

EVALUATING PEACEBUILDING INITIATIVES USING MULTIPLE
METHODOLOGIES: LESSONS LEARNED FROM A GREEK-TURKISH PEACE
EDUCATION

by
GENCO ORKUN GENC

Submitted to the Faculty of Arts and Social Sciences
in partial fulfillment of
the requirements for the degree of
Master of Arts

Sabancı University
Spring 2006

© Genco Orkun Genç 2006

All Rights Reserved

EVALUATING PEACEBUILDING INITIATIVES USING MULTIPLE
METHODOLOGIES: LESSONS LEARNED FROM A GREEK-TURKISH PEACE
EDUCATION INITIATIVE

Genco Orkun Genç

Conflict Analysis and Resolution, MA, Thesis, 2006

Thesis Supervisor: Esra Çuhadar Gürkaynak

Keywords: Contact Hypothesis, Peace Education, Evaluation

This paper discusses the results of the evaluation of a Greek-Turkish peace education initiative at the grass-roots level titled Learning Conflict Resolution and Producing Peace: Turkish-Greek Civic Dialogue (TGCD). The purpose of the initiative was to build and improve relations among Turkish and Greek youth. TGCD incorporated contact and peace education as the primary tools for change. This research evaluated the program by combining a two-way evaluation methodology. The first part investigated the program's theory of change through structured interviews with the organizers and participant observation. A process map has been created as a result of this. In the second part, we conduct an experiment involving the treatment group and a control group to assess the outcomes from the workshop at the inter-personal level. We measure the sustainability of three major traits in the treatment group: attitudinal empathy, behavioral empathy, and trust. The findings of the experiment suggest that there are significant differences between the treatment group and the control group with regard to the development of attitudinal empathy and trust. Finally, we compare the results from the mapping of the program's theory of change and the findings from the experimental design. This study contributes to the literature at large in a sense that it assesses and tests a program's theory of change with multiple methodologies using qualitative interviews, mapping, and a field experiment.

BARIŞ GİRİŞİMLERİNİN ÇOKLU METODOLOJİ İLE DEĞERLENDİRİLMESİ: BİR TÜRK-YUNAN BARIŞ EĞİTİMİ GİRİŞİMİNDEN ÇIKAN DERSLER

Genco Orkun Genç

Uyuşmazlık Analizi ve Çözümü, MA, Tez, 2006-07-06

Tez Danışmanı: Esra Çuhadar Gürkaynak

Anahtar Kelimeler: Kontak Hipotezi, Barış Eğitimi, Değerlendirme

Bu çalışma gençlik seviyesinde gerçekleştirilen bir Türk-Yunan barış eğitimi girişiminin sonuçlarını tartışmaktadır. Değerlendirmeye tabi tutulan projenin ismi Uyuşmazlık Çözümünü Öğrenmek ve Barış Yaratmak'tır. Projenin temel amacı Türk ve Yunan gençleri arasında bir ilişki kurmak ve kurulan ilişkilerin iyileştirilmesidir. Bu proje planlanan değişikliklere ulaşmak için iki ana araç kullanmıştır: barış eğitimi ve kontak. Bu araştırma projeyi iki yönlü bir değerlendirme metoduyla incelemiştir. İlk kısımda röportajlar ve katılımcı gözlem kullanılarak projenin değişim teorileri incelenmiştir. Bu değerlendirmenin sonucu olarak bir süreç haritası elde edilmiştir. İkinci kısımda ise uygulama ve kontrol grubu ile birlikte bir deney yapılmıştır. Bu değerlendirmenin sebebi ise çalıştayın kişisel seviyedeki etkilerini ölçmek olmuştur. Bu deneyde üç temel duruma bakılmıştır: düşünsel empati, davranışsal empati ve güven. Bulgulara göre iki grup arasında düşünsel empati ve güven seviyelerinde ciddi bir fark vardır. Son olarak ise, deney sonuçları ile süreç haritası karşılaştırılarak bir analiz yapılmıştır. Bu çalışmanın disipline katkısı ise değişik metodolojileri bir araya getirerek harmanlamasıdır.

ACKNOWLEDGEMENTS

This thesis has been an incredible journey from me and the most important thing is that I have never felt alone. How could I? First of all, there is Nilgün, my mom, my inspiration, role-model, basically my everything. She encouraged me to take the giant step and made me believe that I can write this thesis: against all odds...Of course my mom was not alone; there are my sister and my brother that completed my home.

Then there is my guarding angel, who stood with me, held my hand and reminded me that everything is OK and will be OK now and for the rest of my life..

There is another hero in this battle of mine: my roomie Evren Bey. I know it must be the hardest thing to put up with an angry room-mate struggling with a thesis, but I'll be there for you as you've been there for me too..

How can I even forget my best friend Naz, who escorted me during those long nights where the sleep went astray.

There have been a lot of times where I felt lost or beat but there she was: my Super supervisor Esra Cuhadar Gürkaynak. It has been a pleasure working with you. As for Benjamin Broome and Nimet Beriker, it has been a privilege to have you in my committee..

My co-fighters Andreas and Katharina: Guys, our names will always be remembered along the corridors of the Sabanci building, we have been a hell of a team..

I also need to give my thanks to Istanbul Policy Center and ELIAMEP for letting me in their project and allow me to evaluate their project.

A big bunch of thanks goes to my friends from the TGCD project: Devrimsel, Yorgos, Aspurçe, Ayşe, Ria, Tassos and all of the rest..

Last but not the least, there are two very special people that I would like to remember in these lines.. Dad, you always wanted me to go and learn how to survive in the big Istanbul,

I guess I did it and I know that you are proud of me with that unique smile on your face, I miss you..

And my Grandma, who silently passed away this year.. You have always been my biggest fan, my most fierce supporter. I have been the luckiest grandchild to have so much love. I still feel your love.

I love both of you and yes, this degree is for you...

Annem Nilgün Tez'e...

TABLE OF CONTENTS

	Page
Chapter One: Introduction	10
Chapter Two: Literature Review.....	13
2.1. Intergroup Contact Theory.....	15
2.2. Peace Education	21
2.3. Evaluation Research.....	25
2.3.1. Defining Evaluation	25
2.3.2. Incentives for Evaluation.....	26
2.3.3. Different Approaches to Evaluation.....	27
2.3.4. Challenges in Evaluation Research.....	29
2.3.5. Toolkit for Evaluators	32
2.3.6. Theories of Change	33
Chapter Three: Research Design.....	36
3.1.Participants.....	36
3.2.Methodology	40
3.2.1. . Program Evaluation.....	41
3.2.2. . The Experiment.....	43
3.3.Data Gathering	46
3.4.Data Analysis	49
3.5.Limitations in Research Design	49
Chapter Four: Findings	51
4.1.Summative Evaluation.....	51
4.1.1. Attitudinal Trust.....	51
4.1.2. Attitudinal Empathy	57
4.1.3. Behavioral Empathy	58
4.2.Process Evaluation	62
Chapter Five: Conclusion	71
5.1.Discussion of the Findings.....	71
5.2.Recommendations	76
5.3.Future Research.....	77
Bibliography.....	80
List of Tables	82

Appendix A: The Program of the TGCD	83
Appendix B: Final Interview Protocol for Organizers	88
Appendix C: Questionnaire.....	89

1. INTRODUCTION

It all started with this single question: Did it work?

It was only a week before I have started the MA program in Conflict Analysis and Resolution (CA&R) in Sabanci University when I received an e-mail from one of my professors informing the prospective students about a workshop on Turkish-Greek relations with an emphasis on Peace and Conflict Resolution. It was a Greek-Turkish peace education initiative at the grass-roots level titled: Learning Conflict Resolution and Producing Peace: Turkish-Greek Civic Dialogue (TGCD). The purpose of the initiative was to build and to improve relations among Turkish and Greek youth. I thought it was a brilliant chance for me to combine the theoretical framework-that I would get from the program-with a field experience. So, I applied right away.

As soon as the project took off, I realized that I had already started to evaluate the program on my mind based on my very basic knowledge on CA&R only after two months of classes. I had received the theoretical framework for the primary tools of change that the TGCD incorporated, namely peace education and contact, and it was a real opportunity to see the practical implications of these concepts. As the project progressed along with my MA program, I made up my mind: I was going to focus on Greek-Turkish relations. Next stop was only after the end of the first year of the MA program when I discovered my specific interest on evaluation research. After seeing the negative effects of poorly designed interventions in the field, I found a way to combine my two interests together.

Thus, the simple question in the beginning, asking whether the program actually worked, is transformed to be appropriate for a scientific research question:

How effective was the initiative of 'Turkish Greek Civic Dialogue: Learning Conflict Resolution and Producing Peace' in improving relations among its participants? Is it possible to sustain those improved relations after a year?

This research evaluated the program by combining a two-way evaluation methodology. The first part investigated the program's theory of change through structured interviews with the organizers and participant observation. A process map has been created as a result of it.

In the second part, I conducted an experiment involving the treatment group and a control group to assess the outcomes of the workshop at the inter-personal level. I measured

the sustainability of three major traits in the treatment group: attitudinal empathy, behavioral empathy, and trust.

Finally, I compared the results from the mapping of the program's theory of change with the findings from the experimental design. This study contributes to the literature at large in a sense that it assesses and tests a program's theory of change with multiple methodologies using qualitative interviews, participant observation, mapping, and a field experiment.

Turkish-Greek Civic Dialogue

“Learning Conflict Resolution and Producing Peace: Turkish-Greek Civic Dialogue” (TGCD) was organized jointly by two prominent think-tanks in Turkey and Greece: respectively IPC and ELIAMEP. Funded by the grant of European Representation to Turkey, the Project brought fifteen Turkish and fifteen Greek students together for three meetings.

The first meeting took place in Istanbul between 5th and 7th of November, 2004. During this meeting, students had the chance to meet with their counterparts from other national group and provided by a series of seminars on theories of peace and conflict together with the history of Turkish-Greek relations. The significance of this initial meeting was that the students were given a list of issues between Turkey and Greece and they were encouraged to select the one topic they would like to work on. Later on, the organizers reviewed the requests of the participants and came up with teams comprising students from both ethnicities. The goal was that these teams would work together throughout the project and produce a paper for the final conference and present it.

The second meeting took place in Athens between 3rd and 6th of December, 2004. Throughout this meeting, the students continued to work with their team members on the final assignment and were provided assistance by the members of the organizing committee. The topics of the seminars in the second meeting were more policy oriented compared to the theoretical focus on the first meeting.

The students continued corresponding with their team members and the organizers during the gap between the second and third meeting and rehearsed their parts of the presentation in their own countries under the supervision of the organizers.

The third, and final, meeting took place in Athens between 18th and 20th of February, 2005 and was dedicated to the student presentations that were conducted as a panel session. Once a group finishes its presentation, the ground was open for debate and constructive

criticisms from fellow students as well as the organizers and some professors attending to these meetings.

The project ended by the submission of the final papers to the local organizers by the early March, 2005.

2. LITERATURE REVIEW

It has been mentioned in the introduction chapter that this research aims to explore the following question: How effective was the initiative of ‘Turkish Greek Civic Dialogue: Learning Conflict Resolution and Producing Peace’ in improving relations among its participants? Is it possible to sustain those improved relations after a year? The research question falls under the domain of three major theories; (1) Intergroup Contact Theory, (2) Peace Education Literature, and (3) Evaluation Research. It is essential to explain what the organizers expected as a change and how these expectations were met in the final outcome. In the mission statement of the project, the main goal of the organizers is described as ‘creating a contact group with the aim of prejudice reduction’ among Turkish and Greek graduate students. With this description, as a starting point, the relevant literature on intergroup contact theory in relation to prejudice reduction will provide a grid to evaluate the process itself and set the criteria in evaluation of TGCD.

It is important to indicate once again that this thesis not only focuses on the outcomes of the initiative into creating the desired change, but also the process of achieving the goals. Therefore, the evaluation of the initiative has two dimensions. The first one is process oriented, which explores the program’s theories of change by using qualitative interviews and participant observation. The second dimension is yet outcome oriented where the levels of attitudinal empathy and trust in addition to the behavioral empathy was measured by a quasi-experiment. During the process evaluation, intergroup contact theory and peace education was outlined by the organizers as the two major tools in achieving the program goals. Previously efforts have concentrated on merely bringing together the conflict parties with the fairly undefined hope that such contact would facilitate resolution.(Allport, 1954; Cook, 1978; Pettigrew, 1982) However, in light of the relative lack of success these efforts have achieved, a new line of theorists such as Anna Ohanyan and John. E. Lewis (2005) has offered a somewhat more radical approach in which peace education is a viable part of the process. In other words, no one leaves the table without being exposed to this component of the initiative. The aim of the organizers in combining contact with peace education and the reflections on this method in the literature was analyzed as well in detail in section 2.2.

The final dimension of the literature review section includes an overview of *Program Evaluation* and will summarize why there is a need for such evaluations. In this part, discussions on different types of evaluations, and especially evaluation of a program’s theory

of change are explained. Program Evaluation is a tool that is widely used in this research especially while analyzing the first part of the research question on the effectiveness of the project.

2.1. Intergroup Contact Theory

The Turkish-Greek Civic Dialogue (TGCD) started with the assumption that the role of prejudice is significant in shaping the Turkish-Greek conflict and the current era of détente. Thus, to the organizers the remedy to tackle this obstacle in the bilateral relations was considered to be ‘contact’, which was to be the most dominant tool used so far in the specific Turkish-Greek conflict. In this section, the relationship between prejudice and contact is described via reference to the debate in the literature starting from the late 40s.

Although a huge body of literature, mainly within social psychology, on prejudice exists, the aim of this thesis is not to examine the concept itself. An operating definition is adequate in moving forward. Thomas Pettigrew (1982) defines prejudice as:

...an antipathy accompanied by a faulty generalization. It may be felt or expressed. It may be directed toward a group as a whole or toward an individual because he is not a member of that group. Thus, prejudice simultaneously violates two basic norms – the norm of rationality and the norm of human-heartedness. (pg. 3)

Prejudice was a defining feature of the situation that many groups were facing in the United States especially after the World War II. Among these groups, the Black, Indian and Asians were subject to segregation and discrimination due to the prejudices by White Americans. These circumstances brought forward the need to come up with a remedy for the exacerbating situation.

The *human relations* movement that unfolded in the United States in the aftermath of World War II is thus considered to be the primary stimulus for the contact hypothesis theory. (Pettigrew, 1986) The first social scientists that worked on theorizing the intergroup contact were R.M. Williams, Jr. and G. Watson in 1947 and they identified education as the remedy to tackle the root causes of prejudice and ignorance. Pettigrew links this starting point of the theory to the premise that group members who have more contact with the out-group are less likely to stereotype them. (Pettigrew, 1982)

The problem with theorizing contact was the strong expectation of the concept as the remedy itself to the rising prejudice among certain groups in the United States followed by a disappointment resulting from some experiments. With the rise of negative results coming out of contact situations, it became obvious that contact does not serve the same positive results under every circumstance. This understanding is how the acclaimed ‘contact hypothesis’ of Gordon Allport emerged. Allport (1954) proposed four ‘optimal conditions’ in order for contact to create positive results in intergroup settings. These are:

- *Equal Group Status*
- *Common Goals*
- *Intergroup Cooperation*
- *Support of Authorities, Law, or Custom* (Allport as cited in Pettigrew, 1998)

By setting these conditions for successful contact, Allport became one of the most prominent figures in the literature by establishing the fundamental tenets of contact theory. His hypothesis has been the focal point of debate with a competing line of researchers arguing for and against the validity of this approach. A line of researchers such as Cook 1978; Sherif 1966; Smith 1994; Powers & Ellison 1995 tested the validity of these conditions, either by taking the whole set or a single one into consideration, and came up with results that corroborate the importance of these conditions. Another line of researchers in Europe and the Middle East elaborated on new conditions such as active participation (Maoz 2005), common-language, voluntary contact, prosperous economy (Wagner & Machleit 1986), positive initial views (Ben-Ari & Amir 1986) to be added to list provided by Allport. Following the expansion of conditions, Pettigrew criticized the contact literature following the work of Allport and compared the recent literature to “an open-ended laundry list of conditions.” (Pettigrew, 1986) In his later works, Pettigrew expanded upon this problem, calling it the *independent variable specification problem*, where he argues that the writers overburden the hypothesis by suggesting alternative conditions suitable in different settings all over the world. Pettigrew points out that these additional conditions are not essential for the intergroup contact, yet essential for facilitation in specific settings. (Pettigrew, 1998) While Allport's work provides the 'skeleton', the framework for which modern day theory rests, Pettigrew has supplied the 'flesh', by pruning the massive amount of literature, and reducing the theory to its more essential nature.

This part of the literature contributed to this study by providing the necessary tools to evaluate the goals of the organizers in the body of TGCD. There is a need to discover the perception of the organizing team itself on the contact hypothesis and their understanding of optimal conditions leading to a successful intergroup contact. Later on, while mapping out the TGCD'S theories of change, these conditions were used to clarify the mechanisms of change and correlate them with the program activities.

There is yet another dimension of the 'Intergroup Contact Theory' that was useful in evaluating the effectiveness of the micro objectives of the project, which has actually arisen from an accumulation of criticisms. As Pettigrew states: “The original hypothesis ...predicts

only *when* contact will lead to positive change, not *how* and *why* the change occurs. A broader theory of intergroup contact requires an explicit specification of the process involved” (Pettigrew, 1998, p. 70). The need to put emphasis on process rather than simply measuring the outcome is now being supported by a recent line of researchers. (See the works of Bar and Bargal 1995; Maoz 2000; Salomon 2002)

As a consequence, Pettigrew (1998) suggested that intergroup contact sets into motion four kinds of processes of change. This idea of four different processes operating through a contact situation emerged as a result of the debate between the initial supporters of the contact hypothesis and the contrary views of the cognitive researchers. The initial theory was that as the subjects of the intergroup contact gain new information about the out group, their negative views would be altered. Thus, contact would be effective in reducing prejudice. On the other hand, the cognitive researchers found that learning is limited when faced with material that contradicts the attitudes and stereotypes held by the in-group. Pettigrew then raised the question of how contact still seems to be working under these circumstances. He answers his own question by referring to this four-step model of change for an intergroup contact subject

1. *Learning about the Outgroup:* The debate mentioned above is perceived by Pettigrew as inadequate since learning is only the first step in the process of change; learning does not simply by itself produce successful outcomes.
2. *Changing Behavior:* A change in behavior is argued to be the first step towards attitudinal change. When new situations formed in the intergroup setting alter the expectations of the ingroup in a way to accept the outgroup, this change in behavior can lead to attitudinal change. Jackman and Crane also suggest that “This behavioral process also benefits from repeated contact, preferably in varied settings.” (Jackman & Crane cited in Pettigrew, 1998)
3. *Generating Affective Ties:* In this process, the importance of acknowledging the presence of anxiety and its possible consequence in leading to negative reactions is stressed. The role of positive emotions, and even intergroup friendship, is described to be a crucial process in the contact situations. (See also the works of Amir 1976; Oliner 1988; Rippl 1995 on the role of intimacy)
4. *Ingroup Reappraisal:* In this step of the contact process, Pettigrew suggests that as the bond between self and the other is formed, the ingroup would start to question its own knowledge prior to the contact. Pettigrew further

claims that the emergence of outgroup friends leads to less national pride. (Pettigrew, 1997) Another supporting suggestion comes from Mullen, arguing that “less ingroup contact leads to bias toward the outgroup”. (Mullen cited in Pettigrew, 1998)

In addition to these, Pettigrew proposes a fifth essential condition to be added to the list created by Gordon Allport. This fifth condition is the ability of the contact group to provide its participants with an opportunity to form strong affective ties, namely a friendly setting. Pettigrew claims that cross-group friendship is a key factor in reducing prejudice and, therefore, must be included in the generic framework of contact hypothesis rather than a facilitation tip. (Pettigrew, 1998)

Different from this traditional line of research on contact, Anna Ohanian and John E. Lewis (2005) raised another issue with the interethnic contact in Georgian-Abkhaz Peace Camp. They challenged the traditional line of research on contact which sees the role of attitudinal change as a prerequisite for successful contact outcome. These authors argued that the importance of prejudice reduction, tackling ignorance and the lack of knowledge about the outgroup, before any attempts for future collaboration is overemphasized. They vaguely claim that prejudice reduction should not be the only goal of contact but rather suggest willingness for future cooperation as an alternative and yet independent goal of contact. This point is demonstrated as a result of the research in their own words:

...Willingness to cooperate between the two groups materialized without the interpersonal attitudinal change serving as a primary catalyst or a precondition....The overemphasis on attitudinal changes at the expense of willingness to cooperate is a limitation that the relevant literature on contact hypothesis needs to overcome. (Ohanian and Lewis, 2005, p. 76)

Ohanian and Lewis in the same work move forward to criticize the bottom-up approach in theorizing contact, which they perceive to be the dominant approach in the literature. The bottom-up approach in contact suggests that the outcomes of contact are gradual and attitudinal change among the recipients of contact is a prerequisite without moving any forward with any behavioral change. According to Ohanian and Lewis, this approach overemphasizes the micro level variables such as the individual characteristics of the program participants. Even though they do not mention the work conducted by Pettigrew (1998), it seems that the authors are against including the fifth condition that Pettigrew suggested, the opportunity for cross-group friendship, based on its origins in interpersonal relationship. Instead of the bottom-up approach, Ohanian suggests a top-down theorization, where the attitudinal change is not a necessary condition to achieve future collaboration. In

this approach, once the recipients of contact understand the need for future collaboration and grow a will for this, attitudinal changes will flourish eventually. (Ohanyan, 2005)

Ohanyan's last point brings the case study of TGCD back to focus. Since the effectiveness of the project in creating change among its participants was measured, the scientific quality of the work carried out by the organizers should be investigated. Thus the importance of evaluating contact lay in two criteria: a.) if the goals are process or outcome oriented, and b.) if their goals are to achieve future cooperation or just pure attitudinal change that would eventually lead to future cooperation. This thesis therefore focuses only on the second question where the expectations of the organizers are analyzed. In addition to this, the experiment presented in Section 3.2 will reveal more on the attitudinal and behavioral changes among the participants of TGCD.

Among these attitudinal changes, an operating definition of attitudinal trust needs to be provided in order to clarify the findings in Section 4.1.1. Lewicki and Stevenson (1999) define trust as "an individual's belief in, and willingness to act on the basis of, the words, actions, and decisions of another." (Lewicki, Saunders and Minton, 1999, pg. 711) Lewicki and Stevenson suggest that there are three different types of trust: calculus-based trust, knowledge-based trust, and identification-based trust. The authors also divide their trust scale, which was used for preparing the questionnaires in this research, in accordance with these different levels of trust.

"Calculus-based trust is based on consistency of behavior that people will do what they say they are going to do" (Lewicki, Saunders and Minton, 1999, p. 711). This type of trust is based on a cost-benefit analysis between creating and sustaining a relationship, and maintaining or severing it.

"Knowledge-based trust is grounded in other's predictability; knowing the other sufficiently well so that the other's behavior can be anticipated...In knowledge-based trust regular communication and courtship are key processes" (Lewicki, Saunders and Minton, 1999, p. 712). Under this light of the defining features of knowledge-based trust, the levels of knowledge-based trust within the experimental group should be higher than the control group, who did not receive contact as the treatment.

"Identification-based trust is based on a complete empathy with or identification with the other party's desires and intentions...One comes to learn what really matter to the other, and comes to place the same importance on those behaviors as the other does" (Lewicki, Saunders and Minton, 1999, p. 712).

Lewicki and Stevenson underline also four major characteristics of this trust scale. They believe that these different types of trust are sequentially linked, calculus-based trust develop first, followed by knowledge-based trust and finally identification based trust. They also highlight the fact that most relationships don't pass the lines of calculus-based trust and identification-based trust is acquired over a long time period. They also mention that while trust develops slowly over time, it can rapidly decline in cases of trust violations and repairing violated and broken trust is a very complex, difficult process. (Lewicki, Saunders and Minton, 1999, p. 713)

2.2. Peace Education

Peace education is the second major tool used by the organizers of TGCD. The project brought together a list of acclaimed academicians and several policy makers in order to help the student participants learn more about peace and conflict resolution. These series of seminars took place only in the first two meetings: one in Istanbul and the other in Athens. After providing a short summary of TGCD's peace education plan, the major lines of debate in peace education literature relevant to this study will be presented.

In the first meeting that took place on November, 2004, the participants received training mainly on some theoretical concepts such as *War and Conflict in History*, *Three Approaches to Peaceful Settlement*, as well as some practice-oriented sessions like *Conflict Resolution Toolbox* and *Mediation Process*. In the second meeting, that took place in Athens on December, 2004, the variety of topics presented to the participants was richer and more specific in nature. The program included interactive panels on topics such as: *The Role of Public Opinion in Greek-Turkish Relations*, *Greek-Turkish Economic Relations*, and *the Role of European Union in Transformation of Greek-Turkish Relations*. The participants were assisted by their advisors assigned by the organizers, a list of recommended readings, and the prominent academicians attending the meetings on their final throughout the whole process. The final papers of the participants were presented to public with a conference that took place in Athens on February, 2005. The whole program of three meetings of the TGCD can be found in Appendix I.

A broad definition of the concept of peace education is provided by Harris and Morrison. (2003) "Peace education is currently considered to be both a philosophy and a process involving skills, including listening, reflection, problem-solving, cooperation and conflict resolution. The process involves empowering people with the skills, attitudes and knowledge to create a safe world and build a sustainable environment" (Harris, Morrison, 2003, p. 9). This definition almost fits the understanding of TGCD especially in the sense that the project had the aim to empower the students with conflict resolution skills, which would lead to improved relations and change in attitudes of trust and empathy. Although, the literature in peace education is extensive, it mostly deals with students of younger age rather than the older aged participants of TGCD. Peace education, in general, presents a variety of notions such as security and peace, differing religious traditions, cultural values, and linguistic concepts. (Harris and Morrison, 2003) Despite this diversity of practice in peace education, the common denominator lies in teaching the root causes of conflict and presenting

alternatives to violence to students of all age, gender, or race. Among different types of peace education, Harris and Morrison outline five: *human rights education, environmental education, international education, conflict resolution education, and development education*. Among these five, the only suitable type of peace education to TGCD is conflict resolution education, therefore, while discussing peace education from now on I will only be mentioning this type among the other five.

Under the title of conflict resolution education, the authors of the book 'Peace Education', Harris and Morrison, state ten main goals that an initiative has to achieve. These are:

(1) to appreciate the richness of the concept of peace, (2) to address fears, (3) to provide information about security, (4) to understand war behavior, (5) to develop intercultural understanding, (6) to provide a future orientation, (7) to teach peace as a process, (8) to promote a concept of peace accompanied by social justice, (9) to stimulate a respect for life, and (10) to manage conflicts nonviolently (Harris and Morrison, 2003, p. 32).

Even though the TGCD did not have a structured peace education design, it managed to provide the students with a variety of speakers, expert on their areas of interest, which was the major tool of the projects peace education design. In order to be more specific, the TGCD was able to address only four of the goals stated above: numbers 2, 4, 6, and 10 in particular. The lectures on *Greek-Turkish Relations: Past and Present* and *EU's role in transforming the bilateral relations* along with the discussion sessions on sensitive issues enabled the participants of the TGCD to address their fears. (Number 2) In terms of understanding war behavior (number 4), there were two lectures: *War and Conflict in History* and *Conflict and National Identity*. Another generic goal of peace education that the TGCD managed to address was to provide future orientations (number 6) with a panel on *Greek-Turkish relations: prospects for the future* despite the fact that it lacked providing future orientations for the participants in interpersonal level. The last point the project addressed was the non-violent ways of managing a conflict (number 10). Under this category, the seminar on *three approaches to peaceful settlement* and practice-oriented sessions on *mediation process* and *simulation exercises* can be considered. This situation signals that the TGCD aimed to focus more about the Greek-Turkish relations in macro-level with an emphasis on past, present, and future fears and concerns present with each side. However, peace education in general puts a significant importance on the concept of peace and its effects on both micro and macro levels.

A thorough discussion of an ideal peace education design can be found in Section 5.3., where the results of the study are discussed.

The major theoretical background behind the idea of peace education is the socialization theory. According to Spence (1978), there are four main approaches to socialization theory.

1. *Freudian-psychoanalytical Approach*: argues that what is learned in the beginning is unlikely and very hard to change. Thus, peace education aiming to challenge the initial learning of the students is unlikely to be successful. Therefore, peace education will possibly function most effectively with the younger students with a fresh mind.
2. *Culture-personality Perspective*: argues that motivational uniformity and cognitive conformity are two essential elements for a society to exist. The views challenging this uniformity will be eliminated for the sake of the well-being of the society in general. Therefore, a peace education program can only be successful if it corresponds to the general discourse of the society.
3. *Social-learning Perspective*: argues that human brain is passive and receptive to the knowledge that is given. This perspective considers learning process as absorption.
4. *Cognitive-Developmental Approach*: argues that people tend to learn in accordance with their personal affiliations. “It holds certain basic frameworks or orientations, like personality, identification with a particular political party, ideology, social class, interest group, or ethnic community, will determine or structure the acquisition of specific beliefs.” (Spence cited in Ohanyan, 2005, p. 78)

Among these four approaches, the most appropriate ones for this study are the last two, *social learning perspective* and *cognitive developmental approach*. The facilitators and organizers of TGCD held the *social-learning perspective* in designing the tool of peace education. Their assumptions were that given the basic skills in conflict resolution and peace education, participants will challenge their assumptions and views about the conflict and this will trigger attitudinal change towards each other. On the other hand, *cognitive-developmental approach* argues that learning is political in nature by stating that “the participants will internalize only information that conforms to their belief system, ideology, and values and peace education is not likely to have a significant effect in terms of attitudinal change among participants.” (Ohanyan, 2005, p. 79)

When applied to the context of TGCD, the *cognitive developmental approach* implies that ethnic identity and the political status of the participants in the conflict determine the level of responsiveness of participants to peace education so there is either small or no room to tackle any established prejudice and stereotypes since they are internalized by the participants. As a response to this pessimist approach, Aboud and Levy (2000) state that “an important achievement in such peace education programs is the opportunity for the participants to achieve self-insight in order to challenge the learned norms and stereotypes to which they are exposed in their respective social setting.” (Aboud and Levy cited in Ohanyan, 2005, p. 79) According to this argument, the subjects of peace education learn to challenge their established knowledge as a result of socialization before changing their attitudes towards the members of the other ethnicity. This argument combines the two opposite approaches, social learning and cognitive developmental, by stating that with peace education as the tool, change among the conflicting parties is possible but it does not occur overnight. “Attentiveness to the processes and mechanisms through which the students develop certain positions in regard to the conflict, and become aware of their biases will create a strong ground for changing the structure of intergroup relations.” (Ohanyan, 2005, p. 80) The respect for the role of process in peace education is a key element in the success of initiatives like TGCD.

Peace Education is still a growing field where there is a significant amount of work conducted in a variety of contexts. However, the changing nature of conflict around the globe should also be taken into consideration, and, instead of attempting to devise a generic framework; new tools for alternate cases should be developed. This thesis evaluates peace education as a treatment together with contact, which attempts to improve relations between Greek and Turkish youth and provide them with certain skills and knowledge on peace.

2.3. Evaluation Research

This final section of the literature review focuses on the importance of research on Evaluation Research since it constructs one of the major pillars of this study. Chapter 3 will also shed light on the theoretical aspects of the methodology used while evaluating the efficiency of the micro-project on Greek-Turkish relations. In this section, an operating definition of evaluation will be presented followed by the different drives, as well as types and challenges in evaluating an intervention. After a section on different toolkits for evaluation, a brief summary of the literature on theories of change will be presented.

2.3.1. Defining Evaluation

Carol H. Weiss offers an operating definition of an evaluation with five crucial elements: “Evaluation is the **systematic assessment** of the **operation** and/or the **outcomes** of program or policy, compared to a set of explicit or implicit **standards**, as a means of contributing to the **improvement** of the program or policy.” (Weiss, 1998) The first element is the *systematic assessment* which defines the scientific characteristic of evaluation that is to say, the research that is conducted should be compatible with the general social science criteria.

The second and third elements, *operation and outcome*, can be clarified jointly. While some evaluations choose to investigate the process of an intervention, others simply choose to focus on particular outcomes that the intervention aims to achieve. This thesis however, assesses both operation and outcome. The research will examine the prescribed practices that the organizers of the project set in advance, and how they designed it, as well as their efficiency in providing the intended benefits to the participants, namely an increase in the levels of trust and empathy.

The next crucial element while evaluating an intervention is the need for standards of comparison: the relation between the initial stated goal and the outcome or compatibility with the goals altered during the process. In this research, the standards for comparison are two-fold: 1) the success of the project in creating a significant level of trust and empathy among the participants, when compared to a set of non-participants, and 2) the theories of change, where the expectation of the organizers set for themselves will be compared with the actual outcomes.

The fifth and final element is the notion of improvement in the program evaluated. The results of a particular evaluation point at the strengths and weaknesses of an intervention,

which will form a base for similar future projects. A crucial aim of this study is also to improve the practice in those grass-roots projects, which aim to improve Turkish-Greek relations.

Even though Carol Weiss provides a clear operating definition, she is still one of the few scholars studying the need for evaluation as a strong component of social sciences. Her work *Evaluation* can be considered as the key work that evaluators from many different practices might find very useful but her efforts to date have not yielded a good deal of research with itself. Lewis (2004) also criticizes the lack of evaluation research, pointing out to the fact that most evaluation models and theories are created to cover fields such as humanitarian assistance and development. Lewis furthers her arguments by mentioning that peace and conflict resolution fields lag in this respect mainly due to the obstacles and difficulties of assessing conflict environments. The existence of such a gap thus is a motivational factor for this thesis as well and evaluation of a micro-project in this field will contribute to fill this gap in the literature. Evaluation is a powerful tool in linking theory with practice and increasing effectiveness in conflict-resolution interventions. This can be achieved by the function of evaluation requiring explicitness regarding goals, process and theory. (Elliot; d'Estree; Kaufman 2003)

2.3.2. Incentives for Evaluation

Often, there is not a single incentive for evaluation. According to Lewis (1998), there are three types of drives for evaluation. The first one is the funder-driven evaluation, where the funder requires evaluation as a means to improve efficiency so as to lead to a decrease in unnecessary expenditures in future interventions and also to increase the accountability of the funding agency. Despite these positive effects of evaluation, the funder of the TGCD did not choose to carry out an evaluation at any stage of the intervention. The second source of the driving force for evaluations is the practitioners. Practitioner-driven evaluations are the type of evaluations that serve as agents of feedback and recommendation during the intervention in order to help the practitioner control the process. This type of evaluation was also absent in the project analyzed in this research. There was only one Greek observer during all three meetings but there was no report produced by that observer. The third and last drive for evaluation is the most appropriate one for this study. It is the evaluation conducted with the drive for good Public Relations. Since the evaluation serves the purpose of validating certain interventions, an evaluator should be very careful about neutrality when engaged to a project. In the conflict context, doing evaluation well matters greatly in pragmatic terms because poor

interventions cost lives; moreover, competent evaluation matters in ethical terms because the practice helps “weed out” poor interventions before they exact such a cost. (Scriven, 1991) Instead of using a harsh term like “weeding out”, the proper evaluations of poor interventions can in fact improve the quality of intervention and provide them with another tool in their interventions.

2.3.3. Different Approaches to Evaluation

One major issue in evaluation literature is the presence of a large quantity of approaches. While many approaches are not field specific, innovative conflict resolution scholars have come up with different approaches to evaluation. Michael Elliott, Tamra Pearson d’Estree and Sandra Kaufman have created ‘Evaluation Utilization Continuum’ that constitutes four different approaches. This continuum differentiates between evaluations used as a tool for intervention on one side and on the other side evaluation for research. In this model the evaluator has a choice to pick the best approach to the case, subject to evaluation among four alternatives as following:

1. *Conflict Assessment*
2. *Formative Evaluation*
3. *Summative Evaluation*
4. *Knowledge-oriented Evaluation*

The view of evaluation as a tool for intervention starts with *conflict assessment*. This approach is conducted prior to the intervention and helps the parties set the goals and process themselves to build relationship at the very beginning. Next in line is *formative evaluation*, a structured process of reflection that seeks to provide input into program planning and revision. (Patton, 1997) Different from the previous approach, formative evaluation allows the evaluator to function during the entire process of an intervention and, if necessary, alter the course of events before it is too late. The third approach is summative evaluation, conducted at the end of the intervention and measures the overall effectiveness of the process. Examining the effectiveness of a series of similar interventions, might help the evaluator derive broader lessons from interventions. At the other end of the continuum where evaluation is conducted as a tool for research lies knowledge-oriented evaluation.

This approach seeks to accumulate lessons across cases and to build theory, contributing to the overall understanding of conflict. The products of knowledge-oriented evaluations are often aimed at understanding conflict

dynamics and improving general practice of conflict resolution rather than attempting to improve a specific intervention. (Elliott; d'Estree; Kaufman, 2003)

In this continuum, there are two approaches that define this case best: *knowledge-oriented evaluation* and *summative evaluation*. Knowledge-oriented evaluation is relevant in the sense that the evaluator in my case is not requested by the project organizers or the participants so he has no stake in the case other than creating a piece combining theory with practice. It is anticipated that the results of this research will serve similar interventions with the long term goal of improving the general practice of contact and peace education based interventions. On the other hand, summative evaluation was used to test the outcomes of the TGCD in creating significant levels of trust and empathy among its participants.

Apart from the approaches defined in the work of Elliott, d'Estree and Kaufman, another list of conflict resolution compatible approaches to evaluation is given by Helen Lewis (2004). The approaches in her study include:

- *Participatory Evaluation*: carried out by all possible stakeholders of a certain intervention. The capacities of these stakeholders are aimed to be developed during the course of events such as gathering and collecting data as well as generating recommendations. This approach seems to disregard the fact that most stakeholders lack the time and resources for even the intervention alone. There is a clear lack of a feasibility principle in this approach.
- *Utilization-Focused Evaluation*: a group of 'intended users' define an 'intended use' of the evaluation data. The aim of this approach is defined as "building the capacity of stakeholders to think and act evaluatively." (Lewis, 2004)
- *Impact Evaluation*: determines both positive and intended impacts as well as the negative and unintended impacts of the interventions evaluated. Instead of evaluating the short-term outcomes of a project, this type of evaluation measures the long term consequences and when doing so, acknowledges challenges such as the effect of external environment. This type of approach best suits the overall methodology of this research since the evaluation at hand is carried out a year after the intervention. Additionally, similar to the impact evaluation approach, this research aims to inform the organizers of the micro-project whether to expand, modify, or eliminate their interventions.

- *Action Evaluation*: slowly becoming a very popular approach among peace and conflict resolution researchers. “Action Evaluation joins a project by helping participants define and then formatively redefine success, to forge effective action and make success a self-fulfilling prophecy.” (Rothman, 2003) The method used in this approach to evaluation can be very briefly summarized as asking questions on organizers’ goals, their values and beliefs followed by their own suggested action strategies. Action Evaluation follows three basic steps: first, *establishing a baseline* for individual, group and organizational goals; second, *formative monitoring*, meaning that refining the goals in order to tackle the obstacles during the process of intervention. It is important to note that the goal-setting phase is also repeated at this step and the whole life-cycle of the intervention according to this approach. The final step is *summative evaluation* where the evaluators check how well the intervention met previously set goals. It is important to note that Action Evaluation is not just an evaluation approach. Different from other methods mentioned here, Action Evaluation is a form of collaborative social intervention. (Rothman, 2003) Additionally, Lewis notes that this type of evaluation is especially suited to the volatile conflict contexts.
- *Macro-Evaluation*: investigates the link between micro-level interventions and the macro processes such as track one diplomacy. (See Fisher 2005, Cuhadar-Gurkaynak 2004) An alternative definition more suitable to the field of conflict resolution defines the approach as the effect of only conflict resolution practices on the general dynamics of conflict resolution.

2.3.4. Challenges in Evaluation Research

Several scholars discuss the challenges in evaluation research but only some are relevant to evaluating conflict-resolution interventions. One particular author is Helen Lewis, who sees five main challenges in conflict case interventions. Lewis names the first as *timing evaluation* where she argues against Church and Shouldice’s idea of describing evaluation to be most fruitful during the life-cycle of an intervention. Lewis(2004) notes that “it is also important to re-evaluate interventions in order to track changes in their impacts and to determine their sustainability”. Therefore, the hesitancy of the organizers in conducting post-evaluations is addressed as a main challenge in conflict-case evaluations. This challenge is also relevant to the main research in a sense that the organizers planned no evaluations before, during, or after the workshop. This research varies at producing results on the long-term

effects of this type of initiatives and the sustainability of the effects that they aim to create at the first place.

The second challenge suggested by Lewis is the difficulties in *tracking change*. Lewis argues that most of the time the goals stated by the organizers are inflated or unclear. This discrepancy between the stated goals, objectives and the actions taken emerges as an important challenge for the evaluator. Lewis suggests the evaluator in that case to guess the real objectives of the interventions or disaggregate these goals and objectives into separate components for evaluation. However in my case, in-depth interviews with the organizers will be used in order to clarify the initial and modified goals through the course of the intervention. In addition, participant observation I conducted helps with this aspect of evaluation. The participant observer role will also serve as a useful factor in increasing the sincerity of organizers while sharing their experiences. Another challenge in tracking change is once again related to the sustainability of the results achieved after an intervention. There is a need to track long-term changes, for example in relationships, attitudes, and behaviors, that are triggered by conflict resolution interventions. (Lewis, 2004) In determining these long-term changes, the necessity of using both quantitative and qualitative indicators of change is stressed as well. These indicators are mainly social indicators (intermarriage between groups), security indicators (the rate of conflict-related deaths), and psychological indicators (groups' perceptions of one another). In seeking the long-term change that TGCD created among its participants, the main indicators used are psychological ones measuring the rates of trust and empathy among both communities.

Another crucial activity when conducting an evaluation is *attributing change*. It is the evaluators' job to map the connections between certain interventions and impacts. The challenge lies with the difficulty of mapping the transfer. Tracking the path and attributing change is the duty of the evaluator, however in this case, I will not investigate the transfer effects of the project, because it would widen the scope of the research to a great extent. The next set of challenges is also irrelevant to this research project but beneficial to briefly summarize them. One challenge is the dilemma of the evaluator in *engaging stakeholders* where the evaluator is bound with the time constraints of the funders while trying to provide useful feedback. The last challenge Lewis mentions is the vague terms in *ownership of evaluation results*. Both of these challenges are not present in this research, since the evaluations conducted on this project are totally self-funded with academic purposes.

Elliott, d'Estree and Kaufman (2003) also suggest a list of possible challenges during an intervention. These challenges can be briefly listed as following:

- *Large Scale:* When the conflict case of the evaluation is an intractable one, the effectiveness of a single intervention and its evaluation will not give broad lessons about that particular conflict.
- *Inflated Expectations:* Ambitious expectations lead to inflated goals, harder to realize at the end.
- *Complex Causality:* In certain types of conflicts, the bond between micro level interventions and macro-level outcomes is vague.
- *Need for Confidentiality:* Cases when participation in a conflict resolution process might be seen as synonymous to treason. In such cases, it would be difficult to collect necessary information for evaluation.
- *Unclear Indicators of Success:* Interventions with small but specific goals might seem insignificant to outsiders. Additionally, the change on goals during the course of events in some interventions will cause extra burden for the evaluator. The authors also provide an exhaustive list of multiple criteria applied to measure success. Some of these indicators of success are:
 - a. Achievement of an outcome
 - b. The quality of conflict resolution process
 - c. The quality of outcome
 - d. Satisfaction with outcomes
 - e. The quality of the parties' relationships
 - f. Improved decision making ability
 - g. Increased social capital¹

Among the challenges presented by Elliott, d'Estree and Kaufman, only one of them presented a significant challenge in the evaluation process of this research. *Inflated expectations* of the organizers resulted with a long list of goals with inadequate matches in activities. Other than that, among the criteria listing the indicators of success, the relevant indicators include bullet points (a), (b), (c), and (e). This point can be operationalized to include elements such as “new relationships resulting in increased trust and an improved emotional climate, reductions in hostility, an increased ability to resolve future disputes, new conceptualizations of the relationship and increased empathy between the parties.” (Elliott,

¹ In reviews of environmental and public policy disputes, inter communal conflict resolution and consensus-building processes, d'Estree, Beck, and Colby (2003), d'Estree, et. al. (2001) and Innes and Booher (1999) identified the criteria.

d'Estree and Kaufman, 2003) The first indicator of success, *achievement of an outcome*, will be revealed with the results of the experiment. (See Section 4.2) *The quality of outcome* will be determined after the thorough mapping of the TGCD's theories of change combined with the results of the experiment. The participant observation is the tool in evaluating *the quality of conflict resolution process*. The last indicator, *the quality of the parties' relationships*, will be tested among the participants and non-participants of the intervention and the qualities of relationships will be compared to check the efficiency of the micro-project.

The last discussion that I would like to elaborate on is the question of who should conduct a specific evaluation. Elliott, d'Estree and Kaufman (2003) come up with a model of three possibilities, each effective at a certain type of conflict. The first option is the self-evaluations conducted either by the facilitator or the participants themselves. Since the purpose of this evaluation is to improve a certain intervention, and the evaluator has also a stake in the success, there will be a significant amount of bias. Similarly, second option, peer evaluation, is conducted by conflict resolution practitioners, who are also inclined to have bias, since the success of an intervention will be beneficial to the field itself. Evaluations that seek to inform participants about the process and to promote active learning within that group can often be conducted by either of the two options mentioned above. On the other hand, the third alternative is an outsider with a strong grasp of both theory and practice conducting the evaluation. In that case, scientific rigor will be the priority and the results will not necessarily benefit the organizers. Evaluations that seek to influence outsiders or to determine effectiveness should usually be conducted by professional evaluators.

2.3.5. Toolkit for Evaluators

Elliott, d'Estree and Kaufman's idea of a 'toolkit of an evaluator' which combines the different works of prominent evaluation researchers is as following;²

1. *Observations*: two types: participant and field. In this study, the evaluator also has participated in the project making him a participant observer with a chance to observe the process of the intervention as well. The participant observer role is also useful in overcoming the problem of gate-keepers since the evaluator had a chance to personally know the stakeholders and organizers.

² The pieces used in preparing the toolkit are by the Works of Patton (1997 and 2002), Rea and Parker (1992), Marshall and Romsan (1999), House (1993) and Morgan (1998). The proper citations of these pages can be found in the detailed bibliography.

2. *Qualitative Interviews*: also classified as either less-structured or more-structured. In particular to this research, the evaluator chooses to use a more structured interview protocol derived from the theories of change that is applied to the organizing team of the project.
3. *Fixed-Response Interviews and Surveys*: evaluates the rates of trust and empathy among the participants of the project. Two very specific models of survey are used. (further discussed in the methodology part.)
4. *Focus Groups*: evaluator brings up to eight participants together and discusses the consequences of an intervention. There was no need to use this tool in this research.
5. *Document and Media Analysis*: reviews the meeting notes, supporting documents, newspaper and news accounts, government documents, and similarly recorded material. The documents that are analyzed in this research are the correspondence in a specific mail group and the grant proposal of the project.
6. *Dialogue with Participants*: takes corrections and clarifications coming from the participants into consideration. This is also a tool that will not be used in my research.

Among six basic tools that an evaluator has, this study covers four, used in different parts of the evaluation. Such a toolkit is extremely useful for evaluators who have little experience in the field and the specific methodology.

2.3.6. Theories of Change

Rein (1981) argues that every program is a theory and an evaluation is its test. However, in order to get a full grasp of the theory tested, an evaluator should get to know the program. Carol Weiss (1998) posits several reasons on the necessity for program knowledge, such as its use in developing a good sense of the issues, understanding the data, interpreting the evidence and making sound recommendations. Since my role in this research includes elements of participant observation, strong knowledge of program gave general sense of the program theory at hand. On a similar note, Ilana Shapiro provides a very brief definition of the concept as being “drawn from the literature on program evaluation, a theory of change refers to the causal processes through which change comes about as a result of a program’s strategies and action” (Shapiro, 2005, p. 1). Carol Weiss, a pioneer in evaluation research, however more specifically defines by differentiating between two types of theory first: program theory and implementation theory. Program theory can be described as a set of

hypothesis linking the program inputs to the expected outcomes, or as Wholey puts it to identify “program resources, program activities, and intended program outcomes, and specifies a chain of causal assumptions linking program resources, activities, intermediate outcomes, and ultimate goals.” (Wholey cited in Weiss, 1998) On the other hand, implementation theory focuses on the accurate delivery of the planned activities for an intended outcome. The assumption is that if the activities are conducted as planned, with sufficient quality, intensity, and fidelity to plan, the desired results will be forthcoming. (Weiss, 1998) After illustrating each theory with very useful figures, Carol Weiss defines the combination of program and implementation theory as a program’s theories of change. In addition to these definitions, Andrea Anderson (2004) provides the necessary terminology in assessing a program’s theories of change. One of these terms is a *pathway of change*. This term is used to describe a map that illustrates the relations between a variety of *outcomes* that are each thought as *preconditions* of the long-term goal. The second term suggested by Anderson is the *indicators*, which helps in recognizing success. *Intervention* is another term meaning the actions required to fulfill each precondition in the map. The last term is *assumptions* used to explain why the whole theory makes sense. In this research, it is important for the evaluator to come up with a *pathway of change*, a map, derived from the information collected after the interviews with the organizers of the project. Weiss also outlines the need for such a map to be created. According to Weiss, “a theory map provides a picture of the whole intellectual landscape so that people can make choices with full awareness of what they are ignoring as well as what they are choosing to study” (Weiss, 1998, p. 62). This is exactly the case in this study where some specific parts of the project were chosen to measure their effectiveness. A broader map of the whole project at that point makes this process of choosing easier for the evaluator in a sense to be aware of the trade-off being made.

There is also a debate among the evaluation authors as to who should prepare the final version of this map. Wholey (1987) and Patton (1989) leave the final decision to the project organizers and the key stakeholders, whereas Chen and Rossi (1980, 1983) believe in the rigor of the social sciences and the importance of theory rather than practice and gives the last word to the evaluator in creating this ‘pathway of change’. Chen (1990) followed by Weiss (1995, 1997) creates an alternative where both practitioners and the social scientists have to work together in creating a final map. Weiss (1998) also underlines the benefits of the communication between the practitioners and evaluators. In this research, the evaluator

follows the last point and collaborates with the organizers and stakeholders of the project, who had no ‘pathway of change’ created for the project, derived from their hypothesis.

In helping the organizers bring their ‘theories of change’ to life, the evaluator has used the model of Ilana Shapiro (2005). This model highlights the crucial importance to know at which level of analysis the intervention is intended to function. She provides certain tools of individual, relational, and social change which will further be discussed in the methodology part.

As a conclusion of this section, I have to say that the literature on evaluation is scarce and many approaches, tools and theories need continuous testing among the intervention cases around the globe. There is also a specific need to come up with rigorous, scientifically backed evaluation models in conflict resolution and peaceful intervention cases.

3. RESEARCH DESIGN

This chapter comprises three sections. The first one gives a detail account on participants of this research. The second section presents the multiple methodologies incorporated in this research together with the tools for data gathering and data analysis. The third and last section outlines the process operating throughout this research and reflections on how to improve the current design adopted for this research. Before moving any forward, the following figure outlines the general lines of the research design used in this thesis. For summative evaluation, a quasi-experiment was conducted using a control group to measure the outcomes of the workshops. For process evaluation, which aimed at understanding the planning and implementation of the project, I conducted interviews with the organizers of the project as well as participant observation due to my participant role in the project.

1. Summative Evaluation

- * Quasi-Experiment

2. Process Evaluation

- * Structured Interviews
- * Participant Observation
- * Mapping

3.1. Participants

The participants in this study will be examined in two categories: the experimental and control group. The experimental group consists of twenty-six ($n=26$) *upper level* students who participated in the TGCD project and thus received the treatment. The term *upper level* is operationalized by the organizers of TGCD in their mission statement in order to refer to a group of students who are either at least senior undergraduate or graduate students. In terms of age, the participants in this group range from twenty-two to thirty in which the average is twenty-four. The participants of the TGCD are determined by the organizers to include an equal number of students with regard to their ethnicity and gender: fifteen Greek and fifteen Turkish participants even though there is only a slight difference when we consider gender distribution. Overall female participants comprised 46% and male participants comprised 54% of the group. Originally, the ratio of male to female participants in each ethnic group was the same: namely 8:7. While forming the experimental group out of the TGCD participants, a Turkish female participant refused to be a part of this research stating that she did not want to make generalizations in life according to her perspective. There were three Greek participants,

one female and two male, that did not take part in this study. At the end, the experimental group resembled the same levels of gender distribution.

The participants of the TGCD were admitted after an application and then personal interview process conducted by Dr. Phillipos Savvides in Athens and by Sansel Ilker in Istanbul. The turnout rate in the application process differed. There were fifty applicants in Greece whereas the turnout was limited to twenty in Turkey. According to Dr. Savvides, the low turnout rates were due to the late announcement of the project on both sides, which is because of the uncertainty related to the grant from the European Commission until early September, 2004. The main requirements from the applicants were a certain level of command in English and basic knowledge of Greek-Turkish relations. Even though there was no requirement, the majority of the participants were selected from social science departments of the participating schools. Since there was insufficient time for a proper announcement, there was not a variety of different schools participating in the project. Among the Turkish participants, students from Sabanci University seem to dominate the group with a ratio of 4:7, where as the University of Athens seems to be the dominant school among Greek participants with a ratio of 1:2. Table 1 provides additional information on the participants' school affiliations.

TABLE 1: School Affiliations of the Participants in the Experimental Group (n=26)

TURKISH PARTICIPANTS		GREEK PARTICIPANTS	
Name of the School	No. Of Participants	Name of the School	No. of Participants
Sabanci University	8	University of Athens	6
Bilgi University	2	Panteion University	3
Istanbul University	1	Aristotle University	1
METU	1	ASOE	1
Bosphorus University	1	University of Essex	1
University of London	1		

The second group used for this research is the control group. The control group was comprised of thirty (n=30) *upper level* students. The term *upper level* students was borrowed from TGCD in order to keep the control group similar to the experimental group. This was also helpful in controlling the average age in each group as a constant variable. These

similarities were also valid for ethnicity and gender distributions within the group. The control group is ideally comprised of equal numbers of students from each ethnicity and gender. As a consequence, the control group in this study consisted of eight Turkish female students, seven Turkish male students, eight Greek male students and seven Greek female students. In a distribution like this, male participants comprised 51.7% of the sample (n=56) but this rate changes when we look separately at each group: 54% among the experimental and 50% in the control group. The distribution of ethnicity was also the same where Greek participants comprised 48.3% of the sample (n=56) with the percentage of Greek participants differed in each group: 46.2% in the experimental group and 50% in the control group. Table 2 provides information on the distribution of gender and ethnicity among each group.

TABLE 2: Demographic Comparability of Experimental Group and Control Group

	Experimental Group (n=26)	Control Group (n=30)	Total Sample (n=56)
Greek Students	46, 2 %	50 %	48, 3 %
Turkish Students	53, 8 %	50 %	51, 7 %
Male Students	53, 8 %	50 %	51, 7 %
Female Students	46, 2%	50 %	48, 3 %

Finally, a note needs to be made regarding the selection of the students in the control group. In both countries, there were two main schools which provided most of the participants for the experimental group. In Turkey, there were eight students from Sabanci University which comprised more than 50% of the Turkish students in the experimental group. Next was Bilgi University, which sent two students to TGCD project. Both of these universities in Turkey are private and the student profile and education qualities are similar. This fact justifies the selection of ten students in Sabanci University from similar departments to the students in the experimental group and five students from Bilgi University. In Greece, on the other hand, the University of Athens was the dominant provider with seven students, who participated in TGCD. Panteion University was the closest institution with three students among the experimental group. Since the difference in the number of students was not as high as in the Turkish case, the control group was divided to include nine students from University of Athens and six students from Panteion University. Once again, the student profile and educational quality in these schools are similar to each other since both of these universities

rank among the top universities in Greece. The distribution of the students in the control group can be seen in Table 3.

TABLE 3: School Affiliations of the Participants in the Control Group (n=30)

TURKISH PARTICIPANTS (Control)		GREEK PARTICIPANTS (Control)	
Name of the School	No. Of Participants	Name of the School	No. Of Participants
Sabanci University	10	University of Athens	9
Bilgi University	5	Panteion University	6

3.2. Methodology

Before proceeding to describe the measures used in this thesis, the research question and the hypothesis of this study should be discussed. The research question is:

How effective was the initiative of ‘Turkish Greek Civic Dialogue: Learning Conflict Resolution and Producing Peace’ in improving relations among its participants? Is it possible to sustain those improved relations after a year?

In light of this question, it should be highlighted that the research question has two components which required a combination of several different methodologies. The first part of the research question required the use of certain tools such as structured interviews and participant observation in order to come up with a map that outlines the TGCD’s theories of change. Thus, the first part is a process evaluation, where I examined whether the TGCD project achieved their initial goals stated at the beginning. On the other hand, the second part of the research question was examined by conducting a summative evaluation. In this part, a quasi-experiment with time-series design was used to measure the levels of attitudinal trust and empathy, as well as the behavioral empathy between the experimental and the control group after a year from TGCD. Where the first part is qualitative in nature, the second part necessitates the use of more quantitative measures in order to work on the three main hypotheses given below:

Hypothesis 1: Students who met through the workshop experience will have greater trust in members of the other ethnicity than will non-participants after a year.

Hypothesis 2: Students who met through the workshop experience will have greater empathic attitude for members of the other ethnicity than will non-participants after a year.

Hypothesis 3: Students who met through the workshop experience will have greater empathic behavior for members of the other ethnicity than will non-participants after a year.

All of the three hypotheses above are stated in one direction³ and thus the problem is a one-tail test of the hypothesis. It is important to mention that a t ratio⁴ of 1.94 is significant at the 0.5 level for a one-tailed directional test for the purposes of the experiment, which will be explained in greater detail in Section 3.2.2. of this chapter. The operational definitions of trust and empathy will also be provided in the ‘*The Experiment*’ section.

³ The hypotheses are one-directional because they measure the effects of the treatment itself without combining it with any other variables.

⁴ The t-test is a ratio of the difference between the means (averages) of two conditions to a measure of variation within each of the conditions, referred to as the Standard error of the difference. (Druckman, 2005. pg. 87)

It is important to mention that most of the researchers in the field choose one of two methods: namely they either just provide a summative evaluation of a program to check the compatibility of goals with outcomes, or they only measure the acquisition of certain skills after a workshop. In contrast, this investigation is unique in that it provides a bigger picture by bringing these two questions together. It starts with a process evaluation, where TGCD's theory of change is mapped. "Drawn from the literature on program evaluation, a theory of change refers to the casual process through which change comes about as a result of a program's strategies and action" (Shapiro, 2005, p.1). The research then moves forward with a summative evaluation to measure the acquisition of two major skills, trust and empathy, after the workshop experience.

3.2.1. Program Evaluation

The first part of the research question investigates whether TGCD was effective in creating the patterns of change that they have anticipated in the beginning. In order to conduct such inquiry, the most suitable research method is process evaluation, also known as formative evaluation. Process evaluation therefore is conducted after an intervention is carried out and focuses on the overall effectiveness of the program. "It draws general lessons from interventions in order to improve conflict resolution practice over time" (Elliott, d'Estree, Kaufman, 2003, p. 4). Since the research in this thesis is not funded by TGCD itself, the main purpose here is to contribute to the general practice of grassroots Turkish-Greek youth initiatives by providing an analysis of a single program with its achievements and shortcomings. In addition to contributing to the field, this study will also be useful for the organizers of TGCD project in their future projects.

Speaking on the issue of program evaluation, Carol Weiss (1998) suggests that evaluators should familiarize themselves with the program area and with the specific program they will be studying. It was relatively easier for me to accomplish this important step, mainly due to two reasons. First, as a Masters level student in conflict resolution, I am familiar with the program area and second, I had the chance to meet the organizers of the project, which allowed me to conduct in-depth interviews. My role as a participant observer together with the interviews I conducted with the organizers allowed me to comment on the program area as Weiss suggests.

The next step involves mapping a programs' theory of change. It is important for the practitioners, especially in the field of peace-building and conflict resolution, to have a theory

of change in order to use it as a map when designing an intervention. According to Dayton and Cuhadar (2005)⁵ theories of change have three basic features:

1. Assumptions about the underlying root(s) of the conflict/ problem being confronted (and how they are linked)
2. Beliefs about the conditions under which these root causes can be transformed (either in a positive or negative direction)
3. Beliefs about what kind of programmatic interventions bring about what kind of changes.

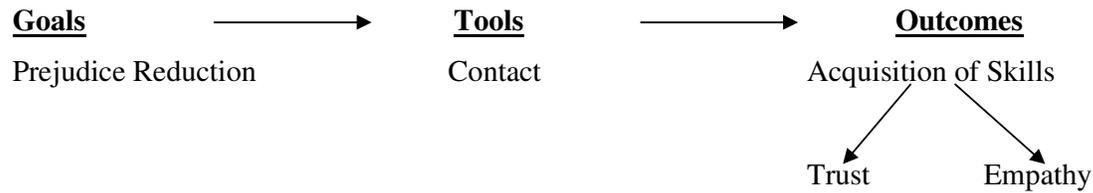
This thesis will also attempt to ascertain whether the TGCD project developed its own theory of change. This information was obtained through conducting interviews with the planning committee of TGCD, studying their action plans designed prior to the project. In the case of the lack of such a theory of change, it is again the duty of the researcher to retrieve information from the organizers and map the theory of change himself.

Initial interviews with practitioners can elicit programs' theories of change as well as both the explicit and implicit logic of an intervention design. This includes how they: 1) frame the specific problems to be addressed; 2) frame their intervention goals; 3) identify processes through which change happens; 4) describe their strategies, principles and specific methods for intervention; and 5) delineate short-and long-term intended effects (Shapiro, 2005, p. 3)

Covering most of these essential points, an interview protocol created by Cuhadar and Dayton (2005) was used in this research. The protocol prepared for the organizing committee of TGCD consists of questions such as: *What do you believe is at the heart of the Greek-Turkish conflict?* and *What were your anticipated outcomes?* Appendix II contains the interview protocol used in this study.

According to Weiss, the theory map provides a picture of the whole intellectual landscape so that people can make choices with full awareness of what they are ignoring as well as what they are choosing to study. (Weiss, 1998, p. 62) Once the mapping of the theories of change of an intervention is over, it is the evaluator, who chooses the lines of inquiry to pursue. A thorough map of a program's theory of change is provided in the results section of my thesis, but, for the readers' benefit, the simple linear diagram below plots the basic stages of the mapping process.

⁵ This piece has not been published yet, but presented in conferences such as "Oslo and Its Aftermath" and "Organizing for Peace and Its Education: Comparing Notes". This piece is used by the permission of Esra Cuhadar-Gurkaynak.



Under the light of this basic pathway of change, this study aims to investigate the level of sustainability in the acquisition of two major skills: Trust (attitudinal) and Empathy (both attitudinal and behavioral). However, there is an important debate among evaluation researchers at this stage. The problem is whether the program is responsible for whatever outcomes are observed. “Many things other than the program might give rise to desired outcomes. The usual way to tell whether the program was responsible is to compare units that received the program with equivalent units that did not. (Weiss, 1998, p. 60-61) This question takes me to the second part of the research question where I use an experimental design as previously mentioned to question whether TGCD is responsible for any change created among the participants.

3.2.2. The Experiment

There are three main hypotheses that this study aims to evaluate. In order to reduce the numerous concerns of internal and external validity, experimentation is one of the most appropriate methods that can be used at this stage of the study. Before moving on with the specific experimental tools that were used in this research, I would like to acknowledge another research project that inspired my work. Deepak Malhotra and Sumanasiri Liyanage (2005) evaluated a four-day peace workshop conducted with a group of Sri-Lankan participants similar to those involved in this research. Their work discusses the long-term impact of such interventions on attitudes and behaviors in the context of protracted conflict. Since a very similar research was conducted in this study, their approach was adopted with a main difference that they have only measured attitudinal and behavioral empathy rates among their participants, whereas attitudinal trust measures were added to this study.

Creswell (1994) makes a clear distinction between three different types of experimental design: pre-experimental, quasi-experimental and classical experiment. In this study, a quasi-experimental design was chosen, which generally takes the form of a pre-test,

post-test comparison of two groups with a control group serving as the additional feature of this design. (Druckman, 2005)

Unfortunately, due to certain limitations and restrictions, the research at hand has some flaws. The experiment can be illustrated as following, where X signifies the intervention and 0 signifies the tests.

Group A: (EXP)	0	X	0
Group B: (CNT)			0

In a perfect design, there should have been a pre-test, an immediate post-test, and the final test after a years time. This figure can be illustrated as the following.

Pre-Test → Treatment → Post-Test (Immediate) → Post-Test (A Year After)

Even though the TGCD administered questionnaires to be given to the participants after each workshop, none of these worked as an accurate pre-test for this research. The organizers used a questionnaire prepared by Esra Cuhadar-Gurkaynak (2005) after each meeting including questions such as: *'Is this the first time you are participating in a joint Greek-Turkish workshop?'*, *'Tell us a little about your views on Greek-Turkish conflict before attending the IPC-ELIAMEP workshop,'* and *'What did you gain from the first workshop?'* Such questions manage to address the initial views of the participants in order to serve as an acceptable pre-test even though it can't be considered as a proper experimental instrument. The complete set of questions can be found in Appendix III.

Since it is hypothesized that there would be differences in both attitudes and behaviors between the experimental group (TGCD participants) and, the control group (TGCD non-participants) the participants filled out a questionnaire. The questionnaire comprises four sections: Attitudinal Trust (first 15 questions), Attitudinal Empathy (following 5 questions), Behavioral Empathy (a single question), and the final section asking a series of demographic questions.

Since there is a related procedure for measuring attitudes devised by Likert (1932), his five-step scale was used, ranging from strongly agree to strongly disagree, which is widely used in social science. (Druckman, 2005) The first twenty-questions in the questionnaire were developed so that on a scale one to five, 1 referring to strongly disagree and 5 referring to strongly agree.

The trust scale used in this experiment was developed by Roy J. Lewicki and Maura Stevenson in 1999. The authors defined a sequential model of three different types of trust, which is used in this research: Calculus-based, knowledge-based, and identification based trust. (For the operational definitions of trust and the sequential model see Section 2.1.)

The first five questions in the questionnaire measure the levels of calculus-based trust among the participants by asking them to rate five statements on a five-point Likert scale. Some examples of these statements are: *'Greeks (Turks) are known as people who keep their promises and commitments'*, and *'Greeks know that the benefits of maintaining trust are higher than the costs of destroying it.'* Following the first set, there are four statements, questions 6-9, measuring knowledge-based trust such as: *'I think I really know Greeks,'* and *'I can accurately predict what Greeks will do.'* Finally, there are six statements, questions 10-15, that measure the levels of identity-based trust. Sample statements from the questionnaire are: *'Turks and Greeks share the same basic values,'* and *'I know that Greeks would do whatever we would do if we were in the same situation.'*

As far as the measures for empathy are concerned, the empathy scale was adapted from Davis's (1980) "empathic concern" (EC) subscale of the Interpersonal Reactivity Index and is designed to measure the degree to which a person has concern for the other's well-being. There is a slight difference between the model adapted by Malhotra and Liyanage, and the one employed in this research in the sense that they have used a 7-point Likert-type with the same questions where the 5-point type was used in this study. There are five statements in this section of the questionnaire regarding empathy such as: *'I would get very angry if I saw a Greek (Turk) being ill-treated,'* and *'I would get emotionally involved if a Greek (Turkish) person that I knew were having problems.'* Appendix II contains the empathy scale as used in this study.

Malhotra and Liyanage also included a behavioral measure of empathy in addition to collecting the attitudinal measure of empathy. "When the respondents arrived to fill out the questionnaires, they were given Rs. 200 as a show-up fee. (Rs. 200 is roughly equivalent to US \$2, but Rs. 200 is worth significantly more in Sri Lanka given the local economy. A *purchasing power parity* conversion in 2002 would yield an amount closer to \$9.) After filling out the questionnaires, the final question in the questionnaire packet informed respondents that a fund-raiser was underway to help poor children of the other ethnicity." (Malhotra and Liyanage, 2005. pg. 917) The organizers in this study actually asked the participants to write down the amount of money they were willing to donate for record-keeping purposes and drop their money to a box on their way out of the room once the questionnaire was complete. The participants were only told in the briefing session that this was a behavioral exercise and that the amount of money written on the questionnaire would be donated by the organizers themselves. Unfortunately, this was not the case in this research and the reason for such limitation is outlined in Section 3.5. However, a behavioral question was still used in this

research, asking the participants the amount of money they would donate in case of a tragic event that the members of the other ethnicity are facing.

The final part of the questionnaire asks the participants demographic questions as well as whether they have interacted with anyone from the other ethnicity before. This last interaction question was only given to the control group since the information was already collected for the experimental group in a pre-test evaluation survey.

3.3. Data Gathering

There were two different processes running simultaneously throughout this study, mainly due to the fact that the research question itself is two-fold. In the beginning, this study started as an attempt to improve the grass-root initiatives dealing with the Turkish-Greek conflict, which exceeded the limits of a Masters thesis. This was basically the reason behind narrowing down the question to involve a single case and evaluate the effectiveness of a project operating in grass-roots level. The role of the evaluator as a participant in the TGCD project determined the case to cover this project as the focus of the study.

To assess the program's theory of change, structured interviews were conducted only with Dr. Savvides, who has the sole knowledge on program design, implementation and assessment. Although he provided useful data, the face-to-face interview that took place in Istanbul revealed that there were no theories of change employed during the design of TGCD. Despite the initial disappointment, this problem is mentioned as a common feature of such projects in the literature. "Evaluators wonder why program goals are often stated in fuzzy terms...They often have an intuitive rather than an analytic approach to program development and concentrate on formulating activities rather than objectives." (Weiss, 1998, p. 52) This is why a map for TGCD's theories of change was prepared by the evaluator himself. In preparing such map, the sequential model of Cuhadar and Dayton (2005) adopted from the works of Weiss (1998), on design, implementation and assessment provided the grid. There are seven essential steps in this model:

1. Conduct Conflict/Program Analysis
2. Develop Project Goals
3. Conceptualize a Theory of Change
4. Design Project Activities
5. Implement Project

6. Disseminate/ 'Transfer Outcomes'

7. Assess Impacts

Even though the organizers did not plan their project in this fashion, the model presented above served as a guideline in categorizing the data received from the interviews with the organizers. The crucial part was to conceptualize the project's theory of change, where the information gathered from Dr. Savvides was grouped into three. These groups were:

- Assumptions about the underlying root(s) of the conflict/problem being confronted and how they are linked. (Questions 1, 2 and 3 in the interview)
- Beliefs about the conditions under which these root causes can be transformed; either in a positive or negative direction. (Questions 4 and 5)
- Beliefs about what kinds of programmatic interventions bring about what kind of changes. (Questions 6 and 7)

The work of Shapiro combined with the interview results proved a useful basis for this study. In her study, Shapiro examines "a variety of program perspectives on how change happens in individuals, intergroup relationships, and social systems and points to some divergent theories of change that are prevalent in conflict resolution work." (Shapiro, 2005, p. 4) A process evaluation helps determining how the mechanisms for change developed, which would further commit to the improvement of future projects. (For the detailed description of certain mechanisms for change employed within the TGCD see Section 2.1.)

The second major process operating in this research is the experiment. Malhotra and Liyanage (2005) compared three groups of students regarding their empathic attitudes and behaviors toward members of the other ethnicity. With concerns of contribution to the literature, trust was added as a new component of this research. At the same time, the idea of having a control group reduces the threats of internal and external validity, this three group model made sense. Druckman (2005) argues that an experimenter's confidence in findings that support a hypothesized relationship between alternative treatments and outcomes increases, to the extent that threats to internal validity are reduced. However, due to certain limitations, which are discussed in Section 3.5., this research used only one control group instead of one.

The quasi-experimental design with two groups reduces threats to internal validity to a certain extent but in an ideal design randomization would have been used which brings out the lowest rates of threats to internal validity. The experimental group was naturally the participants of TGCD who received contact along with peace education throughout the

process as the treatment: X. The control group, however, was selected in order to show resemblance with the experimental group in terms of distribution of gender, ethnicity and age. A further trait of the participants in the control group was that they were selected mostly on the condition that they have either no or small levels of interaction with the other ethnicity. In order to secure the independent variable of contact, this trait was crucial.

Once the preparation of the questionnaires was complete, e-mails were used as the communication channel to send the questionnaires to the participants together with instructions. The reason for choosing e-mails is mainly due to the scarce time and resources that did not allow the evaluator to travel to Greece. Other than that, e-mails were expected to be a faster medium of communication. The questionnaires were sent first to the experimental group and then to the control group.

The experimental group was participants of the TGCD project who received contact and peace education as the treatment of this experiment. Another sampling method was used with the control group. The questionnaire process was divided into two with the control group. In dealing with the Greek side of the control group four initial contact persons were selected.⁶ These contacts were informed about the intention of the research to comprise fifteen Greek students, with balanced distribution of gender and ideally from the universities selected in the beginning, namely from University of Athens and Panteion University. Twenty-two responses returned from the initial contacts, therefore seven of the responses were eliminated to keep experimental and control group balanced.

The sampling of the Turkish side of the control group was handled in three steps. In the first step, the questionnaires were sent to my contact, Devrimsel Deniz Nergiz, at Bilgi University, who forwarded it to five *upper level students* there. As the second step, eight questionnaires were given to the MBA students in Sabanci University. As the third and last step, in order to balance the distribution of age between the experimental and control groups, two senior undergraduate students received the questionnaire.

⁶ My first contact was Maria Deca, coordinator of a summer-school forwarded my questionnaire to Professor Stefanou, who is the director of Post-Graduate studies in Panteion University.

The second contact Grigoria Kalyvioti, was one of the participants of TGCD project and also the director of a student association called SAFIA in Panteion University. S.A.F.I.A. (Student Association for International Affairs) is a non-profit, non-governmental student association based in Athens, Greece. S.A.F.I.A.'s aim is to promote the scientific research of international relations and the development of the communication between the Greek and foreign academic society.

The third contact is Aspurce Onay, a Turkish student, who has resided in Greece for several months.

Yorgos Triantafyllou provided the major output for the control group. He forwarded nine responses for the experiment. I would like to thank all of my contacts for helping me out in this process, where collecting information from another country via e-mail is a very difficult task to accomplish alone.

3.4. Data Analysis

After the data gathering process was complete, certain tools were used when analyzing the data at hand. While the process evaluation necessitated the use of qualitative measures, the summative evaluation necessitated a slightly more complex data analysis strategy.

The data gathered as a result of the questionnaires, were analyzed by SPSS using linear regression, one-way ANNOVA, factor analysis, and t-tests. The first and probably the most used statistical term in this study is *significance*. It has become a custom in social science to regard values that occur in less than 5 out of 100 samples, as being in the rejection region of the sampling distribution. This means that these values are unlikely to occur by chance alone. This is the significance threshold indicated by the symbol $p < .05$.

Furthermore, “the t-test is a ratio of the difference between the means of the two conditions to measure of variation within each of the conditions, referred to as the standard error of difference.” (Druckman, 2005, p. 87)

While analyzing the interviews, another method is used. The interviews were coded into three separate categories based on the model of Cuhadar-Gurkaynak and Dayton (2005). These categories are: 1) assumptions about the underlying root(s) of the conflict/problem being confronted and how they are linked, 2) beliefs about the conditions under which these root causes can be transformed; either in a positive or negative direction, and 3) beliefs about what kinds of programmatic interventions bring about what kind of changes.

3.5. Limitations in the Research Design

There are certain limitations that need to be mentioned to help the future research in this area. Although e-mails were selected as the faster medium of communication, the turn-out rate was relatively slow. It would be an ideal design if the researcher has the resources to travel and conduct the interviews in a class environment where multiple data can be gathered at the same time. At the same time, anonymity of the questionnaire participants will be preserved. Even though the responses arrived indirectly to the evaluator, the participants send their responses to the initial contacts with their names appearing on account information. At the same time, it is important to signify that this type of ideal research necessitates both adequate funding and flexible time.

Due to these scarcities of time and resources, I could not meet with each group face-to-face, which had impacts on the behavioral exercise as well. This is why the behavioral exercise used by Malhotra and Liyanage (2005) was altered in order to fit the medium of

electronic communication. The fund-raiser for poor children was changed to a tragic event that both Greek and Turkish participants experienced recently: the earthquake. The last question was formulated as ELIAMEP and IPC, the organizing institutions, awarded each participant 100 Euros for their final papers presented in a conference. When they learn about a tragic earthquake incident taking place in the other country, the participants were asked how much they are willing to donate from that money they received.

The last limitation was that one control group was set instead of two as in the work of Malhotra and Liyanage (2005). The main reason behind this constraint was that the second control group should have comprised of students who applied for the TGCD but could not get in. A test run with such a group would have answered some questions on the role of pre-disposition with such projects. Since the organizers of the TGCD were not in possession of such a list, the idea of having a second control group was postponed for future research.

Yet, there is another important aspect worth discussing in this section, which is the role of the evaluator as a participant observer in this research. Even though there are some limitations that this role brought, such as the question of bias, the advantages of the role of the evaluator is much more significant. First of all, this research began six months after the project has ended. So there was no overlap between these two processes. The major advantage of my role as a participant observer was that I had a chance to familiarize myself to the project, which is an important component of evaluation research according to Weiss. (1998) The problem of gatekeepers was eliminated and the organizers was helpful throughout this research and provided me all the documents I asked for, such as the grant proposal and the results of their own questionnaires administered after each meeting. Another advantage of my involvement with the project was to evaluate the process along with the outcome. I had the chance to talk with all of the participants and have their opinions on the quality of the project and their rates of satisfaction.

4. FINDINGS

The research question necessitated the use of two different evaluation approaches for this study: Summative Evaluation and Process Evaluation. Thus, the findings are also examined in two different sections. The first section comprises the results of the summative evaluation, where a quasi-experiment is used. The second section outlines the findings of the process evaluation, where structure interviews as well as the participant observation served as the main tools.

4.1. Summative Evaluation

As Daniel Druckman (2005) mentions, experiments work best when they serve as vehicles for evaluating hypotheses. The experiment in this thesis yielded several results that should be reported and analyzed carefully. Each hypothesis will be discussed in three different sections. Before moving on with the results, there is a need to explain certain statistical terms that are crucial while interpreting the data. The definition of certain statistical terms used in this research is provided in Section 3.4.

4.1.1. Attitudinal Trust

The first hypothesis stated that students who met through the workshop experience will have greater trust in members of the other ethnicity than will non-participants after a year. In this linear regression, experimental and control groups were treated as independent variables together with gender ethnicity. With attitudinal trust serving as the dependent variable, Table 4 reveals the results.

TABLE 4: Attitudinal Trust for Group, Gender and Ethnicity

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2,739	,139		19,651	,000
	GROUP	,351	,142	,313	2,471	,017
	GENDER	,144	,142	,129	1,019	,313
	ETHNICIT	,230	,142	,206	1,623	,111

a. Dependent Variable: TRUST

These numbers indicate that there is significance only with group as the independent variable where p-value was smaller than .05. ($\beta = .313$, $p = .017$) The positive beta sign allows

us to argue that level of trust is significantly higher in the experimental group compared to the control group. As an additional output, there is no effect for ethnicity and gender regarding the levels of trust. In order to make sure the data was valid; another course of linear regression was run with group as the only independent variable where trust was the dependent variable once again. Table 5 below supports the first hypothesis where the students participated in TGCD project had higher levels of trust for the members of the other ethnicity than the non-participants in the control group.

TABLE 5: Attitudinal Trust for Group

Coefficients^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2,927	,098		29,804	,000
	GROUP	,366	,144	,326	2,537	,014

a. Dependent Variable: TRUST

As I have discussed in Section 3.2.2., trust scale used in this study comprised fifteen questions and the designers, Lewicki and Stevenson, argue that it measures three different types of trust: calculus-based, knowledge-based, and identification based. Even though they explained all of these types of trust in their work, they did not mention which questions fit into which category. To be on the safe side, a factor analysis was used to see the different categories. Table 6 shows the distribution of questions and the results of the factor analysis:

TABLE 6: Factor Analysis for Trust

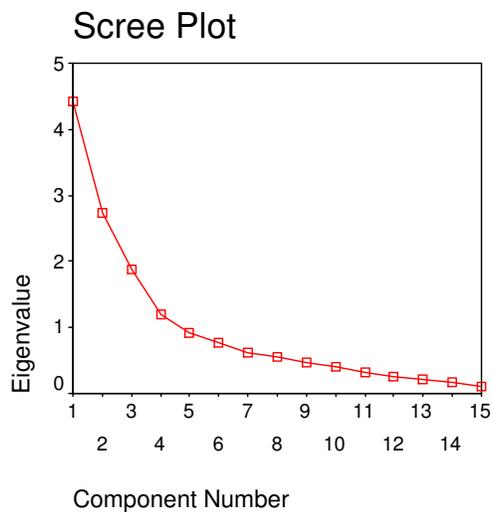
	Rotated Component Matrix^a		
	Component		
	1	2	3
Q1			,330
Q2			,786
Q3		,426	,619
Q4			,783
Q5			,849
Q6		,664	
Q7		,933	
Q8		,864	
Q9		,806	
Q10	,679		
Q11	,828		
Q12	,625		,369
Q13	,681		
Q14	,721		
Q15	,695		

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 4 iterations.

Depending on the descriptions of the authors, which is provided in Section 2.1., and the factor analysis above, it is safe to say that: questions 1-5 measures calculus-based trust, questions 6-9 measure knowledge based trust, and questions 10-15 measures identification-based trust. In the light of this new table, linear regression was used once again to test the results for specific types of trust. Only questions 3 and 12 appeared in two sections, which did not corrupt the results. A scree plot was also created to illustrate the breaking points of difference as can be seen in Table 7.

TABLE 7: Scree Plot for Trust



After seeing the results of the factor analysis, the need to compare the results for three different types of trust emerged. The first category was calculus-based trust. Lewicki argues that this type of trust is usually achieved first in a sequential model. However, the regression results were on the contrary. None of the independent variables of gender, ethnicity, and especially group, did have a significant effect on the dependent variable set as the calculus-based trust. In other words, there is no significant difference regarding calculus-based trust between the experimental and the control group. Table 8 shows the results of the regression.

TABLE 8: Calculus-based Trust**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2,507	,169		14,855	,000
	GROUP	,298	,172	,223	1,731	,089
	GENDER	,195	,172	,147	1,138	,260
	ETHNICIT	,324	,172	,243	1,887	,065

a. Dependent Variable: CALCULUS

The second category defined in the trust scale is knowledge-based trust, achieved as the second step in the sequential trust model. The regression results in Table 9 show that the difference between the experimental group and the control group regarding knowledge based trust was not only significant but also reached a peak with $\beta = .243$ and $p = .00$. There are no detected effects of gender and ethnicity once again. This finding is indeed interesting and will be discussed in further detail.

TABLE 9: Knowledge-based Trust**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2,239	,240		9,333	,000
	GROUP	1,093	,245	,526	4,470	,000
	GENDER	-,112	,244	-,054	-,459	,648
	ETHNICIT	,084	,244	,041	,346	,731

a. Dependent Variable: KNOWLEDG

Last but not the least; identification-based trust was tested with gender, ethnicity and group. The results reveal no effect of independent variables, gender, ethnicity, and group, on the dependent variable. When the experimental group is compared to the control group in Table 10, there was no significant difference.

TABLE 10: Identification-based Trust**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,267	,192		17,043	,000
	GROUP	-,099	,195	-,068	-,507	,614
	GENDER	,273	,195	,188	1,400	,167
	ETHNICIT	,249	,195	,171	1,278	,207

a. Dependent Variable: IDNTFCTN

These data are indeed useful in showing that the high degrees of difference in knowledge-based type of trust between the experimental and the control group, which is an interesting finding. This fact could have been discarded if there was no factor analysis employed in the experiment. The results are surprising in the sense that contrary to what Lewicki argues about the sequential trust model, where calculus-based trust is usually established first, followed by knowledge-based and finally identification-based trust, there are no significant differences observed between the groups regarding the calculus-based and identification-based trust. Instead, the main significant difference between the experimental and control group is detected with the levels of knowledge-based trust. The implications of this finding will be elaborated in Section 5.1., yet it is safe to say that the TGCD managed to elevate the level of trust from calculus-based to knowledge-based trust among the participants of the experimental group. Since both groups are selected from universities with a high socio-economic background and possibly a politically more moderate standpoint in both countries, I conjecture that student in either the experimental and control groups had a basic level of trust. Therefore, there is no significant difference with respect to calculus-based trust. On the other hand, identification-based trust is achieved rarely, usually after an extensive period of friendship and collaboration. So, it was not surprising to see no significant levels of identification-based trust achieved in both groups. This is indeed one of the major findings of this study.

As a last measure on attitudinal trust, a series of t-tests were run in order to see the levels of trust emerged in each group and whether it was possible to comment on the success of the TGCD in being responsible for this acquisition of trust. Tables 11 and 12 below show the results of the t-test for the experimental group. In Table 11, one-sample test, the values where $p < .05$ in Sig. (2-tailed) column, allows the evaluator to comment on the results. In this case, there is significance only with identification-based trust, which allows the evaluator to

look only for this variable in Table 12. The general rule in evaluating the results of a five-point Likert scale is that when the mean is above 2.5, there is significance. The mean for identification-based trust is 3.44 in this case, which gives the confidence to say the participants in the experimental group have identification-based trust but not significantly different from the control group.

TABLE 11: T-Test for Three Types of Trust in Experimental Group

One-Sample Test

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
CALCULUS	,603	25	,552	,0846	-,2046	,3738
KNOWLEDG	1,973	25	,060	,3173	-,0140	,6486
IDNTFCTN	3,304	25	,003	,4487	,1690	,7284

TABLE 12: Average Mean in the Experimental Group Regarding Trust

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
CALCULUS	26	3,0846	,71593	,14041
KNOWLEDG	26	3,3173	,82024	,16086
IDNTFCTN	26	3,4487	,69245	,13580

However, the linear regression results mentioned previously revealed that the TGCD was not responsible for the acquisition of this specific skill. Another strong indicator for this situation is that the t-test results for the control group reveals similar results regarding the identification-based trust. The results are illustrated in Tables 13 and 14 below. With the control group, the evaluator was able to comment on all types of trust since the p value is smaller than .05 with each variable. It would be correct to say that the participants in the control group showed the highest level of trust in identification-based trust with mean 3.5, followed by calculus-based trust with mean 2.7. However, the results show that the knowledge-based trust was not a trait achieved by the participants in the control group. It is still important to note that these additional tests were administered to see the levels of three different types of trust. Still, the initial findings are valid where the levels of calculus-based and identification-based trust are insignificant between the experimental and the control group.

TABLE 13: T-Test for Three Types of Trust in Control Group**One-Sample Test**

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
CALCULUS	-2,117	29	,043	-,2333	-,4587	-,0079
KNOWLEDG	-4,426	29	,000	-,7750	-1,1331	-,4169
IDNTFCTN	3,721	29	,001	,5278	,2377	,8179

TABLE 14: Average Mean in the Control Group Regarding Trust**One-Sample Statistics**

	N	Mean	Std. Deviation	Std. Error Mean
CALCULUS	30	2,7667	,60363	,11021
KNOWLEDG	30	2,2250	,95896	,17508
IDNTFCTN	30	3,5278	,77692	,14185

4.1.2. Attitudinal Empathy

The second outcome this research is interested in is empathy and whether students who met through the workshop experience have greater empathic attitude for members of the other ethnicity than the non-participants after a year. Once again, linear regression test was used to elaborate on data collected after the questionnaires. The results are in Table 15:

TABLE 15: Attitudinal Empathy for Group, Gender, and Ethnicity**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4,166	,154		27,090	,000
	GENDER	-,075	,156	-,063	-,483	,631
	ETHNICIT	-,204	,156	-,171	-1,305	,198
	GROUP	,330	,157	,277	2,106	,040

a. Dependent Variable: EMPATHY

This figure shows that when tested for independent variables, gender, ethnicity, and group, empathy as dependent variable yields a significant result only for group, which means

that the experimental and control groups differ on this measure. ($\beta = .277$, $p = .040$) Therefore, hypothesis two is supported and also the positive sign of the beta value (0.277) shows that empathic attitudes were significantly higher in the experimental group than the control group. Since p value in this case is 0.040, which is lower but also closer to the threshold (0.05), another linear regression was used with putting group as the only independent variable. The result is illustrated in Table 16 below and it is in compliance with the initial results illustrated in Table 15. The value of significance rose to a certain extent but it was still under the threshold of $p < .05$, which verified the second hypothesis once again. There was no factor analysis and t-test conducted regarding attitudinal empathy due to the monolithic nature of questions in this section unlike the one in attitudinal trust section and the satisfying results that linear regression yielded itself.

TABLE 16: Attitudinal Empathy for Group

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4,027	,107		37,759	,000
	GROUP	,319	,157	,268	2,041	,046

a. Dependent Variable: EMPATHY

4.1.3. Behavioral Empathy

The last outcome that this research interested in is that whether students who met through the workshop experience would have greater empathic behavior for members of the other ethnicity than non-participants after a year. Unlike the first two hypothesis that measured the acquisition of attitudinal skills, this hypothesis necessitated a behavioral exercise carried out by a hypothetical questions. The participants in the experimental and the control group were asked how much money they would donate for earthquake victims from the other ethnicity if there were an earthquake. The results of this exercises is presented in Table 17 below:

TABLE 17: Behavioral Empathy for Group, Gender, and Ethnicity**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	70,476	9,826		7,173	,000
	GROUP	16,657	10,020	,220	1,662	,102
	GENDER	-6,623	9,993	-,088	-,663	,510
	ETHNICIT	-15,195	9,993	-,201	-1,520	,134

a. Dependent Variable: DONATION

Third hypothesis suggested that the students who participated in the TGCD project would have donated higher amounts of money than non-participants. However, no significant difference was found between the two groups in this experiment. These results argue against Hypothesis in a sense that there were no meaningful differences detected in terms of group, gender, and ethnicity. Since the main statement in the hypothesis was concerned only with the differences between groups, experimental and control, another linear regression was used with group serving as the sole independent variable. The result illustrated below in Table 18 made no difference where there was no significant difference measured between the experimental and the control group.

TABLE 18: Behavioral Empathy for Group**Coefficients^a**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	59,567	6,864		8,678	,000
	GROUP	15,818	10,074	,209	1,570	,122

a. Dependent Variable: DONATION

Although the difference was not significant, the table below compares the means of the donations made in the experimental and the control group. This is important data showing that the levels of donations in both groups were indeed high but it was the difference between them which was not significant. At this point, it is safe to say that participants in both experimental and control groups had high levels of empathic behavior towards each other. Although there is a slight difference, it is not significant. In Table 19 below, (,00) was an indicator used to refer to the control group, where (1,00) indicated the experimental group.

TABLE 19: Mean of Donations in Experimental and Control Groups

DONATION * GROUP

DONATION

GROUP	Mean	N	Std. Deviation
,00	59,5667	30	41,88793
1,00	75,3846	26	31,90370
Total	66,9107	56	38,09427

To sum, the experiment supported the first two hypotheses of this study. This support means that the TGCD project used contact and peace education sufficiently, which led to a significant difference in the attitudes of trust and attitudinal empathy between the experimental and the control group. The participants of the TGCD, the experimental group, projected higher levels of attitudinal trust and empathy towards the member of the other ethnicity. The insignificant levels of difference when tested for gender and ethnicity as independent variables showed that the TGCD also succeeded to achieve the principle of *equal group status*, which is one of the four optimal conditions offered by Allport in order for contact to create positive results. (Allport, 1956)

The last table summarizes the results for hypotheses 1, 2, and 3 and provides means and standard deviations for each cell.

TABLE 20: Summary of Results
SUMMARY OF RESULTS

	<i>EXP</i>	<i>CNT</i>
Trust (Total)	3.2923 (0.54228)	2.9267 (0.53401)
a. Calculus-Based	3.0846 (0.71593)	2.7667 (0.60363)
b. Knowledge-Based	3.3173 (0.82024)	2.2250 (0.95896)
c. Identification-Based	3.4487 (0.69245)	3.5278 (0.77692)
Empathy	4.3462 (0.50456)	4.0267 (0.64484)
Amount Donated (Euros)	75.3846 (31.90370)	59.5667 (41.88793)

NOTE: Means are reported in each cell. Bold format indicate means that differ significantly across cells. Standard deviations are reported in parentheses. EXP= Experimental Group, Workshop Participants; CNT= Control Group, Workshop Non-Participants.

As I have indicated previously, the interesting findings on three different types of attitudinal trust, especially the knowledge-based trust, is worth future exploration. The details for future research on this issue are mentioned in Section 5.3.

4.2. Process Evaluation

While evaluating the process of the TGCD, there were two main tools: participant observation and qualitative interviews. As a result of the qualitative interviews conducted with the organizers of TGCD, this research found out that the TGCD did not start with a clear vision of their theories of change. Carol Weiss mentions this problem in her book arguing that this lack of a map outlining a program's expected theories of change has become a general trait for similar projects across time. Only after the interview with Dr. Savvides, who designed this project, the evaluator illustrated the programs theories of change with the data gathered from the interview, statements from the grant proposal submitted to the European Representation to Turkey, as well as the participant observation notes.

The results of the questionnaires that were administered by the organizers were also useful in analyzing the project. The questionnaires were used by the organizers after each meeting. Unfortunately, even the first administered questionnaire was also conducted after the participants have met. That is why those questionnaires were not a very reliable source of information in this study. However, there was one interesting question asking the participants whether they have interacted with anyone from the other national group or not. The results suggest that for the 76% of the participants, the TGCD was the first opportunity to meet with another student from the other national group. This fact is indeed important showing once again that contact was one of the major tools of the TGCD.

Initially, model created by Bruce Dayton (2005) was used as one of the major tools in mapping the theories of change applied by TGCD. This model has three major components. Dayton instructs the evaluator to put "assumptions about the underlying root(s) of the conflict being confronted" to the first category. At this point, TGCD identifies two different lines where the roots causes of the conflicts between Turkey and Greece lie. In the first line of thinking, the level of analysis is high politics. At this level, Savvides emphasized the stalemate status in Cyprus which poses a threat and an obstacle in improving the bilateral relations. According to Savvides, a resolution of the Cyprus issue would create a momentum for resolving the bilateral issues such as the Aegean, minorities and casus belli. However, the organizers indicated that TGCD did not put the focus on this line of issues even though preparing students to take active role in the future of Turkish-Greek Rapprochement, and pairing the publics in both countries with a readiness to accept compromises on core bilateral issues were stated in ELIAMEP's website as the goals of the project within the initial definition. The second line of issues, and the ones that TGCD aims to tackle, are related to the

socialization processes in both countries. Savvides argues that Greeks and Turks have been socialized to view each other as “us vs. them.” This construction of perceptions was fed throughout the years with the closed channels of communication and now poses as a threat against fruitful dialogue. This is the specific problem TGCD aimed to change in the beginning.

The second category in mapping programs’ theories of change comprises “beliefs about the conditions under which these root causes can be transformed.” Savvides outlines three main conditions to change perceptions of TGCD participants in a positive direction. One of these conditions is to change how the issues constituting the Turkish-Greek relations are defined. Second one is to change the way two groups view each other and the final one is to change the way each nation constructs their identity in terms of the importance they placed on the other. On the other hand, this data gathered from the interview addresses specific measures where the initial grant proposal includes measures in higher levels of decision making. The conditions mentioned in the grant proposal refer to long-term outcomes such as strong citizenry and pluralist civil societies in each country. The ultimate goal of the workshop is to help a network of young leaders emerge to play an active role in the Greek-Turkish Rapprochement. These goals, however, set in the beginning do not follow a logical action plan and models of project design, implementation as well as assessment. At this stage, I would like to refer to another model designed by Bruce Dayton. In his model of project design, he outlines seven steps:

- Conduct Conflict Analysis
- Develop Project Goals
- Conceptualize a Theory of Change
- Design Project Activities
- Implement Project
- Disseminate ‘Transfer’ Outcomes
- Assess Impacts

Although TGCD managed to conduct the first two steps in this model, they failed in conceptualizing a theory of change which forms the essential link between the second and the fourth step. The lack of this stage in TGCD resulted in numerous goals being listed without proper matches in activities. The project suffers from neglected goals that were set in the beginnings and the inadequate action plans to achieve these goals. To be more specific, although a goal such as flourishing strong citizenry and pluralist civil society both in Greece

and Turkey was set, no activity was designed to achieve this goal. In a linear path, the organizers acted with an assumption that building a common vision for the future among the participants would itself lead to the achievement of this goal mentioned above yet there was no explanation as to how this is possible. Such an absence of follow through activity poses as one of the main challenges for the evaluator in the sense that while creating a proper map to outline the programs theories of change there is a significant number of goals without any match in activities. This challenge also points to one major shortcomings of the TGCD project. As for sixth and seventh steps, even though the program required multiplying the beneficiaries of the program by disseminating the knowledge, which was definitely an anticipated outcome from the beginning, there were no mechanisms present in the project to achieve this goal. In addition, there was absolutely no assessment conducted to measure the effectiveness of the project. Briefly, TGCD was a project with inflated goals but without proper tools of implementation and assessment. The reasons behind the lack of such measures might be due to the financial constraints or organizational difficulties, as Dr. Savvides mentioned during the interview; however, the organizers should at least conduct separate meetings before the project to have a design with a strong theoretical framework. The academic profile of the organizers furthermore would assist in combining theory and practice under a single project.

The third and last category in assessing TGCD's theories of change comprises of the beliefs as to what kinds of programmatic interventions engender what kind of changes. Under this category lie the programmatic interventions used by TGCD in order to achieve the goals set in the beginning. Moving from general to specific, there are two workshops and one conference used as the general intervention tools. The workshops were held in Istanbul and Athens, respectively, providing upper level students with a chance to meet and work together with another student from another ethnicity, and a set of seminars dealing with various topics. Theoretically speaking, the types of intervention used in two workshops were contact and peace education.

Since the organizers did not choose to match these interventions with specific goals, it was the job of the evaluator to illustrate how each intervention might have helped realizing a certain goal. In order to exemplify my point, contact was helpful in opening channels of communication, leading to finding mutual understandings, and sharing a common vision for the future, which ultimately lead to the emergence of strong citizenry and pluralist civil societies in both countries. These paths of causal links were extracted from the description of

the project advertised in the website of the organizers, but were merely mentioned as the goals, furthermore, contact was not mentioned at all as the trigger behind all of these effects.

Peace education, on the other hand, was portrayed to be the driving force behind targeting upper level students who serve as players in the decision-making levels in the future as well as helping them to build conflict resolution skills which would ideally lead to the further development and developing of the Turkish-Greek Rapprochement.

The problem with these goals is that the causal links between the goals and outcomes anticipated by the organizers does not seem to have a scientific background. They indeed seem to be mere generalizations on the minds of the organizers. The organizers can indeed have macro-goals but yet these type of goals necessitate strong mechanisms of monitoring the transfer effects of the project. Unless such mechanisms are employed in the process, the goals cease to be inflated expectations of the organizers with no base.

Despite the general tools of intervention such as contact and peace education, TGCD used another tool to bring out desired changes: a great deal of emphasis was put on the joint papers produced by the participants. The organizers identified these joint papers as the final and tangible outcome of the project and started working on it from the very first meeting. In the first workshop that took place in Istanbul, the organizers set a list of current issues on Turkish-Greek relations debated in each country. The participants were asked to pick the topics they would like to study and the organizing committee formed the groups accordingly. As a result, they came up with eleven groups studying topics such as civil society, earthquake diplomacy, the importance of leadership, as well as the issues in Cyprus and the Aegean. Each group was arranged to include at least one Turkish and one Greek member in order to let them work on a shared assignment. Theory refers to this tool as setting a super-ordinate goal, a major component for the success of contact hypothesis. The emphasis on the joint papers was stressed in the second workshop in Athens as well where the participants were given free time to work on their papers together with their advisor professors assigned by the organizing committee. The final conference was the point when the participants realized their goals and presented their work in a conference held in Athens. The role of this intervention in the map of programs theories of change will be illustrated in the final figure, however, it can briefly be explained as the following: when young people come together, the organizers will try to make them understand how they can write a joint paper. With the help of the seminars, under the name of peace education, students will eventually learn a common vocabulary to be able to discuss the debated issues in both countries. These chains of goals and conditions would finally lead to the final goal of the organizers: two countries learn to accept compromises in

the core issues through these students serving as agents of change in their countries. Once again, the causal links and the goals set in the beginning is far beyond the reach and control of the organizers so the best policy advice at this point would be the importance of setting feasible goals to give participants a better sense of achievement at the end.

Under the theoretical framework created by Bruce Dayton and Esra Cuhadar-Gurkaynak (2005), derived from Carol Weiss (1998), the evaluator has prepared the figures bellow to illustrate the theories of change adopted by TGCD. This map is the result of the first part of this research where qualitative interviews and participant observation. The discussion above helped assessing where the project was successful to create the change they aimed in the beginning and where the project went astray.

Figure 1: Map of the TGCD'S Theories of Change

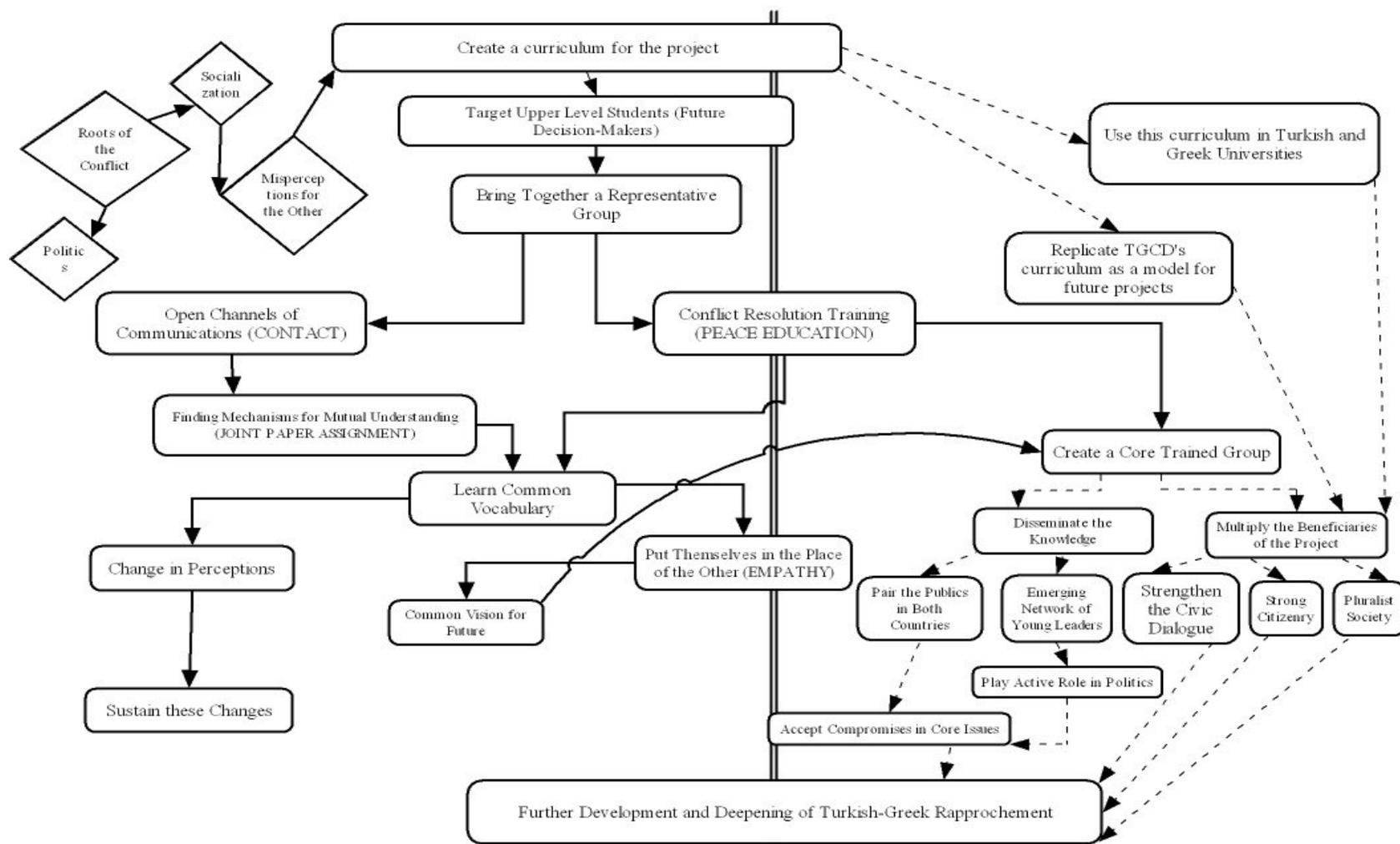
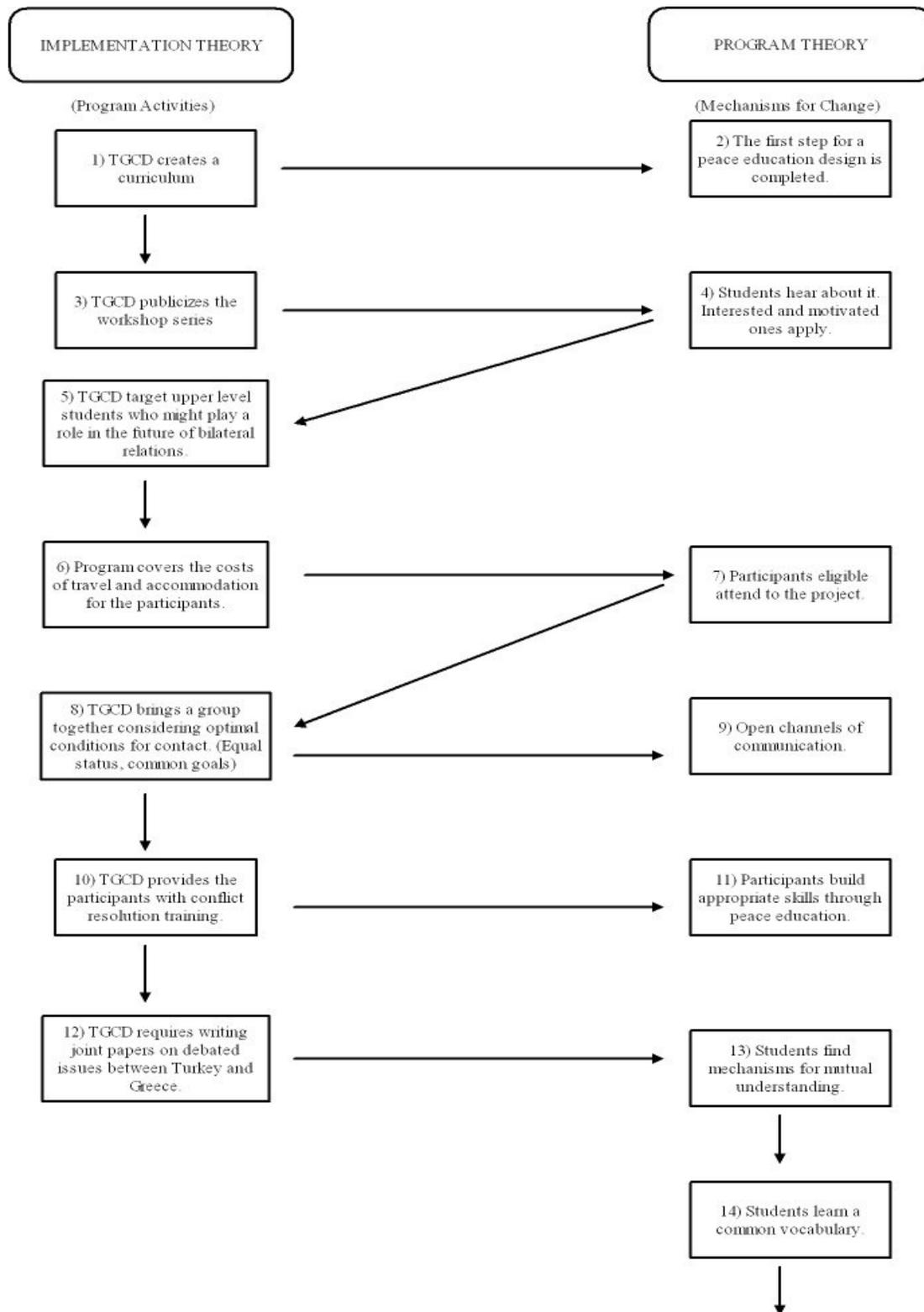
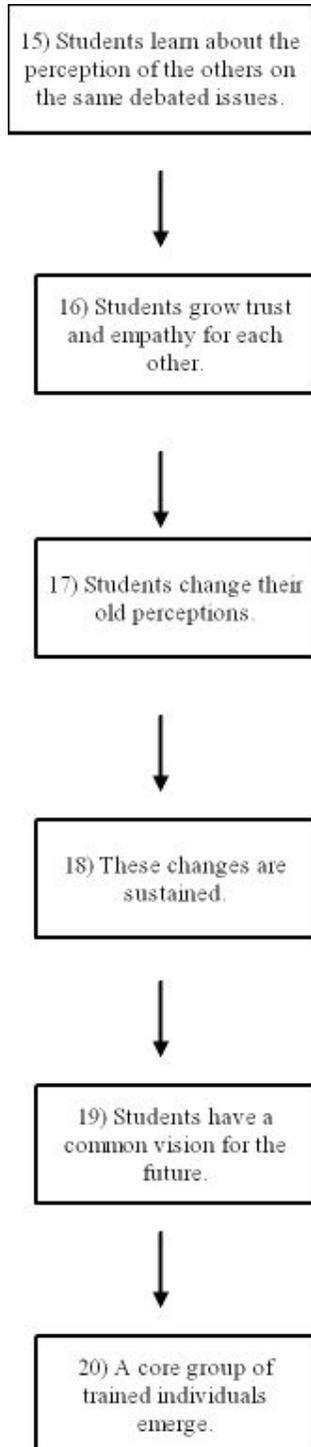


Figure 2: Implementation and Program Theory





The data reported in Figures 1 and 2 are retrieved from two main sources. The first source is the structured interviews with the organizers of the project. The second source is the official statement of the organizers used in the grant proposal. In the figures, it can be clearly seen that the goals at each level of the project are interrelated and thus, difficult to separate visually.

The goals of the project are grouped into two levels: micro-goals and macro-objectives. There is a double line separating Figure I into two. The left side of this dotted line refers to the micro-goals where the right side of the double line comprises of the macro goals of the project. Data concerning these macro objectives of the project are based on the official statements of the organizers as put forth in the grant proposal. The section of the Figure that shows the macro objectives are derived mostly from the official statement. On the other hand, the micro goals of the project reported in Figure 1 are derived from interviews conducted with the organizers of the project. The dotted arrows suggest that there is a gap between the macro objectives of the official statements and micro goals that came out of the interviews. The dotted arrows point the official goals where the continuous ones indicate the theory of change of the program that I retrieved from the interviews with the organizers.

5. CONCLUSION

After outlining the findings in Chapter Four, this chapter starts by discussing those findings and go on with recommendations both for the future initiatives and scientific research in line of evaluation.

5.1. Discussion of the Findings

Evaluation of the TGCD necessitated the use of multiple methodologies such as qualitative interviews, mapping and a field experiment. Each methodology yielded certain results that should be discussed further in detail. When qualitative interviews with the organizers were combined with participant observation, a detailed map of the TGCD's theories of change was created as one of the major outcomes of this study. (See figure 1 in 4.2.) The results drawn from this map can be summarized as following:

1. The TGCD had both micro and macro goals but did not have a proper design for tracking program's theories of change.

After mapping out the program's theories of change, which enabled the evaluator to see the expectations of the organizers, the experiment was carried out to see if the project was successful in achieving the micro-goals set in the beginning. The results from the experiment are as following.

2. There are significant differences between the experimental group and the control group with regard to the development of attitudinal empathy and trust.
3. There is no significant difference between the experimental group and the control group with regard to the development of behavioral empathy.

There is an important finding that stood up while analyzing the levels of attitudinal trust among the participants. Even though the levels of attitudinal trust differed significantly among the experimental and the control group, the factor analysis revealed that the levels of trust also differed when controlled for three categories: calculus-based trust, knowledge-based trust, and identification-based trust. The findings in Section 4.1.1 revealed that among these three categories, knowledge-based trust was the only type that yielded significant differences among the two groups used in this study. This finding indicates that, considering the sequential development of these three types of trust, the participants in control group and the experimental group possess similar levels of calculus-based trust. The reason behind this finding is actually two-fold. Regarding calculus-based trust, the insignificant levels of

difference might indicate that the participants of both experimental and the control group are equally moderate. Regarding the identification-based trust, it is not even surprising to see the insignificant difference between experimental and the control group since this type of trust is very hard to achieve and it usually takes a lot more time. However, the interesting finding is that with the second step in the sequential model of attitudinal trust, which is knowledge-based trust, the experimental group is proved to possess levels of knowledge-based trust significantly higher than the control group. This fact points to the success of the TGCD in providing its participants with significant levels of knowledge-based trust, which is one of the tangible outcomes of the project.

In this section, the aim is to compare the results from the mapping of the program's theory of change and the findings from the experimental design and see how the initial expectations of the organizers met the current outcome. A final result should also be added to the previous four with this comparison.

4. The TGCD succeeded in achieving the micro goal of creating a significant change among the participants' perceptions regarding trust and empathy.

Even though this study is an example of summative evaluation, which puts the focus on the outcome, the role of the evaluator as participant observer allows him to comment on the process of change occurred in the TGCD. Unfortunately, the organizers did not have a clear idea about certain mechanisms of change, which is reflected in the map created as a result of this study. That is basically due to the fact that the tools used for the TGCD, contact and peace education, were not operationalized properly. To be clearer, the organizers presumed that contact and peace education would serve as agents of change themselves without specifying any other activities. This shows that the project lacked an essential theoretical framework on these two concepts. Even though there were certain components of contact and peace education present in the TGCD, these did not seem to be planned beforehand by the organizers. At this point, participant observation proved to be the most effective tool in determining the process of change present in the TGCD.

One major observation was the emphasis the organizers put on the concept of peace education rather than contact. Even the title "Learning Conflict Resolution and Producing Peace" signaled the importance of peace education within the TGCD. However, the design for peace education did not seem to fit into any theoretical framework present in the literature. Even though the program aimed at *creating a curriculum* as the first step in the map, this curriculum was limited to a list of recommended readings. As for the activities, two sets of lectures were the only agents for change. The variety of these lectures also happens to be

problematic in a sense that they fail to cover most of the essential goals for peace education. A comparison of the peace education used in the TGCD and the essential goals is presented in detail in Section 2.2.

One other crucial point in evaluating implications of peace education used in the TGCD is that the theme of the lectures revolved around two major disciplines, International Relations (IR) and History. During the first meeting that took place in Istanbul, the participants were exposed to very theoretical aspects of war and peace along with a comprehensive historical background on Greek-Turkish relations. Despite the undeniable effects of IR and History on the discipline, Conflict Analysis and Resolution (CA&R) is a major discipline with its own conceptualizations of war, peace and security. Any attempt to exclude the CA&R from the program, contradicts with the title of the TGCD itself and would jeopardize the reliability of organizing institutions. This seems to be the main reason the organizers worked with a CA&R graduate, who was responsible for the practice-oriented sessions. Although these practice sessions were distributed equally among the two meetings, it was not very successful in capturing the attention of the participants. Most of the students did not seem to be responsive to these sessions.

This brings out another debate on peace education which is related to the conditions of learning. The organizers of the TGCD adopted the model of social learning, which argues that the participants of the TGCD would absorb the information given with the lectures. However, the skeptics support the cognitive-developmental approach, which mentions the political dimension of learning arguing that the participants will only internalize information that conforms to their belief system, ideology and values. (The entire debate can be found in section 2.2) Almost none of the participants of the TGCD held a strong political position, but that does not automatically refute the cognitive-developmental approach and support the social learning principle held by the TGCD. In this case, the quality of education seemed to replace the role of political affiliation regarding the internalization of the new information received from the project.

The last observation regarding the peace education is once again related to the theme of the lectures. Since IR and History scholars dominated the lecture scene, there was an unbalance between the macro-level and micro-level components of peace education. Even though the TGCD covered most of the macro-level components such as addressing the fears and understanding war behavior, it lacked to provide the participants with key micro-level components such as appreciating the richness of the concept of peace, teaching peace as a process and developing intercultural understanding. These terms also have an important role

in CA&R discipline and since the program promised its participants to provide them with conflict resolution skills, these terms should not be neglected.

With the flaws presented above, it can be argued that peace education was not the most effective tool that the TGCD used. Following the results of this study, depending mostly on my role as a participant observer, I can confidently argue that contact was the most influential tool that the TGCD used. Unfortunately, this tool has been undermined throughout the design and implication of the TGCD and since they did not use any evaluation mechanism, this valuable tool remained under credited. The interviews with the organizers revealed that the TGCD was indeed unaware of the fact that they were successful in creating Allport's (1954) four optimal conditions for contact. (See Section 2.1) These conditions and how the TGCD was successful in providing them is presented below:

1. *Equal Group Status:* The number of Greek and Turkish students was equal. The organizing committee as well as the guest speakers was also distributed fairly among Greeks, Turks and Internationals. The meetings also took place both in Istanbul and Athens. It is not only the number of the participants but also the group is equally distributed in terms of gender. It should be noted as well that the participants were from very similar age and social status groups.
2. *Common Goals:* Each participant had to write a joint paper with a colleague from the other national group on one of the major issues of Greek-Turkish relations with the supervision of the organizing committee.
3. *Intergroup Cooperation:* As mentioned above, each participant had to team up with a student from the other national group and work together. This cooperation theoretically allows the participants to de-categorize their identities dependent on their ethnicities and re-categorize under the common term of being fellow students. In other words, Greek and Turkish students cooperating to undertake an assignment would no longer see themselves as Greeks and Turks but students.
4. *Support of Authorities, Law or Custom:* Even though this condition is sometimes hard to acquire, the organizers of the TGCD managed to arrange receptions respectively in Greek Consulate in Istanbul and Turkish Embassy in Athens, where the participants had the chance to converse with the diplomats and share opinions. There was also another panel that brought two retired ambassadors together and presented the participants with a

discussion on the dramatic difference between the past and present of Greek-Turkish relations.

The activities under these four categories were actually the main activities that the TGCD planned in advance. When looked from this perspective, the TGCD stands out to be a successful case for contact intervention with every necessary condition. However, the organizers underestimated the role of contact and did not publicize contact as the main tool.

Pettigrew (1998) argues that there is no need to overburden the contact hypothesis with numerous other conditions but instead they should be used as facilitating conditions across different settings. However, he argues that there is one condition that should be added to the core four conditions, which is the role of generating affective ties among the participants. Pettigrew claims that cross-group friendship is a key factor in reducing prejudice and therefore must be included in the generic framework of contact hypothesis rather than a facilitation tip. (See Section 2.1) The role of cross-group friendship is an important factor that must be further analyzed in another research. As an outcome of my role as a participant observer, I can confidently say that most of the participants of the TGCD formed strong cross-cultural friendships that might be an important factor affecting the sustainability and even emergence of attitudinal trust and empathy. The organizers did not mention any activities to flourish friendship among the participant, which may be mainly due to the fact that they did not think of it as an important part of the intervention. However, organizers planned for informal activities after each day during the whole three meetings that allowed the participants to have fun together after a hard days work and form closer ties. Even though the organizers closed their channels of communication with the participants after the project, the students formed an electronic group where they were able to celebrate the victories and mourn for the tragedies together. The role of this fifth condition should be further analyzed with different cases in future research.

The last point of discussion regarding the effect of contact hypothesis on the TGCD is that in contact literature there is a current debate on the ultimate goals for contact. As outlined in the Literature Review (Section 2.1), the majority of scholars of contact hypothesis support the premise that attitudinal change is a prerequisite for behavioral change and future willingness to cooperate. On the other hand, there is Anna Ohanyan and John E. Lewis (2005) arguing that the overemphasis on attitudinal change would put serious limitations to achieve future willingness to cooperate and suggest that behavioral change can indeed occur before the attitudinal change. This prioritization is described as a matter of preference of the organizers conducting contact interventions and presents a major dichotomy between

prejudice reduction versus future cooperation as the goal of contact interventions. Even the research conducted by Ohanyan and Lewis was not applied to any other cases until this day, their point is relevant in the case of the TGCD in a sense that it should be obligatory for the organizers to have a clear idea on what they want out of contact situation. Depending on their expectations of an outcome from contact, they should build either a bottom-up approach where attitudinal change would gradually lead to behavioral change or top-down approach where the willingness among the participants to cooperate in the future would force attitudinal change. In the case of the TGCD, the organizers could not manage to choose a proper strategy and resulted in a confusion of macro and micro goals with an inadequate set of activities.

5.2. Recommendations

The interviews and the program's initial proposal indicate that the organizers are willing to conduct follow-ups for the TGCD with new participant. In this case, the following set of recommendations would be useful in improving the quality of intervention with the future projects. These recommendations are derived from the literature review, results and discussions sections of this thesis.

1. The TGCD definitely should have prepared a design for its own theories of change. Even though the map presented as a result of this thesis comprise elements of statements prior to the project, the interviews that were conducted a year after the project still served to be the major output for this map outlining the TGCD's theories of change. A proper map designed before carrying out an intervention would help the organizers see their shortcomings, achievements, and the points that need alteration. This map will be useful both during and after the intervention.
2. The TGCD necessitates a vigorous evaluation mechanism. Although the program distributed questionnaires to the participants after each meeting, the interviews revealed that the TGCD never used the results of this important feedback they received from the participants. An Action Evaluation, which is described in Section 2.3, would allow the organizers to control the process itself and make the participants feel that their opinions count as well and they are indeed the part of the process.
3. The TGCD should have limited its scope to include the micro-goals set in the beginning only. The macro-goals stated in the beginning, such as preparing the future leaders of Turkish-Greek rapprochement, are attractive

catch-phrases but at the end they would not seem realistic or feasible. The proper presentation of the project's micro-goals would also help flourishing a sense of achievement among the participants with the help of scientific evaluations.

4. Contact proved to be the most useful tool the TGCD used through the project. The emphasis on contact as the primary tool should be strengthened by linking the theory with practice. Given the fact that both organizing institutions were founded and run by internationally acclaimed scholars, the difficult task of combining theory with practice would be realized.
5. Peace Education is an important supplementary tool to contact. However, before having a strong theoretical framework, it is impossible to yield any meaningful outcome from peace education. The TGCD should carefully prepare a sound curriculum and a training guide for peace education with goals and activities separate from contact. In this case, their goals and activities were used interchangeably which challenged the validity of the outcome. It is important to note that peace education is a concept with many branches and conflict analysis and resolution is an important type of peace education. The TGCD should take this into consideration and design this tool based on the CA&R discipline rather than letting IR and History dominating it. Another note is that the practice sessions should better be held by different scholars, who have expertise in different spots of the field.

5.3. Future Research

There has been a lot of research conducted on contact hypothesis, peace education and their implications. However, this thesis managed to take a unique path to evaluate contact and peace education and their long-term outcomes through outlining a program's theories of change. This combination multiple methodologies, both qualitative and quantitative, is a new trend in social sciences and should be adopted in future research.

It is true that this research had its flaws and they are mainly due to the scarcities of certain resources such as time and money. The main flaw of this research was with the quasi-experiment. In a perfect design, there should have been a pre-test, an immediate post-test, and the final test after a years time. This figure can be illustrated as the following.

Pre-Test → Treatment → Post-Test (Immediate) → Post-Test (A Year After)

There was a questionnaire conducted by the organizers in the case of the TGCD, yet this questionnaire was applied only after the end of the first meeting. An ideal pre-test should have been administered before the participants met after the interview process to admit the eligible participants. The questionnaire used in this study covers the question whether the participants interacted with members from the other national group prior to the workshop experience, but still a more structured pre-test is a must for minimizing the risk of predisposed participants.

Predisposition is an important factor that future researchers should pay close attention to. A suggestion for the future researchers to test predisposition might be the alternative of having two control groups instead of one. The second control group should comprise the students who applied for the TGCD but could not pass the interviews. Naturally, having such a control group necessitates close collaboration of the organizers with the evaluator. In the case of this research, even though this collaboration was established, the organizers did not happen to have a list of participants who applied but could not get in due to the restraints in budget.

It would also be interesting to follow up on the implications of the findings on the three types of attitudinal trust mentioned in this research. Even though this research yielded certain results, a research design with an extended group size might produce an important output for the literature.

As I have been mentioning continuously, the use of multiple methodologies in evaluating the long-term effects of contact and peace education is a new feature of the field and must be tested in many other conflict settings. In doing such evaluation, it is crucial to see where the expectations of the organizers match with the outcome created among the participants. On the other hand, summative evaluation measuring only the outcome should not be the only type of evaluation used in future research. Process Evaluation is an important tool as well especially evaluating the contact interventions. Pettigrew outlines four main processes of change in contact situations (See Section 2.1) and future research can indeed test which process contributes to the final outcome the most. A gradual evaluation of contact can present the organizers with an invaluable data which shows the steps that yield results and the steps where the process is blocked.

Another important argument posited again by Pettigrew is the importance put on generating affective ties among the participants. Since empathy and trust were the main variables tested in this research, the role of cross-cultural friendship can be investigated

thoroughly whether it can serve as an optimal condition for contact with the rest of the core-four conditions of Allport.

As a conclusion of my two years work, I would like to say that having found significant differences on attitudinal empathy and trust a year after the TGCD ended, is both encouraging and fascinating keeping in mind that the TGCD had no contact with the participants once the project ended. I have tried to elaborate on the possible reasons for these sustained effects of the TGCD on its participants so that the future projects with similar goals can benefit from it. In the current environment, where the future of Turkish-Greek relations, and the so-called Rapprochement, is in a state of critical conjuncture, every project should act with a mission to contribute to a certain aspect of the issues between the two countries and improve the social dialogue which is crucial but at the same time very fragile. The projects like the TGCD should be aware of the full responsibility that the concept of workshop-fatigue is becoming a very alarming threat in conflict resolution interventions, and poorly conducted projects have the potential to harm the process. This is why every initiative should act with a properly planned design, implication, and assessment model which would help flourishing the bilateral relations between Turkey and Greece.

BIBLIOGRAPHY

- Allport, G.W. (1954). *The Nature of Prejudice*, Cambridge, MA: Addison-Wesley Publishing Co.
- Amir, Y (1976). "The role of Intergroup Contact in change of prejudice and race relations", *Towards the Elimination of Racism*, ed. PA Katz. New York: Pergamon
- Bar, H. And Bargal, D. (1995). "Living with the Conflict: Encounters between Jewish and Palestinian Israeli Youth", Jerusalem: The Jerusalem Institute for Israeli Studies
- Ben-Ari, R, Amir, Y (1986). "Contact Between Arab and Jewish youth in Israel", Miles Hewstone and Rupert Brown et al. *Contact and Conflict in Intergroup Encounters*, Oxford: Basil Blackwell
- Chen, Huey-Tsyh, & Peter H. Rossi (1980). "The multi-goal-theory-driven approach to evaluation: A model linking basic and applied social science", *Social Forces*, 59, 106-122
- Chen, Huey-Tsyh, & Peter H. Rossi (1983). "Evaluating with sense: The theory-driven approach", *Evaluation Review*, 7, 283-302
- Chen, Huey-Tsyh (1990). Issues in constructing program theory. In *New directions for program evaluation: Advances in program theory*, ed. Leonard Bickman, pp. 7-18. New Directions for Program Evaluation Series (no. 47) San Francisco: Jossey-Bass
- Cook, S.W. (1978). "Interpersonal and attitudinal outcomes in cooperating interracial groups", *Journal of Res. Dev. Educ.* 12:97-113
- Cuhadar, Esra & Bruce Dayton
- Elliott, Michael, Tamra Pearson d'Estrée and Sanda Kaufman. "Evaluation as a Tool for Reflection." *Beyond Intractability*. Eds. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder. Posted: September 2003 <http://www.beyondintractability.org/essay/Evaluation_Reflection/>.
- Jackman, MR, Crane M (1986). "Some of my best friends are black: interracial friendships and whites' racial attitudes", *Public Opinion Quarterly*. 50: 459-86
- Lewicki, Roy J. And Stevenson, Maura (1999). "The Trust Scale", Roy J. Lewicki, David M. Saunders, John W. Minton et al. *Negotiation: readings, exercises and cases*, 3rd edition, Irwin/McGraw Hill
- Lewis, Helen. "Evaluation and Assessment of Interventions." *Beyond Intractability*. Eds. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder. Posted: September 2004 <<http://www.beyondintractability.org/essay/evaluation/>>.
- Maiese, Michelle. "Setting Goals." *Beyond Intractability*. Eds. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder. Posted: September 2004 <http://www.beyondintractability.org/essay/setting_goals/>.
- Maoz, Ifat (2005). "Evaluating the communication between groups in dispute: equality in contact interventions between Jews and Arabs in Israel", *Negotiation Journal*, January, 2005.
- Nan, Susan Allen. "Formative Evaluation." *Beyond Intractability*. Eds. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder. Posted: December 2003 <http://www.beyondintractability.org/essay/formative_evaluation/>.
- Ohanyan, Anna, Lewis JE (2005). "Politics of Peace-Building: Critical Evaluation of Interethnic Contact and Peace Education in Georgian-Abkhaz Peace Camp, 1998-2002", *Peace & Change*, vol. 30, no. 1.

- Oliner, SP, Oliner PM (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press.
- Patton, Michael Quinn (1989). "A context and boundaries for theory-driven approach to validity", *Evaluation and Program Planning*, 12, 375- 377
- Pettigrew, Thomas F. (1982). "Prejudice", Stephan Thernstrom et. al. *Prejudice: Dimensions of Ethnicity*, (p. 1-30), The Belknap Press of Harvard University Press, Cambridge, Massachusetts
- Pettigrew, Thomas F. (1986). "The Intergroup Contact Hypothesis Reconsidered", Miles Hewstone and Rupert Brown et al. *Contact and Conflict in Intergroup Encounters*, Oxford: Basil Blackwell
- Pettigrew, Thomas F (1997). "Ingroup reappraisal: another intergroup process that reduces prejudice"
- Pettigrew, Thomas F. (1998). "Intergroup Contact Theory", *Annual Review of Psychology*, vol. 49, p. 65-85
- Powers, DA, Ellison CG (1995). "Interracial contact and black racial attitudes: the contact hypothesis and the selectivity bias", *Social Forces*, 74: 205-26
- Rothman, Jay. "Action Evaluation." *Beyond Intractability*. Eds. Guy Burgess and Heidi Burgess. Conflict Research Consortium, University of Colorado, Boulder. Posted: October 2003
<http://www.beyondintractability.org/essay/action_evaluation/>.
- Scriven, Michael. *Evaluation Thesaurus, Fourth Edition* (London, New Delhi: Sage Publications, 1991).
- Sherif, M (1966). *In Common Predicament*, Boston: Houghton Mifflin
- Smith, CB (1994). "Back and to the future: the intergroup contact hypothesis revisited", *Sociol. Inq.* 64:438-55
- Wagner, U, Machleit U (1986) "Gastarbeiter in the Federal Republic of Germany: contact between Germans and migrant populations", Miles Hewstone and Rupert Brown et al. *Contact and Conflict in Intergroup Encounters*, Oxford: Basil Blackwell
- Watson, G. (1947) *Action for Unity*, New York: Harper
- Weiss, Carol H. (1995) "Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families", In *New approaches to evaluating community initiatives*, eds. J.P. Connell, A.C. Kubisch, L.B. Schorr, & C.H. Weiss, pp. 65-92. Washington, DC: Aspen Institute
- Weiss, Carol H. (1997) "How can theory-based evaluation make greater headway?", *Evaluation Reviews*, 21 (4), 501-524
- Williams, RM Jr. (1947) *The Reduction of Intergroup Tensions*, New York: Social Sciences Research Council
- Wholey, Joseph S. (1987). *Organizational excellence: Stimulating quality and communicating value*. Lexington, MA: Lexington Books

LIST OF TABLES

TABLE 1: School Affiliations of the Participants in the Experimental Group (n=26).....	27
TABLE 2: Demographic Comparability of Experimental Group and Control Group.....	28
TABLE 3: School Affiliations of the Participants in the Control Group (n=30).....	28
TABLE 4: Attitudinal Trust for Group, Gender and Ethnicity.....	40
TABLE 5: Attitudinal Trust for Group.....	41
TABLE 6: Factor Analysis for Trust.....	42
TABLE 7: Scree Plot for Trust	43
TABLE 8: Calculus-based Trust.....	43
TABLE 9: Knowledge-based Trust.....	44
TABLE 10: Identification-based Trust.....	44
TABLE 11: T-Test for Three Types of Trust in Experimental Group.....	45
TABLE 12: Average Mean in the Experimental Group Regarding Trust.....	45
TABLE 13: T-Test for Three Types of Trust in Control Group.....	46
TABLE 14: Average Mean in the Control Group Regarding Trust.....	46
TABLE 15: Attitudinal Empathy for Group, Gender, and Ethnicity	47
TABLE 16: Attitudinal Empathy for Group.....	47
TABLE 17: Behavioral Empathy for Group, Gender, and Ethnicity.....	48
TABLE 18: Behavioral Empathy for Group.....	48
TABLE 19: Mean of Donations in Experimental and Control Groups.....	49
TABLE 20: Summary of Results	49

APPENDIX A

Learning Conflict Resolution and Producing Peace

Seminar I

Organized by the Hellenic Foundation of European and Foreign Policy (ELIAMEP) and the Istanbul Policy Center (IPC)

Funded by the Micro Project Programme for Turkish-Greek Civic Dialogue of the European Commission Representation to Turkey

**Istanbul, Turkey
Nippon Hotel
5-7 November, 2004**

Friday, November 5th

12:05 Arrival in Istanbul
13:30-16:00 City Tour
17:30 Meeting in hotel lobby for orientation – ALL STUDENTS
18:30 Depart for Consulate of Greece
19:00 Reception at Consulate of Greece

Saturday, November 6th

09:00-11:00 War and Conflict in History
11:15-12:30 Conflict Resolution Toolbox (Dealing with the Past, the Parties)
12:30-14:00 *Lunch*
14:00-15:30 Role Play
15:45-17:00 Mediation Process
17:15-18:00 Introduction to Topics for final Paper
19:30 Meet in lobby for dinner

Sunday, November 7th

09:00-11:00 Three Approaches to Peaceful Settlement
11:15-12:30 Conflict and National Identity
12:30-14:00 *Lunch*
14:00-15:30 World Mediating Actors
15:45-17:00 Personal Experiences
17:45-18:00 Assignment of Topics
19:30 Meet in lobby for dinner

Monday, November 8th

06:15 Depart from hotel to airport

Learning Conflict Resolution and Producing Peace

Seminar II

*Organized by the Hellenic Foundation of European and Foreign Policy (ELIAMEP)
and the Istanbul Policy Center (IPC)*

*Funded by the Micro Project Programme for Turkish-Greek Civic Dialogue of the European
Commission Representation to Turkey*

Athens, 3-6 December, 2004
Central Athens Hotel

Friday, December 3rd

09:50 Arrival in Athens (TK 1845)

12:30-16:30 City Tour

18:00 Meet in Hotel's restaurant

19:00 *The European Future of Turkey* a panel discussion organized by The Kokkalis Foundation and ARI Movement Brussels with the participation of **Mr. Kemal Dervis**, Member of Parliament, Turkey and **Ms. Anna Diamantopoulou**, Member of Parliament, Greece (The Hellenic American Union, 22, Massalias street).

21:00 Dinner at *Dirty Str-eat "Ginger"* (12, Triptoleμου str., Gazi)

Saturday, December 4th

09:00-10:30 Greek-Turkish Relations: Issues of Past and Present

Prof. Alexis Heraclides, Panteion University

10:30-11:00 *Coffee Break*

11:00-12:30 Greek-Turkish Relations: Prospects for the Future

Prof. Theodore Couloumbis, Director General, ELIAMEP

Prof. Ahmet Evin, Sabanci University/IPC, Onassis Fellow, Athens

12:30-13:30 *Lunch*

13:30-15:30 The Role of Public Opinion in Greek-Turkish Relations

Ms. Christina Bandouna, Consultant Partner at VENTRIS

Prof. Ali Carkoglu, Sabanci University

15:30-16:00 *Coffee Break*

16:00-17:00 Discussion of papers

19:00-20:30 *Reception hosted by H.E. the Ambassador of Turkey to Greece Mr. Tahsin Burcuoglu*

Sunday, December 5th

09:00-10:00 Personal Experiences

Prof. A. J. R. Groom, University of Kent

10:00-11:00 Simulation

Ms. Sansel Ilker, Research Fellow, IPC

11:00-11:30 *Coffee Break*

11:30-13:00 Greek-Turkish Economic Relations

Dr. Charalambos Tsardanidis, Director, Institute for International Economic Relations

13:00-14:30 The European Union and the Transformation of the Greek-Turkish Conflict

Dr. Panayotis Tsakonas, Ass. Professor, University of the Aegean

14:30-15:30 *Lunch*

15:30-18:30 Discussion of papers

18:30-19:00 Conclusions and Evaluations

Mr. Philippos Savvides, Research Fellow, ELIAMEP

Ms. Rana Zincir, IPC

21:00 Meet in lobby for dinner and music at *Lithos* (17, Taki and Aesopou str., Psirri)

Monday, December 6th

08:30 Depart from hotel to airport

10:50 Departure from Athens (TK 1846)

Learning Conflict Resolution and Producing Peace **Final Conference**

*Organized by the Hellenic Foundation of European and Foreign Policy (ELIAMEP)
and the Istanbul Policy Center (IPC)*

*Funded by the Micro Project Programme for Turkish-Greek Civic Dialogue of the European
Commission Representation to Turkey*

Athens, 18-20 February, 2005
Electra Palace Hotel

Friday, February 18

09:50 Arrival in Athens (TK 1845)

15:00-17:30 *Preparation Session*

20:30 *Dinner at “Palia Athina” Tavern (Nikis, Plaka)*

Saturday, February 19

09:00 Registration

09:30 Opening Remarks

Prof. Theodore Couloumbis, Director General of ELIAMEP

Prof. Ustun Erguder, Director General of IPC

09:45-11:30 *Dynamics of Crisis and Cooperation*

Chair/Discussant: **Dr. Thanos Dokos**, Director of Studies, ELIAMEP

Papers: “Sources of Friction in Greek-Turkish Relations: the Aegean Dispute”

Olga Borou and Egemen Ozalp

“Earthquake Diplomacy and Public Opinion: A Real or a Needed Rapprochement?”

Serra Makbule Hakyemez and Athanasios Theocharis

“International Crisis Theory and the Greek-Turkish Dispute Over Imia/Kardak
Islets: What Lessons for the Future?”

Yildirim Kayhan, Anastasios Sykakis and Ioannis Tsantoulis

11:30-12:00 Coffee Break

12:00-13:30 *The European Union and Greek-Turkish Foreign Policy*

Chair/Discussant: **Prof. Ustun Erguder**, Director General, IPC

Papers: “The Role of the European Union in the Cyprus Issue”

Anastasios Chatzivasileiou, Christina Christodoulidou, Murat Karaege and Ayse Kesler

“Greek-Turkish Rapprochement: The Role of Decision Makers”

Katerina Christodoulaki, Maria Ikonomaki and Genco Orkun

“The EU Role in the Greek-Turkish Rivalry and Cooperation”

Grigoria Kalyvioti, Devrimsel Nergiz and Panagiotis Sakkas

13:30-14:30 Lunch

14:30-16:30 *Crafting the Agenda for Greek-Turkish Cooperation*

18:30 *Greek-Turkish Relations: Then and Now*

Moderator: **Prof. Theodore Couloumbis**, Director General, ELIAMEP

Speakers: **Amb. Byron Theodoropoulos**

Amb. Iler Turkmen

21:00 *Dinner and music at “Mandra” (Lepeniotou 7, Psirri)*

Sunday, February 20

10:00-11:30 *Religion and Culture in Greek-Turkish Relations*

Chair/Discussant: **Dr. Hercules Millas**, Athens University

Papers: “The Greek and Turkish Arguments about the Ecumenical Character of the Greek Orthodox Patriarchate”

Pantelis Touloumakos and Akin Unver

“Religious and Cultural Perceptions of the ‘Other’ in Greek-Turkish Relations”

Sinan Ciddi, Lida Dimitriou, Gonul Evren and Muftugil Seda

11:30-12:00 Coffee Break

12:00-13:30 *Civil Society, Social Issues and Democratization*

Chair/Discussant: **Prof. A. J. R. Groom**, University of Kent at Canterbury

Papers: “Civil Society and Conflict Resolution: The Case of Greece and Turkey”

Theocharis Papadopoulos and Defne Paker

“The Asylum Policy of Greece and Turkey”

Elif Renk Ozdemir and Vasiliki Sotiropoulou

“The Democratic Impact of EU on Greece and Turkey”

Aspurce Onay and Vasilis Kyriazis

13:30-14:30 Lunch

14:30-16:00 *Evaluation and Lessons Learned*

16:00-16:30 Coffee Break

16:30-18:30 *Prospects for the Future: Presenting a New Agenda*

Monday, February 21

09:00 Departure of participants from hotel for “El. Venizelos”

APPENDIX B

FINAL INTERVIEW PROTOCOL FOR THE ORGANIZERS:

1. What do you believe is at the heart of the Greek-Turkish conflict?
2. What aspects of this conflict did your project try to address?
3. Why did you choose to address those aspects of the conflict?
4. What activities did you carry out in your project? [What did you do?]
5. What were your anticipated outcomes?
6. How did you think these activities would lead to the outcome you were anticipating?
7. Do you consider this project to be successful? If so, in what terms?

APPENDIX C

The Questionnaire

Instructions: Please answer the questions below on a 1-5 scale, 1 referring to strongly disagree and 5 referring to strongly agree. Please read each statement first and then click on the box below each statement in order to enter a number from 1 to 5 that resemble most how you feel about the preceding statement.

1 2 3 4 5

**Strongly
disagree**

Undecided

**Strongly
agree**

1. The way Turks behave does not bother me.

2. Turks are known as people who keep their promises and commitments.

3. Turks know that the benefits of maintaining trust are higher than the costs of destroying it.

4. Turks do what they say they will do.

5. I hear about the good “reputation” of Turks in keeping their promises.

6. I have interacted with Turks a lot.

7. I think I really know Turks.

8. I can accurately predict what Turks will do.

9. I think I know pretty well what Turkish reactions will be.

10. Greeks and Turks have a lot of common interests.

11. Turks and Greeks share the same basic values.

12. Turks and Greeks have a lot of common goals.

13. Greeks and Turks pursue many common objectives.

14. I know that Turks would do whatever we would do if we were in the same situation.

15. Greeks and Turks stand for the same basic things.

16. I would get very angry if I saw a Turk being ill-treated.

17. I could not continue to feel okay if Turkish people near me were upset.

18. It upsets and bothers me to see Turkish people who are helpless and in need.

19. I can understand how certain political issues might upset Turkish people very much.

20. I would get emotionally involved if a Turkish person that I knew were having problems.

B. Imagine that today you received a letter along with a 100 Euros check from Sabanci University, Turkey rewarding you for participating in this questionnaire. At the same time, you see the terrible news on your TV that another dramatic earthquake incident took place in Turkey that morning. There was also an aid campaign to help the earthquake victims. In this case, would you donate any money from the check you received from Sabanci University? If yes, how much would you donate?

C. Please specify the following points:

Age:

Gender:

Have you ever interacted with a Turk before:

Latest school graduated (or will graduate) from:

Thank you for you collaboration...

