Dialogue and interethnic trust: A randomized field trial of 'sustained dialogue' in Ethiopia

Isak Svensson and Karen Brounéus

Journal of Peace Research 2013 50: 563 originally published online 20 August 2013
DOI: 10.1177/0022343313492989

The online version of this article can be found at:
http://jpr.sagepub.com/content/50/5/563
Dialogue and interethnic trust: A randomized field trial of ‘sustained dialogue’ in Ethiopia

Isak Svensson
Department of Peace and Conflict Research, Uppsala University

Karen Brounéus
National Centre for Peace and Conflict Studies, University of Otago

Abstract
The growing field of peacebuilding has tried to mitigate interethnic conflicts by creating various sorts of dialogue programs, aiming to build social bonds and bridges between individuals from groups with a history of violent interaction. Yet, little is known of the effect of dialogue initiatives on interethnic relations and peacebuilding. Previous research on dialogue programs has suffered from the serious problem of selection bias: in other words, by not having comparable control groups it has not been possible to separate selection effects (that a program attracts certain types of people) from process effects (that programs have an effect on people). The present study is the first to examine the effects of a dialogue process in a context of political tension and ethnic violence through a randomized field experiment, thereby eliminating this problem. Using a stratified randomization process, participants were selected to a two-term Sustained Dialogue program at Addis Ababa University, Ethiopia, in 2009–10. Immediately following the dialogue intervention, an attitudinal survey and a behavioral trust game were conducted with a group of 716 participants and non-participants. We found that the program had a positive effect on participants’ attitudes: it worked for decreasing mistrust and increasing the level of trust between people of different ethnic origins. Concurrently, however, participation in the dialogue program increased the sense of importance of ethnic identities as well as the perception of being ethnically discriminated – a somewhat counter-intuitive finding. Participation in dialogue processes had no significant effect on game behavior: participants in Sustained Dialogue were neither more trusting nor trustworthy than non-participants. This study shows the fruitfulness of randomized field-experiments in the area of peace and conflict research and finishes by identifying some important paths for future research.

Keywords
dialogue, Ethiopia, field experiment, peace, randomization

Introduction
Trust and dialogue are essential aspects of group relationships in the contexts of group-based tension, violence, and ethnic conflicts and are therefore an integral part of peacebuilding processes around the world. Dialogue is considered to be fundamental for creating harmonious relationships in multi-ethnic societies and donors invest considerable amounts of resources in programs that aim at promoting trust. The assumption that dialogue is important underpins many efforts to create and maintain positive relations between ethnic groups in various peacebuilding policy programs. This study therefore examines the research question ‘To what extent do dialogue processes carried out in the context of political...”

Corresponding author:
isak.svensson@pcr.uu.se
conflict and ethnic violence contribute to positive change in attitudes and behavior? 

Surprisingly little is known about how people’s propensities to trust can be influenced by policy programs, and in particular, whether dialogue processes indeed enhance trust. Dialogue processes are generally not studied through systematic analysis. Exceptions in terms of empirical studies that have examined the effect of dialogue processes (see Malhotra & Liyanage (2005) for the case of Sri Lanka, as well as Salomon (2004) for the case of Israel–Palestine) have not used a randomized trial design and therefore do not convincingly deal with the potential for selection effects. Recent studies that have utilized randomized trials on peacebuilding interventions in the context of armed conflicts, such as Humphreys & Weinstein (2007) and Fearon, Humphreys & Weinstein (2009), have not examined the effect of dialogue processes but other broader programs. 

This study makes some specific contributions. First, it is the first to use the methodology of randomized trials to explore the effects of dialogue processes. Although previous research has examined general peacebuilding or aid programs aimed at enhancing civil society using this method, none has hitherto examined the effect of programs that have the direct focus of creating dialogue processes in contexts of ethnic violence and political group conflicts. Second, the study utilized two different types of stratifications in the randomized selection process: self-selection and nomination. Through this unique set-up we were able to examine the role played by the selection process and hence determine whether the program had different effects on participants who were identified by others as good candidates for the dialogue process compared to those participants who themselves signed up for the program. Given the importance of selection has had in the discussion on dialogue processes, to know more about the selection process is pivotal. Third, most dialogue studies have been conducted with student populations in relatively peaceful settings. In this study, we study a campus that has experienced much political and ethnic conflict. Recent theses coming out of Addis Ababa University (AAU) show that it has experienced some significant group-based violence in recent years (Hailemariam, 2010; Mulugeta, 2010).

Through a two-term, 12-session Sustained Dialogue program at Addis Ababa University Main Campus – Sidist Kilo (hereafter AAU Main Campus) – in Ethiopia during 2009–10, we study the effect of dialogue on the level of trust. Participants were selected using a stratified randomization process. Immediately following the dialogue intervention, an attitudinal survey and a behavioral trust game were conducted among the group of 716 participants and non-participants. We find that the program has a limited but significant positive effect on some dimensions of trusting attitudes. At the same time, participation in the dialogue program increases the sense of importance of ethnic identities as well as the perceptions of discrimination along ethnic lines. Participation in dialogue processes has no significant effect on the game behavior: participants in Sustained Dialogue are neither more trusting nor trustworthy than non-participants. We also find that the selection process matters: some effects occur only in those groups that were nominated to the pool of potential participants from which our randomized selection was made; other effects occur only in the groups of self-selected participants.

The findings, we think, are important from a policy perspective as well as theoretically insightful. That this dialogue process, implemented in a violent and tense context, had a positive effect on trust is a strong argument for its usefulness. The finding that the dialogue process at the same time as it increases trust seems to solidify ethnic identities and perceptions of discrimination is intriguing. It could suggest that while dialogue can be positive for trust it may also increase possible tensions from the perspective of ethnic group-relationships. However, taken together with recent research (Aggestam, 2011), we interpret these findings as signaling that by raising issues of ethnic identity and conflict in the dialogue process, differences and similarities are both recognized and acknowledged, which enhances the participants’ own sense of identity and in turn facilitates their capacity to trust others.

The article will proceed by having the following five sections. First, the previous theoretical discussion on dialogue and peace education in the context of violence and conflict is laid out. Second, the research design of the field experiment in AAU Main Campus in 2009–10 is explained, as well as the survey and Trust Games conducted in May 2010. Third, the findings from the attitudinal survey and Trust Games are presented. Fourth, we discuss how the seeming contradiction between the strengthening of ethnic identities and increased level of trust can be explained; what the difference between the results from the attitude survey and (lacking results) of the Trust Game tells us; how confident we can be in these findings and how robust they are. Last, we draw

\footnote{A possible exception being Biton & Salomon, as discussed later (Biton & Salomon, 2006).}
conclusions from this study and highlight fruitful avenues for further research.

**Dialogue and trust in the context of armed conflict: A theoretical framework**

Peacebuilding in the context of political and ethnic conflict has emerged as an important field of study and international practice (Francis, 2010; Jeong, 2002; Newman & Richmond, 2006). One common initiative for peacebuilding at the grassroots level is the use of dialogue processes to build trust and peaceful relations (Maoz, 2000). Considering the significant amount of resources going into such initiatives from local and international donors, it is critical to evaluate their effectiveness; however, very little systematic research has been done (Malhorta & Liyanage, 2005; Maoz, 2002). The present study begins to fill this gap. Through its research question and design, it taps into three distinct literatures – trust, contact theory, and peace education – all of which will be briefly highlighted in the following.

**Trust**

Trust is essential in peacebuilding processes. The concept of trust is much debated and is central to the study of human interaction in general (Hardin, 1993; Gambetta, 2000) and in relation to peace and conflict processes in particular (Hoffman, 2002; Kydd, 2000a,b; Mitchell, 2000; Stein, 1991). The importance of trust between people for social and economic development is well documented (Johansson-Stenman, Mahmud & Martinsson, 2009).

Varshney (2002) suggests that a key structure of peaceful multi-ethnic societies is their informational networks of civic life, which bring different communities together. Many multi-ethnic communities have been able to resist elites’ provocations, efforts to radicalize followers, and attempts to mobilize support against ‘the other side’. These communities are characterized by intergroup or intercommunal civil society networks, which create interethnic dialogue and increase the level of trust at the local level. By killing rumors, removing misunderstandings, and identifying perpetrators behind misdeeds, highly decentralized tension-managing organizations and social structures protect peace (Varshney, 2002).

**Contact theory**

As with trust, and as part of building trust, contact is seen as essential for peaceful relations. The theoretical origins for dialogue initiatives, the focus of the present study, can perhaps be traced back to Gordon Allport’s major theory on intergroup relations, the contact hypothesis (Allport, 1954). In short, Allport proposed in this theory that prejudice and discrimination between groups experiencing conflict is a result of over-simplification and disinformation. Contact, under certain specified conditions, will bring about positive attitude change; the more people learn about each other, the less prejudice and the more positive interactions there will be. Contact theory has since been subjected to extensive research in a wide range of settings, for example, concerning interracial relations, interethnic relations, segregation in housing, work places, and schools – mostly in relatively peaceful settings and with the vast majority of studies reporting positive contact effects (Malhorta & Liyanage, 2005; Pettigrew, 1998). Yet, herein lies also one significant problem with contact theory: it is so all-encompassing that it cannot be falsified, a problem also known as the ‘independent variable specification problem’ (Pettigrew, 1998). The list of mediating variables for successful outcomes of contact has become so long – more than 50 have been suggested (Forbes, 2004) – that the theory loses its explanatory power: any variable can explain the outcome. Similarly, as the correlations are essentially the same in all cases – more contact leads to less prejudice – there is not the necessary variation needed for a theory to hold as it should also be able to delineate conditions under which contact will not lead to less prejudice (Forbes, 2004), that is, conditions under which it will be falsified. Another problem for the theory and research on contact theory – perhaps the most significant problem of all – is that of selection bias. Those who are willing to participate in interethnic dialogues or encounters are likely those most interested in positive or improved relations (Pettigrew, 1998); at the same time, those who are most prejudiced are likely to avoid contact initiatives; hence, we do not know if contact in itself has a positive effect on attitudes or not.

**Peace education**

The dialogue project at focus in the present article, Sustained Dialogue (SD), builds on the principles of contact theory: through a series of meetings over a period of time, SD aims at transforming relationships between conflicting parties by addressing the underlying issues causing the tensions in order to end conflict and build peace (Saunders, 1999, 2003). In being a systematic dialogue process facilitated by moderators in conflict environments, SD also falls under the conceptual umbrella of peace education – together with many other forms of peace dialogues, encounters, peace camps, and other peace promoting programs. Peace education in the
context of conflict and tension has been characterized as (1) being educationally and psychologically oriented rather than politically oriented, (2) addressing ways of relating to a threatening adversary, (3) focusing on intergroup rather than interpersonal relations, and (4) aiming at changing attitudes towards an adversary in a particular conflict (Salomon & Cairns, 2010).

With roots in the 19th century, peace education has a long history – but, again, there is a paucity of research on its effectiveness (Salomon, 2004; Harris, 2010; Salomon & Cairns, 2010). The studies that do exist are often constrained by the above-mentioned problem of selection bias or other methodological weaknesses. To the best of our knowledge, no previous study has used random selection and both attitudinal and behavioral measures of change for assessing a dialogue project. However, a couple of studies do stand out for having rigorous research designs. In Sri Lanka, Malhorta & Liyanage conducted a study on the long-term impact of a four-day interethnic peace camp program (Malhorta & Liyanage, 2005). Using seemingly comparable participant and control groups (not randomly selected, however) and measuring the dependent variable both attitudinally and behaviorally, positive effects of the peace camp were found on empathy for the other ethnicity at a one-year follow-up (Malhorta & Liyanage, 2005). In one of the few studies using random selection for assessing the effects of participation in a peace education program, Biton & Salomon studied the change in perceptions of peace among Israeli and Palestinian high school students (Biton & Salomon, 2006). The focus of the program was to increase interethnic understanding and tolerance by learning of the other within the setting of one’s own group; face-to-face meetings formed only a minor element towards the end of the program. Thus, while this study does not offer guidance concerning the effect of dialogue and encounter, it is important by way of having a strong research design.

Research design

In this section, we will describe the field experiment, the selection process, the survey and trust game methodology.2

Field experiment: Sustained Dialogue

Field experiments are increasingly being used to identify causal mechanisms and evidence-based knowledge concerning issues related to the political economy of development (Humphreys & Weinstein, 2009), and more recently, also to peacebuilding in developing and post-conflict settings. For example, Fearon and colleagues studied the long-term effect of aid on social cohesion in Liberia (Fearon, Humphreys & Weinstein, 2009). Likewise, Paluck conducted a field experiment on the effect of peace radio broadcasts on interethnic relations in post-genocide Rwanda (Paluck, 2009). In field experiments, the key objective for the researcher is to control the assignment of treatment, so that treated and untreated groups as far as possible are identical except for receiving – or not receiving – treatment (Humphreys & Weinstein, 2009). In this way, valid estimates can be made of causal effects – an otherwise impossible task, as the risk of selection bias, as discussed above, is paramount.

In this study we, in cooperation with our partner organizations, designed and implemented a field experiment in the AAU Main Campus in Ethiopia. Being the oldest and largest higher education institution in the country, with substantial growth in recent years both in areas of study and student intake, the university has a high standing and is a reputable academic environment. At the same time, the Ethiopian case was selected because it represents a country with violent political and ethnic group-based tensions. The program took place on the Sidist Kilo University Campus, which has repeatedly suffered from interethnic tensions and violent incidents. During the time period 2000–10, these include for instance a campus student revolt in 2001 and student protests and ethnic-based clashes in 2004 and 2007. Even during this research project in 2010, there was an outbreak of violence.3 On the basis of these violent and tense group-relationships, a group of university students, members of a Peace Club organization, took the initiative to organize a dialogue program on campus in order to increase tolerance and trust among students of various ethnic, religious, and political backgrounds. Hence, the intervention that is studied here came about and has been driven fully by the initiative and ownership of the local students.

The programme of Sustained Dialogue (hereafter SD) was founded and developed by Harold Saunders, building on his experiences from the Middle East and Tajikistan (Saunders, 1999, 2003). SD is an example of a dialogue process methodology, but there are, of course, several other approaches. SD was chosen as it has been used

---

2 More details of all aspects of the research design can be found in the online appendix.

3 For more information on these recent events of violence, see the news report (http://nazret.com/blog/index.php?title=ethiopia_skirmish_at_aau_sidst_kilo_camp&more=1&c=1&tb=1&ph=1).
previously in many different contexts of group conflicts and political violence. In SD, individuals from various backgrounds are brought together to discuss and find solutions and approaches to pressing issues. Hence, while SD aims at building improved relationships between groups with a history of conflict and tension, it also aims at solving the problem at hand. However, we should point out that in this study the focus lies solely on the first aim, that of influencing relationships. The dependent variable in the present study is attitudes and behavior of participants, not whether these in turn contributed to or influenced the prospects for political change. This is a limitation of the study. However, individual change is the first logical step to examine.

In short, SD is a series of repeated dialogue meetings, conducted with a small group of people (around ten) in which several defined steps are addressed (identification of problem, cost/benefit, scenario building, action, etc.) in discussions led by trained moderators. Thirty-two moderators were trained at a two-day session, 13–15 November 2009, at the Semen Hotel in Addis Ababa. The training included an introduction to the concept of dialogue and skills training in active listening and in moderating meetings.

Randomized selection procedure

One major motivation of this evaluation study was to avoid selection bias by using stratified randomization for the intervention. Randomization is one of the most powerful methodologies to solve the problem of selection bias (Banerjee, 2007; Duflo, Glennerster & Kremer, 2007). In this study, participants were randomly selected by the drawing of lots from two different pools: one consisting of nominated candidates and one consisting of self-selected candidates.5

| Table I. The selection process and total number of survey participants |
|---------------------------|---------------------------|---------------------------|
|                          | Non-SD participants | SD participants | Total |
| Information              | 588                  | 41              | 629   |
| Nomination               | 51                   | 36              | 87    |
| Total                    | 639                  | 77              | 716   |

The pool of self-selected candidates was created through an information session at AAU Main Campus in October 2009, in which information about SD was presented and some social activities arranged. Those who were interested in participating in the SD program were invited to sign up; from this pool a randomized selection into the program could be arranged.

Of the 1,091 students who were present at the information meeting at the AAU Main Campus in October 2009, 629 students signed up their interest to participate in SD. In addition, after the event, moderators nominated potential candidates and thus created a second pool of 351 potential participants. Seventy-five participants from each of the two pools of candidates (self-selected and nominated) were thereafter randomly selected to participate in SD. Due to the considerable time commitment necessary for participating in the SD program during the university semester, there was a quite high drop-out rate in both of the randomly selected groups. In the end, out of the 75 randomly selected students in both groups, 41 self-selected and 36 nominated students participated in SD. Hence, a total of 77 students participated in the SD program. The selection process and final numbers are illustrated in Table I. When the selection process was finalized, the 12–14 session SD program could begin, and took place during December 2009–June 2010.

Outcome measures: Survey and trust game

To assess the outcome of this field experiment, both attitudinal and behavioral measures were employed. A survey questionnaire on trust-related attitudes was carried out on 28 May 2010 (in the last weeks of SD). The survey included questions from the World Values Survey (Inglehart et al., 2004), two surveys conducted at the Haramaya and Jimma Universities respectively (Pact Ethiopia, 2008a,b), a survey study conducted in Rwanda (Brounéus, 2010), and a set of new questions. Of the original 1,091 students (both SD and non-SD participants) who were present at the information meeting and called for the survey, 716 students turned

---

4 In line with this methodology, an initial survey was made before the present study began in order to identify students’ perceptions of pressing issues. Ethnic violence was ranked the most urgent issue on campus (54% of students); the second most pressing issue was religious intolerance at 17%.

5 To clarify, there are two types of randomization: by assignment of the treatment or in terms of sampling (Druckman et al., 2011). In this project, randomization is made in relation to the assignment of the treatment (in this case, the dialogue process), not in terms of a sampling of a general population, a procedure that more accurately replicates the type of recruitment procedures that are typically used in the field. It would be rare to find peace policy programs that select participants through random sampling. Rather, they are typically recruited through an open invitation (in our case, the information meeting) or through networks (in our case, the nomination by the moderators).
up and participated. This response rate of 65% is well above average and should be seen as very good (Baruch & Holtom, 2008). The results of the survey will be presented in the results section below.

On 5 June 2010, when the SD program was coming to its end, a series of trust games were conducted. In recent years, the use of trust games has increased in studies on economic behavior in both developed and developing nations (Danielson & Holm, 2005), but they have hitherto not been used as a behavioral measure for assessing the effect of dialogue processes in settings marked by ethnic tension. The trust game used in the present study is a binary anonymous game involving two players: Player A and Player B. After receiving an initial endowment, X, Player A can send a chosen amount to Player B. The amount Player A sends is multiplied by 3, the sum of which is given to Player B. Player B then decides how much to send back to Player A, after which the game ends. The amount Player A sends is seen as a measure of trust, the amount Player B sends back a measure of reciprocity or trustworthiness (Ashraf, Bohnet & Piankov, 2006). Of the 120 participants that were called to participate in the trust games, 84 came and participated (70% response rate; 46 from the information group, 38 nominated; Table IX). In addition, in order to control for the possibility that trust game behavior is better explained by a person’s degree of risk aversion/risk propensity, we also included a lottery, in which the selection of different strategies before doing the trust game reveals different propensities for taking risks.

Results

Among the 716 students who completed the survey at AAU Main Campus, ages ranged from 18 to 31, with a mean age of 21. The gender distribution was heavily skewed to males; only 11% were women. Concerning religion, 79% identified as (Christian) Orthodox, 9% Muslim, 8% Protestant, and the remaining 4% included traditional or no religious beliefs. When asked about ethnicity, 27% stated they were solely Amhara, 12.5% were Tigre, and 11% were Oromo. Approximately 30% identified as being of mixed ethnic origin (the majority of whom stated they were part-Amhara). More than ten additional ethnic minority groups were reported accounting for the remainder of the population, including, for example, Gurage (2%), Hadiya (1%), Kambata (0.5%), and Sidama (0.5%).

Results: General

In the following, the empirical results of the survey study will be described. First, we will look at the entire survey population as a whole, comparing those who participated in SD with those who did not. Second, we will narrow the focus to the two subgroups of (1) those nominated to participate in SD and (2) those that were self-selected for SD (by signing up at the information meeting) and see whether any differences emerge within these groups, between those who did – or did not – participate in SD. Third, the findings of the trust games will be presented.

Survey on attitudes. To measure attitudes concerning trust in general, two of three questions were posed from the World Values Survey (Inglehart et al., 2004). As demonstrated in Table II, in response to the question ‘In general, would you say that most people can be trusted or that you need to be very careful in dealing with people?’, significantly more SD participants (27%) than non-participants (17%) answered that most people can be trusted (p < .05). For comparison, these levels are lower than those found in, for example, Indonesia, Iran or Denmark, where between 50% and 65% respond that people can be trusted (Inglehart et al., 2004), but higher than levels found in nearby Uganda (8%) (Inglehart et al., 2004). However, interestingly, when asked ‘Do you think most people would try to take advantage of you if they

Table II. In general, would you say that most people can be trusted or that you need to be very careful in dealing with people?

<table>
<thead>
<tr>
<th></th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not trust (6–10)</td>
<td>528 83%</td>
<td>56 73%*</td>
</tr>
<tr>
<td>Trust (1–5)</td>
<td>109 17%</td>
<td>21 27%*</td>
</tr>
<tr>
<td>Total</td>
<td>637 100%</td>
<td>77 100%</td>
</tr>
</tbody>
</table>

* p < .05.

6 The distribution of the categories for the survey participants can be seen in Table I: 629 students from the information meeting (588 non-SD participants and 51 SD participants) and 87 students from the nominated group (51 non-SD participants and 36 SD-participants). Hence, 639 non-SD participants and 77 SD participants were included in the survey.

7 This gender imbalance mirrors that of the 77 SD participants who participated in the survey, of which 14% were women. Possible reasons for this gender disparity and potential effects on the results will be raised in the discussion.

8 The World Values Survey is a global network of social scientists who survey people’s basic values and beliefs worldwide (Inglehart et al., 2004). More information can be found at http://www.worldvaluessurvey.org.
Table III. Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?

<table>
<thead>
<tr>
<th></th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take advantage</td>
<td>243 39%</td>
<td>31 40%</td>
</tr>
<tr>
<td>Try to be fair</td>
<td>388 62%</td>
<td>46 60%</td>
</tr>
<tr>
<td>Total</td>
<td>631 100%</td>
<td>77 100%</td>
</tr>
</tbody>
</table>

Table IV. One should be on guard towards other people, even if they are neighbors or friends

<table>
<thead>
<tr>
<th></th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>On guard</td>
<td>570 94%</td>
<td>65 90%</td>
</tr>
<tr>
<td>Total</td>
<td>604 100%</td>
<td>72 100%</td>
</tr>
</tbody>
</table>

The third statement on trust was taken from a survey study conducted in Bosnia and Herzegovina in 2004, ‘one should be on guard towards other people, even if they are neighbors or friends’, (Magnusson, 2005). As seen in Table IV, many people, both SD participants (90%) and non-participants (94%), agree with this statement, with no significant difference between groups. The percentage found here is much higher than that found in postwar Bosnia and Herzegovina, where the largest group to agree with this statement were Bosniaks at 46% (Magnusson, 2005). Taken together, while these three questions on trust in general suggest that trust between students at the University Campus in Ethiopia is low, SD participants do demonstrate significantly higher levels of trust than non-participants on one measure (Table II, ‘most people can be trusted’).

Based on two Ethiopian surveys conducted at Jimma University and Haramaya University respectively in 2008, questions were also asked to further explore more ethnic dimensions of intergroup trust (Pact Ethiopia, 2008a,b). In Table V, responses to the question ‘while interacting with students from different ethnic background, how do you feel?’ can be seen. The respondents had three different alternatives to choose from (Aggressive, In-between, and Accommodative) and since we are particularly interested in the level of accommodative interaction with other students (as an indicator of trust) we have merged the first two into one category. Among SD participants, 67% report that they feel accommodative in their interaction with other students, slightly yet significantly more than the 59% of non-participants ($p < 0.1$). These levels are both substantially higher than those found at Jimma University (the question was not asked at Haramaya), where 43% agreed to feeling accommodative (Pact Ethiopia, 2008b).

The general findings hitherto support the notion that dialogue increases trust to a certain extent. However, SD participation also seems to have a quite different effect. A seemingly contradictory result is found when asking about the perception of discrimination on the basis of ethnicity. Respondents were asked to react to the claim ‘my ethnic group is discriminated against by others’, by ranking their agreement on a scale from 1 (Not at all) to 10 (Completely). In Table VI, we have grouped these into two different categories. Participation in SD significantly increases the perception of one’s own group’s discrimination. Interestingly, this result can be replicated for the self-selected (information meeting) group, but not the nominated. Our interpretation of this finding is that when people begin to discuss pressing issues, as is the case in SD, they become more aware of the problems at hand. Those individuals who were in the nominated group were probably more aware of ethnically contentious issues already from the beginning and hence

<table>
<thead>
<tr>
<th></th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive</td>
<td>258 41%</td>
<td>25 33%</td>
</tr>
<tr>
<td>Accommodative</td>
<td>365 59%</td>
<td>52 67%</td>
</tr>
<tr>
<td>Total</td>
<td>623 100%</td>
<td>77 100%</td>
</tr>
</tbody>
</table>

$^1 p < 0.1$.
Thus far, we have analyzed the results without taking into account the different selection procedures – nomination by others versus self-selection through signing up at the general information meeting. We will now disaggregate the analysis into these two different subgroups to investigate whether any differences emerge between the SD participation of the subgroups.

**Results: Nominated**

We will begin by looking at the effect of SD on attitudes within the nominated subgroup. As mentioned, of the 87 candidates who were nominated to participate in SD, 36 were randomly selected, by the drawing of lots, to take part in the SD workshop program (Table I). Survey respondents were asked to rank their answer to the question ‘How important is your ethnic belonging to you?’ on a scale from 1 (‘ethnic identity is not important’) to 10 (‘ethnic identity is very important’). Interestingly, as Table VII demonstrates, significantly more nominated SD participants answer that ‘ethnic identity is very important’ than nominees who did not do the SD training program ($p < .05$). In other words, among those who were nominated to participate, the importance of ethnic identity grew by participation in SD. No other questions resulted in significantly different answers between nominees who had and had not participated in SD.

<table>
<thead>
<tr>
<th>Table VII. How important is your ethnic belonging to you?</th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnic belonging not so important (1–6)</td>
<td>40 80%</td>
<td>20 57%*</td>
</tr>
<tr>
<td>Ethnic belonging very important (7–10)</td>
<td>10 20%</td>
<td>15 43%*</td>
</tr>
<tr>
<td>Total</td>
<td>50 100%</td>
<td>35 100%</td>
</tr>
<tr>
<td>* $p &lt; .05.$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thus, we can summarize the results of the attitudinal part of the study. It is important to point out that the effect of SD participation is quite limited: in terms of most of the attitudes examined there were no significant effects from the SD participation. In some specific questions, however, participation in SD did generate a change in attitudes in terms of both increasing trust and decreasing the level of mistrust. At the same time, participation in SD increased the sense of ethnicity and perceived discrimination. Hence, conflictual ethnic relationships were not decreased through SD, but tensions seem in fact to have become heightened: potentially contentious issues were brought to the surface through the dialogue. We can also see that the selection process was important: some different effects can be reported depending on the selection process.

<table>
<thead>
<tr>
<th>Table VIII. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?</th>
<th>Non-SD participants</th>
<th>SD participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most people can be trusted (1–5)</td>
<td>97 17%</td>
<td>11 27%$^1$</td>
</tr>
<tr>
<td>Need to be very careful (6–10)</td>
<td>489 83%</td>
<td>30 73%$^1$</td>
</tr>
<tr>
<td>Total</td>
<td>586 100%</td>
<td>41 100%</td>
</tr>
<tr>
<td>$^1$ $p &lt; 0.1.$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IX. Trust game participants</th>
<th>Non-SD participants</th>
<th>SD participants</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>24</td>
<td>22</td>
<td>46</td>
</tr>
<tr>
<td>Nomination</td>
<td>25</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>35</td>
<td>84</td>
</tr>
</tbody>
</table>

Trust games. As demonstrated in Table IX, of the 84 participants who did the trust game, 35 had participated in Sustained Dialogue (22 were self-selected at the information meeting, 13 had been nominated by others). The remaining 49 trust game participants had not participated in SD, of which 24 were self-selected and 25 had been nominated. Analyses were conducted to determine whether those who had participated in SD played differently in the trust games than those who had not, both with regards to how much they sent – a measure of trust – and how much they returned – a measure of trustworthiness. We also analyzed whether there were differences between those who were nominated and self-selected.
No significant differences emerged in any of the analyses. Hence, while some attitudinal differences were seen as a result of SD, no behavioral trust game differences were observed.

Discussion

The present study is the first to use randomization and trust games for assessing the effect of dialogue programs on trust. We have shown that the dialogue process studied here, Sustained Dialogue, significantly enhanced trust (by decreasing the perception of the need to be on guard and increasing the level of trust in others in general) and created positive relationships with others (participants to a larger extent than non-participants felt accommodative towards students of other ethnicities). By using a robust research design – thereby solving the problem of selection bias which has previously hindered causal analysis – we can with greater confidence state that this dialogue program indeed seems to work. That is, dialogue projects such as Sustained Dialogue seem promising for influencing the parties’ attitudes, even in situations of political violence and ethnic tensions.

That said, it must be emphasized that while there is a positive effect on attitudes, it is small. In general, the effects of the dialogue program involve an increase in trust or decrease in mistrust of no more than a few percentage points. Although there is no benchmark against which we can evaluate these figures, effects are detected in only a few, specific attitudes; in most attitudes there is no difference between the treatment and control groups. \(^{10}\)

This study finds a discrepancy between self-reported attitudes and behavior in the trust games: changes are seen in SD participants’ reported attitudes but not in their trust game behavior. This difference between trust game behavior and reported attitudes is in line with some previous research (Danielson & Holm, 2007; Johansson-Stenman, Mahmud & Martinsson, 2009). There are three possible explanations for this disjunction. A first way of interpreting the difference between attitudinal and behavioral measures concerns incentives for misrepresenting information. Attitudinal surveys can suffer from biases such as acquiescence (‘yah-saying’) and social desirability rating (the subjects’ desire to be seen in a good light). It may be easier to report that you trust someone than to actually put your trust in another person. For this reason, some researchers have argued that trust games are superior to survey questionnaires, since they are incentivized and can to a lesser extent be disregarded as cheap talk (Glaeser et al., 2000). According to this perspective, the attitudinal difference seen in the present study should not be relied upon, since it is not validated by the behavioral measurement, the trust game, which would be considered more credible. However, this view has been disputed in more recent research (Danielson & Holm, 2005).

A second interpretation is that discrepancies occur because surveys and trust games measure different dimensions of trust, with self-reporting attitudinal surveys measuring a broader, more general form of trust and trust games being more focused on monetary aspects of trust (Johansson-Stenman, Mahmud & Martinsson, 2009). The attitudinal trust change observed in SD participants in this study would according to this perspective be genuine but of the more general kind, and less applicable to direct monetary interactions. If this hypothesis is correct, that surveys and trust games measure different dimensions of trust, it is also plausible that the theme of ethnicity – discussed and emphasized in the Sustained Dialogue program – is better reflected in the attitudinal survey, which directly posed questions on ethnicity. Had the trust game included an ethnic dimension (for example telling players the ethnicity of the other) the results may have differed between SD participants and non-participants along similar lines as observed in the attitude survey. \(^{11}\)

A third way of interpreting the difference in results between the attitudinal survey and the trust game is by applying a psychological framework: dialogue programs may influence attitudes first and behavior second. Cognitive behavioral theory and research demonstrate that attitude and behavioral change happen in cyclical sequence, one reinforcing the other (Butler & Wells, 1995; Dobson, 2002). In the present study, the attitudinal change not observed in trust game behavior may be due to this lag between observed attitude change and

\(^{10}\) Danielson & Holm (2007) argue that the student population is not representative and that it may therefore be problematic to generalize from a student population. However, in our study, the student campus is one of the arenas in the Ethiopian society that experiences the political and ethnic violence most acutely and is therefore an important group to study from a peace and conflict research perspective.

\(^{11}\) In the original design of this study, we planned to manipulate the ethnic composition of the groups playing the Trust Game in order to distinguish these two different explanations; however, the logistical challenge turned out to out to be too great. We leave this to future research to explore.
a corresponding change in behavior (and vice versa). Future research may examine if there are longer-term impacts on behavior among those who have gone through the dialogue program. Also, while the SD program focused on interethnic attitudes, no behavior training was involved; this would be an interesting component to develop and include in future studies. With our current research design, we are not able to identify which of the three explanations above (cheap talk, different dimensions of trust, and lag between attitude change and behavioral change) account for the difference between game behavior and the replies in the survey. However, due to inconsistent evidence of the first explanation, we lean in favor of the theoretical and evidence-based foundation of the two latter perspectives. Therefore, we take both the attitudinal and behavioral results at face value: we attribute the change in attitudes to SD and the non-change in behavior as a result of (1) measuring different types of trust, and/or (2) demonstrating a lag between attitude and behavior change.

In general, participants in the trust game – both SD participants and non-participants – did not invest money. Previous research has demonstrated that trust game participants who make low investments may be more risk averse (Holt & Laury, 2002); however, the propensity to take risks was controlled for (in a lottery game) and no correlation was found between lottery game and trust game behavior, so risk aversion cannot help to account for the results (Eckel & Wilson, 2004). In future research, it would be interesting to study the role of contextual factors such as poverty in trust game behavior. Perhaps perceived poverty level could better explain low risk trust game behavior in developing contexts; lottery games may measure another type of risk taking behavior.

Another major finding in this study is that the dialogue program enhanced ethnic awareness and self-identification. This may seem counter-intuitive. However, the dialogue program was initiated by the students based on the premise that there was a problem with ethnic intolerance and that most (if not all) students were affected by this problem – this presumption provided the necessary legitimacy for students to feel comfortable to start voicing their concerns. Indeed, in the original survey before the program, ethnic violence was identified as the most problematic concern, and formed the focal point for the discussions during the dialogue sessions. Interestingly, the finding that SD created positive relationships with others and simultaneously raised ethnic awareness is in line with Nannestad et al.’s study on non-Western immigrants in Denmark, where a positive relationship was found between so called bridging social capital, which connects individuals across ethnic groups, and bonding social capital, which closely binds individuals to their own ethnic group (Nannestad, Svendsen & Svendsen, 2008).

This study also finds an interesting difference in the outcome of the dialogue program with regards to the selection process. Recall that the study’s unique research design enabled us to examine two processes of selection by creating two pools from which randomization was made: nomination and self-selection. The respondents’ sense of importance of their own ethnic identity was enhanced by the Sustained Dialogue programs, but only among the group that was nominated to the program (see Table VII). In other words, dialogue programs do not seem to generate an increased sense of ethnic identity in a population that is recruited into a program through information and self-selection. The identity-enhancement effect would occur primarily among individuals that are considered to be specially suited for dialogue programs. Future recruitment strategies into dialogue programs may need to take this into account.

The selection process also has implications for our ability to generalize results. Since we do not randomly sample the participants to the program, we cannot say anything about the general effect on the population of dialogue programs such as the one studied here. We can only generalize to the group of selected participants – those that have been identified as suitable candidates or those that have explicitly shown interest in participating by coming to the information meeting, which was the starting point for the program. Although the generalizability is less than would have been the case with a randomized sample, this set-up also has certain advantages. In particular, it resembles the design of most dialogue programs in the field: a randomized selection would be an atypical selection process for this type of program. Typically, dialogue programs select their participants through open calls and self-selection, or through nominations. Hence, with this study, we draw conclusions about the effect of dialogue programs for exactly these types of selection procedures.

The methodology of field experiment has been growing, but has had only limited application to peace and conflict research. Because the main dependent variable (armed conflict) is both rare and inherently difficult to

---

12 This is one of two measurements of enhanced ethnic identity; the other measurement (see Table VI) was on the sense of ethnic discrimination, which was enhanced in all SD participants.
predict, randomized field experiments have had limited applicability, particularly in comparison with the field of developmental economics. Perhaps the profound ethical considerations that need to be made before conducting a field experiment in peace and conflict research have also been an impediment for wider use of the methodology. Yet, this study has shown that there are possibilities to use this methodology in a fruitful way in order to enhance our understanding of ethnic identities in conflicts and shed new light on theoretical debates on dialogue as way of promoting interethnic trust.

Three limitations should be mentioned as caveats. First, one concern in the present study is the skewed gender distribution: only 14% of SD participants were women (11% women in entire study population). It is likely that the results would have been different if more women had been included (Croson & Gneezy, 2009; Hannagan & Larimer, 2010). Men have dominated in the violent group-based clashes on campus. In the study, men and women also report differently on the experience of violence. One reason for so few women being included might be found in the selection process: very few women came to the initial information meeting where interest in participation could be expressed. More women were included through the nomination process, but this did not do enough to address the skewed gender composition. This gender imbalance is an important limitation of the present study and demonstrates a gap that merits further research. In future study, a refined design to ensure a broader inclusion of both women and men would be essential.

Second, the research design used in this study is able to only capture some of the changes that Sustained Dialogue programs may produce among participants—trust is only one possible outcome of the program. Further, just as trust is one of many possible outcomes of the program, how we measure trust will capture some aspects of trust and not others. However, considering the issues at stake (for example, time, resources, and hope in post-conflict or conflict settings) we believe that it is better to study one aspect thoroughly and with scientific rigor, than not at all—only in this way can we begin building evidence-based knowledge devoid of the problem of selection bias and really learn how dialogue can best benefit peacebuilding.

Third, the measurement of ‘success’ in this study is focused on change in attitude in individuals; individuals are our basic unit of analysis. Violence and ethnic tensions are, however, social phenomena and the ultimate test of dialogue programs in the context of conflict is therefore whether and to what extent individual attitudes, preferences, and behavior aggregate in the social interaction in a way that creates peaceful interactions with other groups or, alternatively, tensions, threats, and violence. The development of social measurement indicators of the impact of dialogue programs is pivotal for future research.

Conclusions

This study has examined the effect of dialogue processes in the context of political and ethnic tensions and violence. Three conclusions can be drawn. First, this study yields support for the implementation of dialogue processes in the context of political conflict and ethnic violence. Even if both perceptions of ethnic discrimination and sense of ethnic identity are enhanced—perceptions that can increase the risk for tension and conflict—these perceptions were coupled with increased trust in others. Hence, this study suggests that there is something basically positive about raising concerns and airing issues of contentions, and that there is nothing inherently wrong with an increased sense of ethnic identity as long as this does not influence negatively the perceptions of others. Dialogue programs, such as Sustained Dialogue, have a positive effect on participants’ attitudes: this program works by decreasing mistrust and increasing the level of trust between people of different ethnic origins.

Second, this study suggests that participation in dialogue programs is not sufficient to alter the behavioral aspects. There may be several different reasons for this difference, for example time-lag and different dimensions of trust, which we have discussed at some length above. Hence, this study indicates a need to develop more behavior-oriented dialogue programs, which may be in a better position to create change in how individuals act toward each other. Most dialogue programs seem to be oriented towards verbal communication in the form of talking and listening; fewer programs are devoted to developing skills of cooperative behavior. This study gives some evidence for the need to develop programs with such a behavioral orientation.

Third and more generally, this study shows the fruitfulness of randomized field-experiments in the area of peace and conflict research. There are many challenges and obstacles of a practical, ethical or political nature to overcome in order to implement a robust research design of this kind for studying the impact of dialogue and other forms of peacebuilding programs. However,
when it can be done, this study shows that it can generate new findings and insights, which can guide theory and policy. To some extent, field experiments represent a new way of doing research, which is collaborative between research and practice. This study represents one effort to develop this approach.

**Replication data**

The data used in this article can be found at www.prio.no/jpr/datasets. An online appendix with further details on, for example, the selection process and questionnaire can also be found there.

**Acknowledgements**

We wish to thank our partners at the Life and Peace Institute in Addis Ababa and Uppsala, and Addis Ababa University for great collaboration. In particular, we are indebted to Hannah Tsadik, Habtamu Wondimun, Shane Quinn, Jody Henderson, Amanda Ree, Habtamu Wondimun, Yonas Adeto, Tarekegn Adebo, Kassahun Habtamu, Dalaya Ashenafi, and Demelash Fesehagiorgis Abdi. We also thank Mats Hammarström, Kristine Eck, Jakob Svensson, John Gray, and the late Dorothy Brown for advice and support. Three JPR reviewers gave excellent advice on an earlier draft. Finally, we are grateful to all of the students who participated in this study—we feel deep respect for your dedication and work for peace.

**References**


Pact Ethiopia (2008a) Haramaya University Campus Youth Survey.

Pact Ethiopia (2008b) Jimma University Campus Youth Survey.


