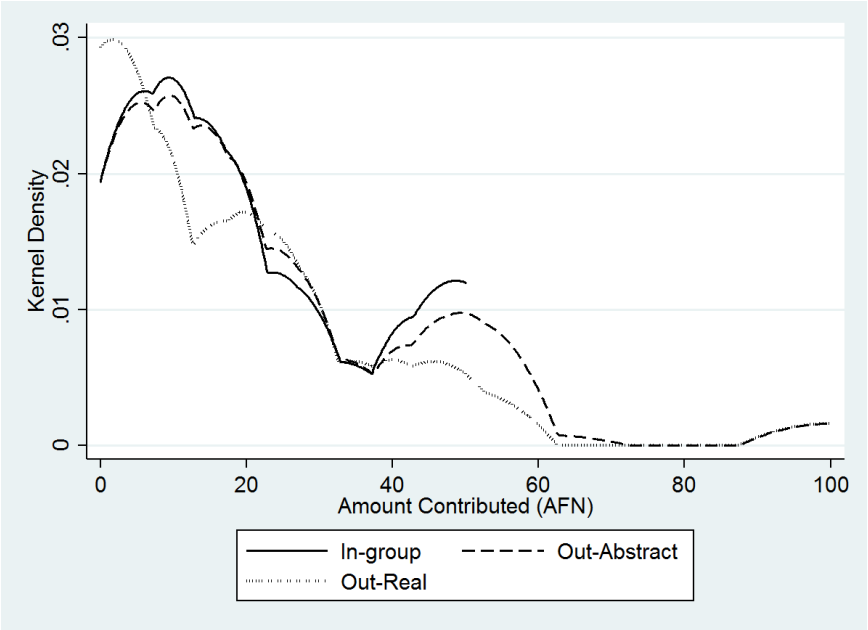
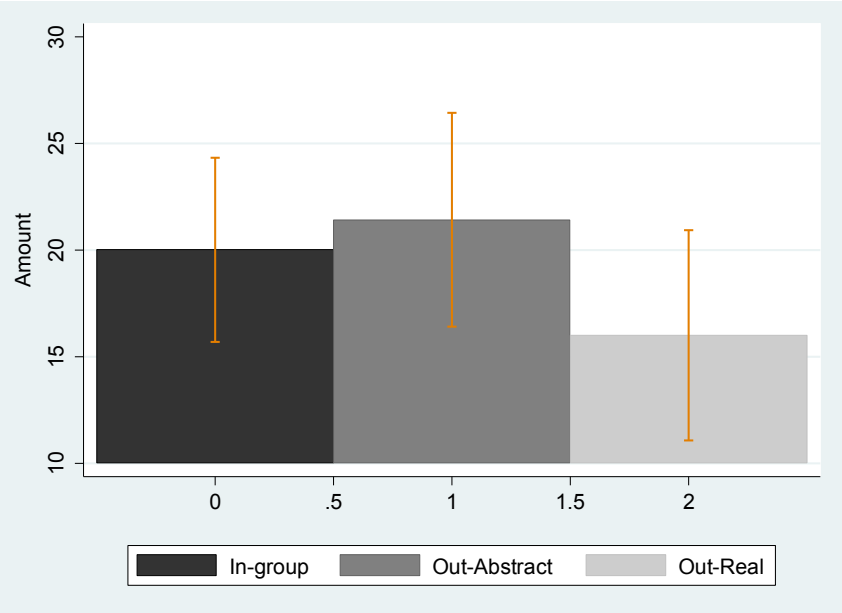
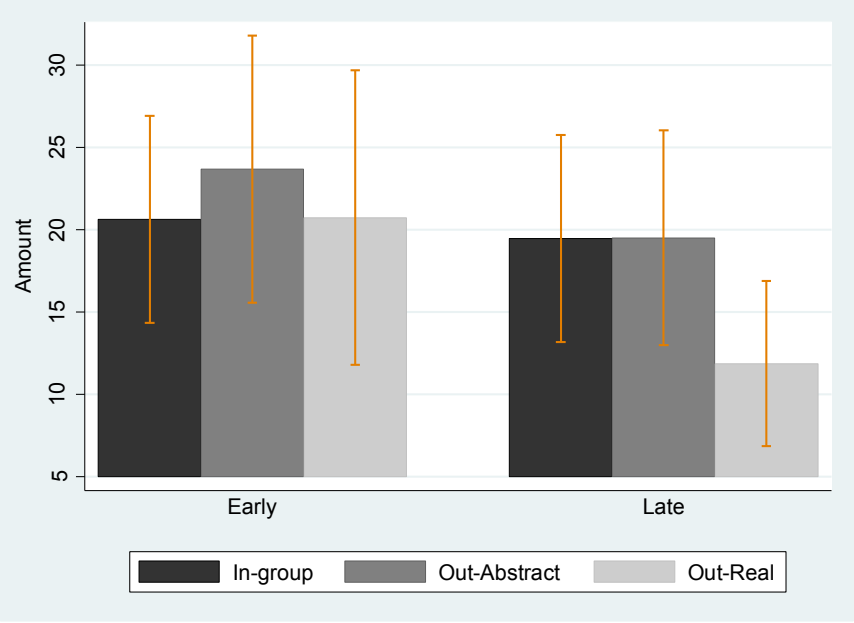


Figure 2a. Average Contributions of Non-Pashtuns to EMERGENCY Hospital across all treatments



Note: Bars depict subjects' mean contribution amounts, by treatment condition, with associated 95% confidence intervals.

Figure 2b. Average Contributions of Non-Pashtuns (Early vs. Late Sessions)



Note: Contributions from non-Pashtun minorities are the same across treatments in Early sessions, but drop in Late sessions in the *Out-group-Real* treatment (with Pashtuns physically present). Bars depict mean subjects' contribution amounts, by treatment condition, with associated 95% confidence intervals.

Appendix Table A1. Balance of Demographics across Treatment Conditions (Dari-speaking subjects)

	Means of Demographic Variables			*p-values of Differences in Means*		
	In-group	Out-group-Abstract	Out-group-Real	In vs Out-Abs	Out-Abs vs Out-Real	In vs Out-Real
N	67	74	73			
Age	36.07	34.27	29.26	0.39	0.01	0.01
Language	1.58	1.51	1.74	0.57	0.05	0.22
Marital Status	0.79	0.78	0.81	0.95	0.72	0.77
No education	0.48	0.49	0.53	0.92	0.57	0.51
Years of education (if educated)	7.03	6.58	6.91	0.58	0.66	0.88
Monthly income	4426	3892	4732	0.13	0.02	0.45
Years in occupation	9.51	11.78	9.72	0.15	0.15	0.87
Days worked	2.59	2.18	2.96	0.04	0.001	0.1
Do not feel poor	0.19	0.135	0.3	0.34	0.15	0.01

Appendix Table A2. Average Contributions across Treatment Conditions

	In-group	Out-group-Abstract	Out-group-Real	In vs Out-Abstract	Out-Abstract vs Out-Real	In vs Out-Real
Panel A: Individual Level						
N	67	74	73			
Amount	20	21.4	16	p=0.96	p=0.04	p=0.04
Panel B: Session Level						
N	10	11	17			
Amount	19.94	21.23	15.88	p=0.65	p=0.04	p=0.08

Notes: *p*-values correspond to Wilcoxon rank-sum tests

Appendix Table A3. Effects of Abstract vs. Real Exposure to Out-group on Contributions

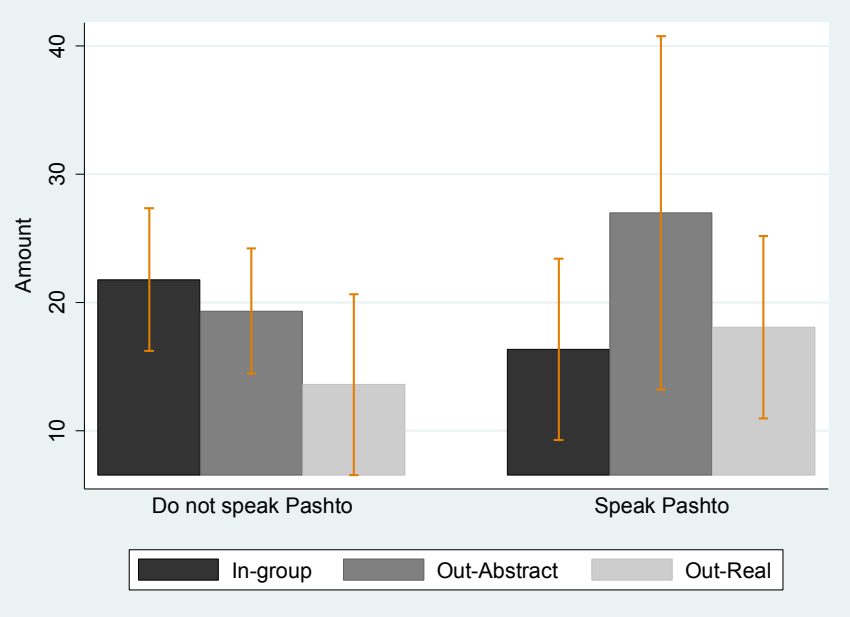
	All	All	Early	Late
	(1)	(2)	(3)	(4)
VARIABLES	Amount	Amount	Amount	Amount
Out-Abstract	1.54	2.44	4.52*	0.38
	(2.36)	(2.46)	(2.29)	(3.20)
Out-Real	-8.40**	-6.66*	-5.34	-12.78**
	(3.41)	(3.93)	(6.14)	(5.67)
Languages Spoken		-0.34	-2.35	2.49
		(2.90)	(6.35)	(2.42)
Age		-0.01	-0.08	0.13
		(0.19)	(0.31)	(0.27)
Married		9.11*	18.83**	0.07
		(4.74)	(7.36)	(6.13)
Years of Education		0.33	0.79*	-0.13
		(0.40)	(0.45)	(0.56)
Monthly Income (AFN)		0.00	-0.00	0.00
		(0.00)	(0.00)	(0.00)
Time in occupation (years)		-0.11	-0.36	0.01
		(0.21)	(0.35)	(0.24)
Days of work last week		0.12	-0.62	0.11
		(1.22)	(2.09)	(1.68)
Do not feel poor		-1.98	2.22	-0.33
		(3.62)	(6.10)	(4.09)
Session	-0.53**	-0.42*	-1.02**	-0.62
	(0.20)	(0.22)	(0.47)	(0.42)
Constant	34.47***	25.23*	25.34	29.10
	(5.28)	(13.63)	(20.39)	(18.82)
Observations	214	212	99	113
R-squared	0.08	0.12	0.20	0.18

Notes: All models include interviewer fixed effects. Robust standard errors clustered on session.
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table A4. Summary Statistics on Pashtun Subjects in *Out-group-Real*

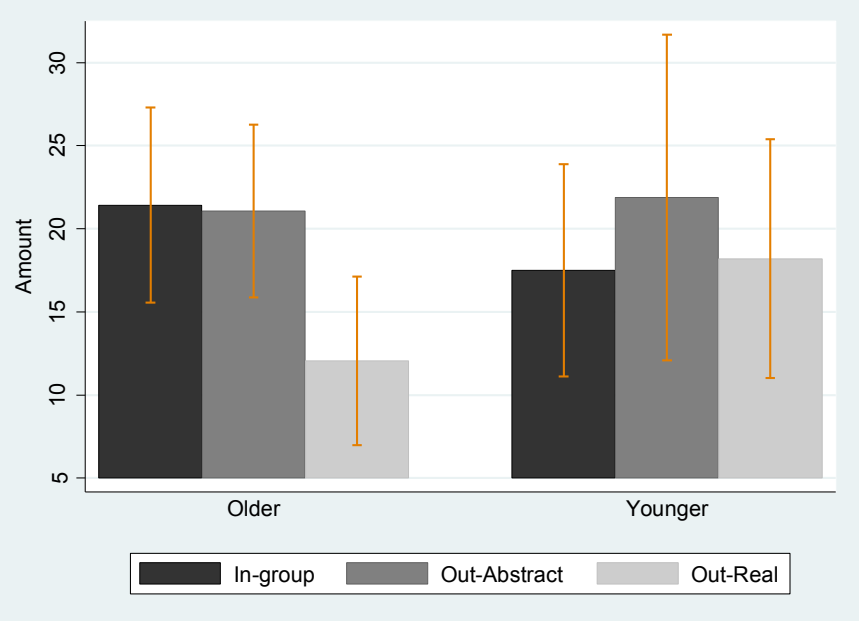
Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Pashtuns contributing to Pashtuns (Out-group-Real)					
Amount Contributed (AFN)	33	9.70	14.47	0	50
Age	33	30.48	9.79	18	55
Languages Spoken	33	1.55	0.56	1	3
Married	33	0.91	0.29	0	1
No Education	33	0.73	0.45	0	1
Years of Education (if educated)	9	6.11	2.42	3	11
Monthly Income (AFN)	33	4078.42	1951.37	88	9000
Time in occupation (years)	33	10.70	8.00	0	35
Days of work last week	33	2.36	0.90	1	5
Do not feel poor	33	0.36	0.49	0	1

Appendix Figure A1. Average Contributions of Non-Pashtuns (by ability to Speak Pashto)



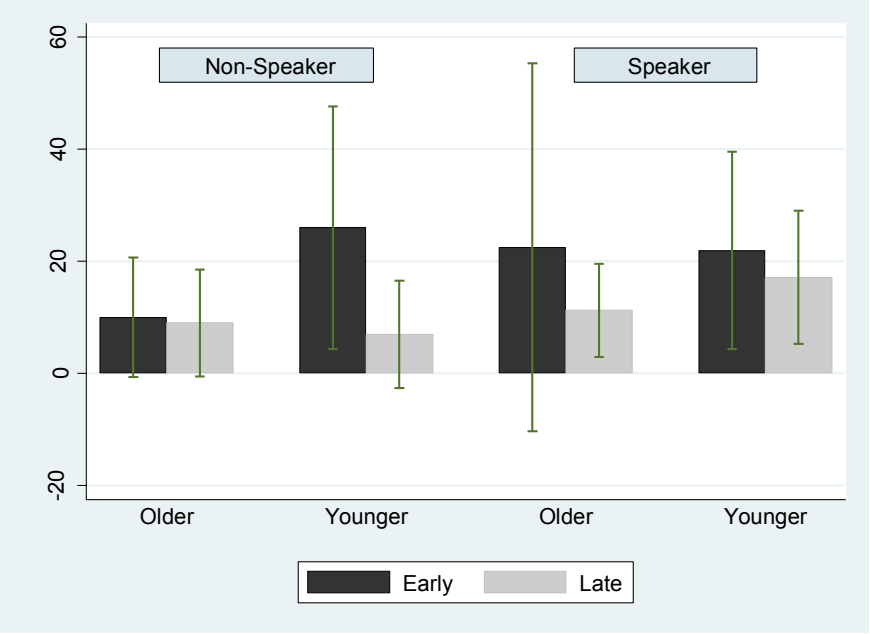
Note: Bars depict subjects' mean contribution amounts, by treatment condition, with associated 95% confidence intervals.

Appendix Figure A2. Average Contributions of Non-Pashtuns (by median age)



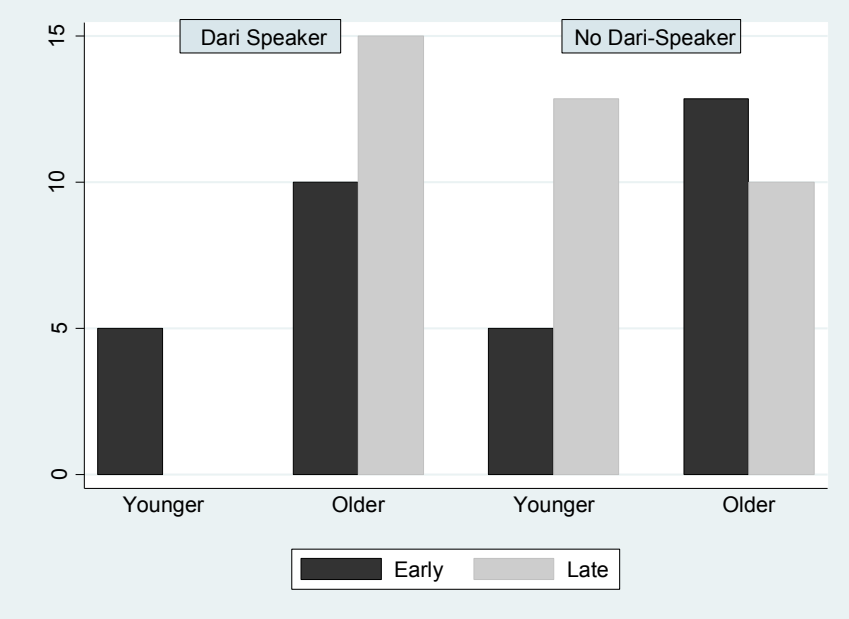
Note: Bars depict subjects' mean contribution amounts, by treatment condition, with associated 95% confidence intervals.

Appendix Figure A3. Average Contributions of Non-Pashtuns to Pashtun-area hospital (by waiting time in *Out-group-Real*)



Note: Bars depict subjects' mean contribution amounts with associated 95% confidence intervals.

Appendix Figure A4. Average Contributions of Pashtuns to Pashtun-area hospital
(by waiting time in *Out-group-Real*)



Subgroups are ordered as follows: young Pashtuns that can also speak Dari, older Pashtuns that can also speak Dari, young Pashtuns that cannot speak Dari, older Pashtuns that cannot speak Dari. Note: contributions of Younger Pashtun who is also a Dari Speaker is 0 AFN in the Late sessions. CI bands are not included because the number of observation represented in each bar is very small (N=33 for all Pashtuns in *Out-group-Real*).

Consistent with Figure A3, the left-most group of bars in Figure A4 is much lower than the right-most group of bars. Also, consistent with Result 6, the length of exposure to the out-group (Dari-speakers) does not affect contributions of groups in the far left or far right of the figure. Rather, it affects the middle group (Older Pashtuns who speak Dari, or Younger Pashtuns who do not speak Dari). As in Figure A3, we see that it is the Younger subjects who do not speak the other group's language who are most affected. While Younger non-bilingual Dari-speakers in Figure A3 *reduce* their giving to the Pashtun-area hospital, the Younger non-bilingual Pashto-speakers here *increase* their giving to the Pashtun hospital. This is suggestive of the threat hypothesis.