"PISCO"

Mobilising gender and environmental awareness in high schools in rural El Salvador

Evidence from a RCT impact evaluation study
by order of *Consciente Switzerland*

Christoph Kühnhanss

January 25, 2019

Chair:	Prof. Dr. Ulf Liebe
Type:	Master Thesis in Sociology
Institution:	Institute of Sociology
	University of Bern
Author:	Christoph Kühnhanss
	Wasserwerkgasse 6a
	CH-3011 Bern
Phone:	+41 79 744 31 93
Mail:	christoph.kuehnhanss@soz.unibe.ch
Student ID:	79-705-851
Submitted:	January 25, 2019

Abstract

Education is central for human and societal development. While school enrollment rates in developing countries have considerably increased over the last decades, recent efforts focus on improving basic skills at primary school level, in particular literacy and numeracy skills. Other abilities and higher school levels are hardly ever addressed. In this study we have examined the effects of workshops that the Swiss NGO *Consciente* conducted at Salvadoran grammar schools in the summer of 2017. Workshops on either gender or environmental issues have been randomly assigned to 98 classes with a frame sample of 2848 out of a population of 5323 students. The innovative aspect of the program was the cost-effective deployment of young volunteers as *folk teachers*. By means of a baseline-survey before and an endline-survey after the intervention we measured changes of the cognitive, affective and conative components of gender and environmental awareness. The RCT-design of the implementation allowed causal inference of the treatment effects. Despite the low intensity of the treatments and the overloaded instrument, the analysis showed a few significant effects. The results do suggest that intensified educational interventions of this kind could produce considerable effects at very low costs.

Contents

1. Introduction and Background

1.1. PISCO in a Nutshell

In this paper we evaluate the impact of *environmental* and *gender* workshops that the Swiss NGO *Consciente* conducted in the summer of 2017 in 98 high schools classes in eastern El Salvador. The expectation and hope of the NGO was that a quite short intervention of half-day workshops could bring about considerable effects and changes in *environmental* and *gender awareness*. The most innovative element of the program was the cost-effective deployment of specially trained young volunteers as *folk teachers*.[1] Our scientific task was (1) to design the intervention in such a way that it would satisfy the criteria of a *randomized controlled trial* (RCT),[2] (2) to develop the measuring instrument, (3) to run the surveys and finally (4) to evaluate and interpret the collected data.

The two thematic workshops were randomly assigned to a random frame sample of 2848 out of 5323 officially registered students. By means of a base-line survey immediately before and an endline survey two weeks after the intervention the "PISCO Research Project"[3] measured changes in *cognitive*, *affective* and *conative* components of *environmental awareness* (EnvA) and *gender awareness* (GenA). We analyzed (1) cross-sectional mean differences in outcomes regarding EnvA and GenA after the intervention between the treatment groups and between the treated and the untreated. (2) We estimated difference-in-difference effects using a panel containing all individuals who participated in one of the assessments.

The basic expectation was that the environment workshop group would show higher EnvA scores and the group participating in the gender workshops higher GenA scores after the treatments. Since knowledge spillovers between the two topics are unlikely, students in the gender treatment functioned as control group for the environment treatment and vice versa. Due to the strict RCT-design, the differences of EnvA and GenA between the groups could be interpreted as the causal effects of the workshops.

Let us anticipate our main results in a nutshell: Unfortunately no differences between the treatment groups have been detected. However, we found significant instrument effects, i.e. differences between those students who participated in both the two surveys *and* the workshops and those who just participated in the endline survey. Despite the lack of significant treatment effects, the study provided important insights many of which are at least as interesting as the missing significance stars. We will discuss this in the results and discussion section.

This paper is organized in six sections:

- In section 2 we will present our population, i.e. who participants are and how they live.

---

[1] We translate the Salvadorian term "educadores populares" with "folk teachers". Alternatives would have been "lay teachers", "volunteers" or similar.

[2] For details see: Glennerster & Takavarasha (2013), Banerjee & Duflo (2011), Gneezy & List (2013).

[3] What does "PISCO" mean? Our investigation was the first major study of *Consciente*. It was baptized simply as follows: "*P*rimer *I*nvestigación *S*ocio-economico de *Co*nsciente" – PISCO.

- In section 3 we review the basic literature on EnvA and GenA as far as necessary, and have a look at the basic concepts *attitudes*, *values* and *intentions to act*. We discuss the theoretical framework of the investigation and formulate a brief theory of change.

- In section 4 we present the RTC-design of the implementation, the schedule, the data and our measurement instrument. In addition, we review the methods used in our data analysis.

- In section 5 we present the most important results.

- In section 6 we discuss the results and the problems that occurred during planning, implementation and data collection. There are many lessons learnt which are probably more important for the NGO's work than the significance stars of our models that we failed to find.

- But let us start with an introduction to the PISCO investigation and the folk teacher program.[4] Let us first take a look at our main stakeholders and the environment in which the project took place. In this introduction we try to present you in a more essayistic way the NGO and El Salvador, this small, little known country on the mountainous Pacific coast of Central America, which is historically burdened and struggling for its future.

1.2. The Problem to be Addressed

1.2.1. The Problem

What is the problem the NGO tried to address with its *sustainability education program*? El Salvador is a country marked by poverty, machismo, (gender)-violence, and rapid environmental degradation. Nevertheless, gender and environmental issues have not yet arrived in the daily discourse and consciousness of people, let alone the official school curricula. The education system is geared towards copying, memorizing and reproducing. Young people, for that reason, hardly develop the ability to critically question societal problems and to advocate for change. Our results show how little even high school students know and care about gender and environmental problems of their country, let alone of the planet. Reading books and newspapers is really not a common practice, the vast majority of people have no knowledge of English, and the Internet is generally used for entertainment and communication rather than information search.

Consciente looked for an efficient way to address these problems and to awake students awareness. Since school systems, teachers and curricula can not be changed with a mouse-click, the idea of trying it out with one's own, self-trained and highly motivated team of folk teachers is promising. The optimistic belief that a relatively small intervention might already achieve a lot was not just naive. Given the usually monotonous and inefficient frontal teaching in El Salvador, a short but intensive and didactically completely novel and surprising intervention might have a lasting effect. So the NGO decided to minimize

---

[4]   We will say *PISCO* when we talk of the research and we will say *RdE-project* when we talk about the "Red de Educadores"-project 2017.

the intervention's intensity and maximize the number of students to be attained. We will see hereafter to what extent this optimism was justified.

1.2.2. Scientific Context

In which scientific discourse can we embed the RdE-intervention and the PISCO evaluation? Scholars and actors in the field of development economics and sociology agree in the importance of education. An educated populace is necessary – though not sufficient – for economic and societal development. Education accumulates human capital which is essential for economic growth. That's why enormous amounts of public, institutional and private spending are invested in education (The World Bank, 2018, 184). Over the last decades, developing countries experienced sharp increases in enrollment and attendance rates (Glewwe & Muralidharan, 2015). So "most enrollment gaps in basic education are closing between high- and low-income countries" (The World Bank, 2018, 5) But did this considerable success improve education? Unfortunately, there are many reasons for doubt. "Schooling is not the same as learning" – the first sentence in the World Bank Report on learning get's to the heart of the problem (The World Bank, 2018, 3). Attending school and learning something are two different things. Learning outcomes remain considerably poor in developing countries.[5] The World Bank even talks of a "learning crisis" which mostly affects the most disadvantaged (The World Bank, 2018, 4).

Educational projects in developing countries try to improve educational systems in four fields: (1) demand side interventions, i.e. households', parents' and children's motivation to go to school; (2) school and students inputs (infrastructure, teacher salaries, textbooks, learning materials), (3) pedagogy, i.e. the technology of instruction; (4) governance (policies, educational systems, school management etc.) (Muralidharan, 2017, 330ff.)[6] Another dimension we may add is the school level an intervention is designed for. As far as we know most investigations primarily concentrate on projects that aim to improve students' basic literacy and numeracy skills at the primary and sometimes secondary school levels. The main concern of most efforts in developing countries is, of course, to get the children to go to school in the first place and teach them the basic skills. Nevertheless, these foundational skills are just the start. Education is not just about reading and calculating and it's not only about the production role of human capital: Education should increase an individual's capabilities (Sen, 1999) and its *agency* to make better choices. Education should make people flourish and give them the chances to live meaningful lives (Wright, 2010). Sure, without basic skills no higher education is possible, but school has to be more than that.

---

[5] Our own research on numeracy, conducted in El Salvador in 2018, shows worrying mathematical achievements by both primary school students and teachers. On average, teachers were only able to correctly answer 50% of the questions on 2nd and 6th grade mathematics material. See: `https://consciente.ch/en/cal-impact-project/`.

[6] A very similar classification of today's efforts to improve learning outcomes is made by The World Bank (2018, 13ff); for an overview of recent investigations and results see: Muralidharan (2017); for a review of education intervention reviews see: Evans & Popova (2016).

1.2.3. Higher Cognitive Skills

Few investigations measure other dimensions of education such as the achievement of higher cognitive and non-cognitive socio-emotional skills or the grade of awareness of local, national and even global societal problems. Moreover, little research is done on secondary or high school level. The RdE-project can be located in the field of improved pedagogy and enriched content as well as at the higher education level of grammar schools. So it perfectly fits the World Bank's claim: "Students entering the workforce need better critical thinking and socioeconomic skills. (...) [T]o navigate a rapidly changing world, they have to interact effectively with others, think creatively, and solve problems." (The World Bank, 2018, 164f.) Education makes people capable to adapt to ever faster changes of society, shifts in labor markets and innovations in technology. Since decades we observe worldwide urbanization, a shift from agriculture to manufacturing and industry, and from industry to services. The changes that the *Second Machine Age* or the *Fourth Industrial Revolution* will bring with them are hardly foreseeable, especially for countries like El Salvador.

Education does not only produces literacy, numeracy and some higher technical skills but also generates future citizens and members of society. Education has a positive effect on *civic agency*, on tolerance of diversity, pro-democratic attitudes, political engagement, protest activism and citizen interest. More educated people are better-prepared to participate and engage in societal and political activities and discourse. "Education increases awareness and understanding of political issues, fosters socialization needed for effective political activity, and increases civic skills." (The World Bank, 2018, 42) That's why Robert Putnam labeled education "the best individual-level predictor of political participation" (Putnam, 1995, 68) (Putnam, 2000). Analysis of the microlevel relationship between education and political participation such as discussing politics, attending lawful demonstrations, voting etc. show that more educated people people tend to exhibit higher levels of political engagement.[7] The contribution of educated people to social and political life can lead to more inclusive institutions by increasing critical observation of politics and imposing checks and balances on officials and institutions (Acemoglu & Robinson, 2012).

Evidence from more than 30 middle to lower income countries show that "[m]ore highly educated citizens are found to be significantly more tolerant of out-groups and more engaged in politics. They are also more likely to understand democracy in terms of free elections, civil rights, gender equality, and economic prosperity." (Chzhen, 2013, 1) Education can change individual attitudes, beliefs and values towards more tolerance. Educated people are more likely to participate in political and democratic processes and they can thereby strengthen the cohesion of society. Furthermore, they generally have better capacities to understand what is going on and how society and politics work.

Beyond that, education reduces the tendency towards crime, a very important issue particularly in El Salvador. It raises the health level and decreases mortality risk of individuals. Education can help to eliminate poverty of families and future generation. As we know today, education plays a crucial role in the reproduction of social inequality (Solon, 1999). Educated fathers and mothers increase their children's chances to better

---

[7] Even though, political participation is, of course, affected by country characteristics; see: Campante & Chor (2012)

education and higher incomes. The effects of education on gender equality are well documented. Gender equality is probably a major driver of development (The World Bank, 2011). Educated women run lower risks of teen pregnancy and have the capability to control their fertility rate because it sharpens the awareness of trade-offs of having children.

In summary: Uneducated people have little chances of adaptation to new developments and risk being left behind. They are more vulnerable and exposed to higher risks of poverty, disease and exploitation of any kind. They have lower political participation opportunities and worse chances on the job market. Without a good education, you really have a bad hand. That's why *Consciente* invests all its energy and resources in improving education.

But who and what exactly is *Consciente*?

1.3. The NGO *Consciente*

1.3.1. The Organization

Let me briefly introduce *Consciente* to you: *Consciente* is a Swiss non-governmental development organization founded in 2012. The association with its headquarters in Bern (Switzerland) and its partner organization *Fundación Consciente* in San Francisco Gotera (El Salvador) mainly works in Morazán, one of the poorest, most rural and remote departments of El Salvador.[8]

Consciente CH was found in 2012 by the young Swiss sociology student Martina S. Jakob from Bern, Switzerland, and some of her Salvadorian friends in San Francisco Gotera, among them the current director of *Consciente ES*, Jasael M. Torres. While volunteering as an English teacher for a couple of months 20-year-old Martina came to experience the daily school life in Salvadorian schools and the lousy quality of teaching. In El Salvador, lessons are usually limited to children copying abstract materials from the blackboard and then memorizing poorly understood study matters to reproduce them in exams. So *going to school* has little to do with *learning*, i.e. discovering new topics, asking critical questions, scrutinizing, finding own solutions, understanding complex matters, or, to put it in good enlightenment tradition: using your mind without the help of another. So Martina and her friends started a simple program of tutoring and extracurricular activities in a small school in San Francisco Gotera, the capital of Morazán. Since then the organization has grown successively. After the typical founders and start-up phase driven by idealism and youthful energy *Consciente* entered the transition phase in 2016 to become a professional NGO with rapid growth in turnover and activities. In 2018 the *Fundación Consciente El Salvador* has been founded. By the beginning of 2019 the organization has 10 permanent employees in El Salvador.[9]

---

[8] The correct legal name of the association in Switzerland is *Consciente - Unterstützungsverein El Salvador, CH-3014 Bern*. In El Salvador, the organization is called the *Fundación Consciente*, SV-San Francisco Gotera. In the following, the former will be called *Consciente CH*, the latter *Consciente ES*. We speak of *Consciente* if we mean the organization as a whole.

[9] For detailed information see the official website of *Consciente*: http://www.consciente.ch/.

1.3.2. *Consciente* Programs

Consciente describes itself as an "initiative for quality education for all" and thus sums up its most important activities. According to its official website *Consciente* today works with three different programs:

1. *Scholarship Program*: *Consciente* gives committed young people from poor backgrounds access to university, technical or high school education. The scholarships are financed by personal godparent-hoods primarily from Switzerland and Germany. In addition, since 2016 and 2017 the foundation runs two student residences in San Francisco Gotera for up to 20 marginalized adolescents from remote areas who could not afford daily transportation, room rent and livelihood outside their village and homes. Another residence is planned to be opened in 2019 in the city of San Miguel, the nearest university town.

2. *Education Innovation Program*: *Consciente* realizes projects to improve the quality of education in public schools. In 2018 the project "CAL-IMPACT" offered interactive computer-based mathematics lessons to about 2400 primary school children. *Consciente* developed an innovative pedagogical concept that combines individualized learning on the computer with games and group work. In autumn 2018 a pilot of the new project CATT (*C*omputer *A*ssisted *T*eacher *T*raining) was realized: 250 teachers were tested for their mathematical skills. The scores were devastating. The CATT-pilot will be realized in 2019.

3. *Sustainability Education Program*: *Consciente* uses participatory educational activities to promote and spread the awareness and critical discussion of important social issues such as gender inequality, violence, environmental degradation etc. It organizes additional teaching in schools and workshops, forums and congresses for students, teachers and representatives of other organizations, associations and NGOs.

Consciente describes itself as one of the first Swiss NGOs that consistently pursues an *evidence-based approach* in its project work. That means that *Consciente* first conducts a pilot trial whenever it makes sense. New projects are carefully designed and scientifically evaluated. The scientific methods applied make sure that the projects achieve the desired results. Based on the outcomes of the pilot projects, the designs are modified, improved and then, whenever possible, scaled up.

By the end of 2018 *Consciente* had approximately 71 members, 80 godparents and 80 scholarship holders, 20 students in two residences, and hundreds of private and institutional donators. In 2018 the budget is about 350 000 USD. The Swiss board consists of eight volunteers who work on a voluntary basis. In 2017 about 99.5 percent of the donations have been spent in El Salvador.

1.3.3. Sustainable Education Program

Let us have a closer look at the *Sustainable Education Program* in the context of which the RdE-project and PISCO were realized:

In the past four years *Consciente* has built up a network of about 30 young volunteers prepared as *folk teachers*. These volunteers were carefully trained in regular educational

camps and weekend workshops. Training focused on new creative didactic methods as well as societal issues such as gender inequality, violence and gender violence, environmental devastation and sustainability. In 2016 these folk teachers realized more than 200 one-day workshops in more than 40 high schools with about 4000 students. In December 2016, the board of *Consciente CH* decided to have the effectiveness of these workshops evaluated given the significant effort and costs of the RdE-project.

However, the deployment of folk teachers is a cost-efficient and low-budget form of knowledge dissemination in situations with scarce financial or human resources. The method is not completely new in El Salvador. Rather, it has been applied since the Civil War: In the refugee camps in Honduras during the civil war 1983 – 1992 Salvadorian refugees with exceptional knowledge like literacy, numeracy or foreign language skills (e.g. English) taught other refugees by simply sharing their knowledge at no charge. These activities were undoubtedly inspired by the success of the Cuban Literacy Campaign in the sixties, which was largely carried out by lay teachers. The folk teacher concept has several advantages:

- Folk teachers normally are volunteers and for this reason they are cheap or they even work for free. In the *Consciente*-program the team of folk teachers consists of (1) scholarship recipients who work off their compulsory social hours and (2) of young volunteers, mostly university students or high school graduates.

- Folk teachers may be more motivated and teach with more enthusiasm than established teachers on the regular payroll of the schools.

- Folk teachers are "people of the people" and they will disseminate their knowledge not only within the program but also in their entire social environment.

- The folk teachers of *Consciente* are virtually the same age as their students and may be perceived as their peers and therefore relate easier and establish an ambience of confidence.

On the other hand, some of these advantages also have their drawbacks:

- Perceived as peers, folk teachers may not be taken serious. They may lack the authority required to have their classes under control or may not be able to assert themselves.

- Folk teachers have little or even no pedagogic and didactic experience and they may hardly master the subjects as well as an older teacher with long experience.

- Folk teacher may stop teaching faster than employed teachers so that folk teacher training costs may be expensive in the long run.

1.4. El Salvador

At an international level, common knowledge about El Salvador is limited. If anything comes to peoples' minds when confronted with the name of the country, it is pictures of refugee treks to the United States or cruelly murdered people. Indeed, pictures of blood-covered tattooed corpses seem to be one of El Salvador's most important export goods.

El Salvador exhibits one of the highest homicide rates in the world. As a matter of fact, the "war of the gangs" – the "maras" – casts a dark shadow over Salvadorian society. But the negative headlines in international media make us forget that a 6.5 million people is struggling to build up a modern democratic state and find its place in the world after a long conflictual violent history of colonialism, imperialism, military dictatorship, ethnic cleansing and civil war.

1.4.1. The Maras

One of the predominant topics in modern Salvadorian society is the cancer of organized criminality. Large criminal gangs, the so-called "maras", earn their money with "normal" mafia businesses: protection racketeering and trafficking of drugs. The maras do everything necessary to assert their interests, including blackmail, violence and murder. Two Maras dominate the "market": The *Mara Salvatrucha* (*M-13* or simply *13*) and the *Pandilla Callejera* or *Barrio 18* (*M-18* or simply *18*). The media-effective face of the Maras in the international press are their young members with tattoos all over their bodies including their faces and the bloody corpses of young men murdered in every bloodthirsty way imaginable. The maras' war for territories is the main reason why El Salvador has been leading the list of the most criminal countries in the world for years. Since a couple of years the "13" and the "18" have been engaged in a bloody war with thousands of people killed and the endless spiral of retaliation and counter retaliation can hardly be broken. In 2015 the wave of killings pushed the murder rate above 100 for every 100,000 inhabitants, probably the highest in the world among non-war countries. According to UN-figures the rate dropped to 82.84 per 100,000 in 2016. The figures look shocking when compared with those for Switzerland (table 1, page 8).[10]

Table 1: UN OCSD: Homicide rate El Salvador and Switzerland

	2012		2013		2014		2015		2016	
	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
ES:	2,594	41.70	2,513	40.20	3,921	62.42	6,656	105.4	5,257	82.84
CH:	45	0.56	57	0.70	41	0.50	57	0.69	45	0.54

Source: UNODC United Nations Office on Drugs and Crime; http://www.unodc.org

Where do these gangs come from? In short: During and after the Civil War (1980 - 1992) thousands of Salvadorians fled to the United States. In the US mega cities like Los Angeles young men learnt how to survive by organizing themselves in gangs and fighting other gangs. After the war remigration began - unfortunately also that of the gangs. With the gangs came the battle for territories and control over lucrative businesses. Nowadays their membership figure is alarmingly. It is estimated at 70,000 gang members. Almost 13,000 are imprisoned in one of the overcrowded jails of the country. Possibly more than half a million Salvadorians depend on the income of the gangs, a considerable part of the total population. Hence, the maras are not a marginal problem.

---

[10] c.f. Reuters: http://www.reuters.com/article/us-elsalvador-violence-idUSKBN14M15F; or Swiss newspaper *Neue Zürcher Zeitung NZZ* https://www.nzz.ch/international/amerika/auf-dem-weg-zum-gewalttaetigsten-land-der-welt-1.18601176; last date accessed: 26.08.2018.

For more information on this sad subject, let me refer to the award-winning work of *Consciente*-board member Livia Jakob, who has traced the history of the Maras in the worth seeing story map *Pay or Die? – Prison or Cemetery? A Story about the Maras – The Mafias of the Poor in El Salvador*, winner of the ESRI contest 2018 as "Best Culture, History and Events Story Map 2018" (Jakob, 2018).

When working with people from El Salvador, you can *feel* the presence of the Maras in their daily lives. Nobody really likes to talk about the issue, it's the "Voldemort-word that no one says". But everybody has stories to tell about threatened, endangered or even killed neighbors, relatives and acquaintances. Everybody knows, where which Mara rules and which neighborhoods not to enter. When implementing its programs, *Consciente* is sometimes advised against going to certain schools or villages because "13" or "18" are controlling the region. The consequence of this uncanny, diffuse presence of the Maras is an immense annihilation of social capital such as trust, security and hope. The negative impact on economic development, investments, political stability, the process of democratization and the emigration and refugee rates can hardly be overestimated. Who invests, who opens a business, if he exposes himself to the risk of protection money extortion or even the threat of life and limb? Donald Trump's decision to expel 200,000 Salvadorans back to *their* country, which the younger ones have never seen, might further exacerbate the problem.[11] On the other hand, despite the huge problem of organized criminality in the country, we should never forget: There are millions of Salvadorians who simply try to make the best of their lives, for their families and children.

1.4.2. Facts and Figures

The figures in table 2, p. 10 are intended to give an overall impression of what El Salvador is like. The contrast to the same figures of Switzerland is striking. The country ranks on place 121 of 181 in the human development index HDI 2018. It belongs to the group of the "medium human development" countries together with his neighbors and countries like Vietnam, Bolivia, Iraq, Bangladesh. The per capita income is just 12% of that of Switzerland, so poverty is widespread. Maternal mortality is 10 times higher than in Switzerland. The trust in national government is low and perceived corruption very high.[12][13] The low feeling safe score reflects the presence of the Maras and the everyday violence.

1.4.3. Short History and Current Situation

History is important to understand the present. However, tracing El Salvador's history "from the Conquista until today" would be beyond the scope of this paper.[14] However, some important cornerstones of the Salvadoran history are briefly mentioned here to

---

[11] NYT, 04.01.2018: https://www.nytimes.com/2018/01/08/us/salvadorans-tps-end.html; last date accessed: 16.08.2018

[12] Transparency International: http://www.transparency.org/cpi2015#results-table; last date accessed: 03.01.2017.

[13] Transparency International:: http://www.transparency.org/gcb2013/country/?country=el_salvador. Newer figures are not available.

[14] See for example the extensive article on Britannica Online: https://www.britannica.com/place/El-Salvador; last date accessed: 17.01.2019.

Table 2: Human Development Indices El Salvador – Switzerland

Index	El Salvador	Switzerland
Human Development Index 2018:	0.674	0.944
Rank in 189:	121	2
Gross national income per capita:	6,868	57,625
Expected and mean years of schooling:	12.6 / 6.9	16,2 / 13.4
Education quality satisfaction:	70%	83%
Health care quality satisfaction:	59%	93%
Trust in national government:	31%	79%
Feeling safe:	36%	87%
Satisfaction about actions to preserve the environment:	31%	84%
Gender equality index GEI:	0.384	0.04
GEI-rank:	84	1
Maternal mortality ratio (deaths per 100,000 live births):	54	5
Share of women in parliament:	32.1	28.9
Female labor force participation:	49.1	62.7
Male labor force participation:	79.2	74.8
Adolescent births per 1,000 women ages 15-19:	65.2	2.9
CPI Corruption perception index	33	85
CPI rank 2017 (2016) of 180	112 (95)	3 (5)

Sources: UNDP United Nations development programme; http://hdr.undp.org

Transparency International - the global coalition against corruption; http://www.transparency.org

better understand the current context of our project. El Salvador shares the history of Spanish conquest and subjugation in the 16th century with other Central and South American countries. The country became independent from Spain in 1821 and was part of the "Federal Republic of Central America" from 1824 – 1841. The 19th and 20th centuries were characterized by the political struggles of small elites for control and power, which mostly resulted in unstable, short-living dictatorships supported by military forces. It is difficult to keep track of constantly changing governments in the 19th and beginning 20th century. In December 1931, a coup d'état organized by junior officers led to one of the pivotal events in modern Salvadoran history: the uprising of Salvadoran peasant in February 1932. The rebellion was headed by Farabundo Martí who later became the eponym of the left-wing party FMLN, which is in power today.[15] The riot resulted in bloody repression by the government that was later referred to as "La Matanza – The Massacre". Tens of thousands of peasants were killed, many of them indigenous.

Another bloody event is particularly important for the border department of Morazán: the so-called *Football War* with Honduras in 1969. It only lasted a couple of days, but as many as 130,000 Salvadorans who lived in Honduras were forcibly expelled or streamed

---

[15] FMLN: Farabundo Martí National Liberation Front Frente Farabundo Martí para la Liberación Nacional; Britannica Online: https://www.britannica.com/topic/Farabundo-Marti-National-Liberation-Front; last date accessed: 17.01.2019.

back to their country. But the most important event in recent history, of course, had much deeper consequences: The *Salvadoran Civil War* between 1979 and 1992. The UN reports that more than 75,000 people were killed, and thousand were wounded and left disabled. Besides the terrible human suffering of direct war victims the country experienced a complete economic collapse. The war was triggered on March 24, 1980, when Monseñor Romero, Archbishop of San Salvador, was killed by a death squadron while holding a mass. Several major guerrilla groups of the Salvadoran left formed the FMLN in October 1980 and organized the rebellion against the right-wing regime. On December 1981 the US-trained *Atlacatl Battalion* was responsible for the "El Mozote massacre" in the North of Morazán where more than 800 civilians were murdered, over half of them children. A frightening monument in Mozote keeps the memory of atrocity alive to this day. Only in January 1992 the parties signed the "Chapultepec Peace Accords" and two years later the UN published the "Truth commission for El Salvador" report.[16] In 1993 the national parliament declared a general amnesty for war crimes committed before 1992.

Unsurprisingly, the civil war has naturally left deep wounds in Salvadoran society to this day. None of the parties had achieved a victory, but the country lay in ruins and was still divided into two hostile camps. I visited the country for the first time in 1995 and was horrified by the state it was in. What followed was a kind of restoration and sluggish rebuilding. A democratic system has been established, but it is still in its infancy.

Between 1989 and 2009 the right-wing "Nationalist Republican Alliance ARENA" won presidency in every election. In the 2009 elections the tide turned against ARENA and Mauricio Funes became the first FMLN president. In 2014 he was succeeded by the ex-guerilla official Salvador Sánchez Cerén. In the 2015 National Assembly elections ARENA and its coalition partners won the majority again and, together with their allies, became the strongest party. This has led to a legislative stalemate and a government blockade.The next elections in June 2019 are fiercely contested.

What does all this mean for our project and investigation? In summary: The history of El Salvador is bloody and violent. The civil war is not too far away and the various massacres and ethnic purges are vividly anchored in the people's memories. El Salvador is a place of great divides – between right and left, rich elites and masses of poor, large landowners and dispossessed. The rift runs along the line of the former warring parties. Simply put, the parents of our students lay opposite each other in the trenches of the war. Most families had victims to mourn and everyone knows to tell stories about the atrocities of the "others". The present situation poses great challenges for the development of a strong civil society. Democracy is young, the governmental institutions are weak, corruption is a plague, poverty, violence and a lack of prospects are huge problems, especially for the young.

That is the context in which *Consciente* works and in which the RdE-project and PISCO were realized.

---

[16] Truth Commission for El Salvador; United States Institute for Peace: https://www.usip.org/publications/1992/07/truth-commission-el-salvador; last date accessed: 20.01.2019.

2. Population Description

So let us take a closer look at our population: Who are the girls and boys we examined? Where and how do they live? What is their socio-economic situation? We collected comprehensive socio-economic data in the baseline questionnaire (W1).[17] In W1 we surveyed 2242 participants, 1176 (54%) girls and 1009 (46%) boys.[18] In the following, we will present the most important variables – i.e. those we assume to have the strongest influence on the students' EnvA and GenA: variables about students' parents and family, religion and socio-economic situation.

2.1. Family

Not even half of our students live with both of their *parents*. Almost 12% live without any of them, over 47% without their fathers. Poverty, irresponsibility or machismo are possible explanations as to why so many fathers leave their children. However, we do not have any data to further explore this matter.

We asked for *siblings* and *household members* to estimate family- and household-size. The mean household consists of 5.22 people. The richer the household the fewer people live in it and the lower the portion of very large households (>8 persons) (table 3 p. 12).
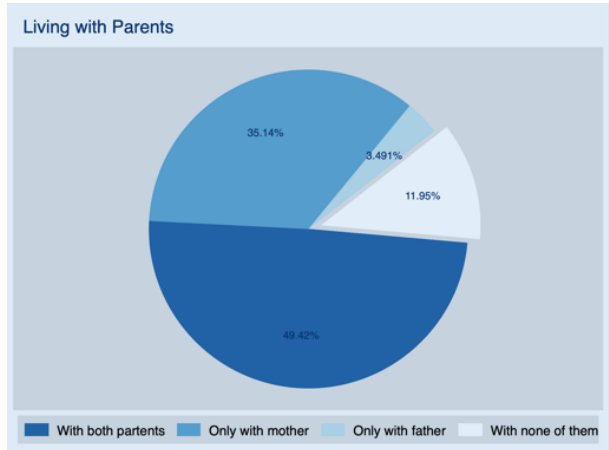
Figure 1: Living with Parents

Table 3: Household Members by Poverty Level

	Very poor	Poor	Somewhat poor	Not poor	Total
0-4	45	168	452	176	841
	39.47	39.53	38.31	43.67	39.63
5-8	54	231	650	211	1146
	47.37	54.35	55.08	52.36	54.01
>8	15	26	78	16	135
	13.16	6.12	6.61	3.97	6.36
Total	114	425	1180	403	2122
	100.00	100.00	100.00	100.00	100.00

---

[17] For details see codebooks in the *07_Questionnaires*-folder of the documentation.
[18] The figures are rounded to the nearest 0.5-digit amount to make the text more readable.

Richer families tend to have less children than poorer families (table 4 p. 13). The share of individuals with 5 – 8 siblings drops from 25% for the very-poor to 13.5% for the non-poor. The share of students with 0 – 4 siblings raises from 70.5% for the very-poor to 84% for the non-poor.

Table 4: Siblings by Poverty Level

	Very poor	Poor	Somewhat poor	Not poor	Total
0-4	81	307	887	340	1615
	70.43	72.24	75.68	83.74	76.25
5-8	29	100	237	55	421
	25.22	23.53	20.22	13.55	19.88
>8	5	18	48	11	82
	4.35	4.24	4.10	2.71	3.87
Total	115	425	1172	406	2118
	100.00	100.00	100.00	100.00	100.00

Let us now take a look at the important level of education of the students' mothers and fathers. As discussed, the educational attainment children will achieve usually strongly depends on the educational attainment of their reference persons. In environments characterized by low educational attainments, the young people tend to receive little academic support. If we assume that the parents of our 15 to 18-year-old students belong to a cohort that was socialized during the civil war and had to go to school in the '80s and '90s, the results shown in figure 2 p. 13 are not surprising: The vast majority of our students' parents do not have any higher education degree. Almost half of them did not even finish primary school.
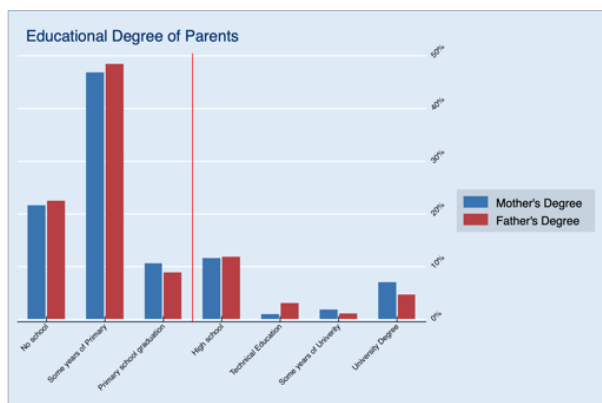
Figure 2: Parents' Education Grade

Furthermore, we can see in tables 5, p. 14, and 6, p. 14, that the disadvantaged children of poor families are confronted with even lower levels of parental education than their richer peers. 83% of the very poors' fathers and 82% of their mothers did not even finish primary school. In the very-poor and poor households very few parents have an education degree above the primary school level.[19]

---

[19] The values in the categories *High School*, *Some years of University* and *University Degree* in the very-poor column must be considered with caution as the case numbers are small.

Table 5: Father's Degree by Poverty Level

Father's Degree	Very poor	Poor	Somewhat poor	Not poor	Total
No school	37.89	32.47	20.86	11.20	22.15
Some years of Primary	45.26	49.35	51.79	40.00	48.67
Primary school graduation	2.11	5.71	9.40	12.27	8.86
High school	14.74	8.31	11.28	16.00	11.78
Technical Education	0.00	2.60	1.97	7.20	3.02
Some years of Univerity	0.00	0.78	0.94	1.60	0.99
University Degree	0.00	0.78	3.76	11.73	4.53
Total	100.00	100.00	100.00	100.00	100.00

Table 6: Mother's Degree by Poverty Level

Mother's Degree	Very poor	Poor	Somewhat poor	Not poor	Total
No school	35.51	28.50	21.28	10.49	21.43
Some years of Primary	46.73	51.93	47.73	38.11	46.68
Primary school graduation	6.54	8.45	11.13	13.30	10.76
High school	6.54	7.00	11.13	18.16	11.40
Technical Education	0.00	0.48	0.62	2.30	0.88
Some years of Univerity	0.93	0.72	2.05	2.30	1.77
University Degree	3.74	2.90	6.06	15.35	7.08
Total	100.00	100.00	100.00	100.00	100.00

2.2. Religion

A topic that is particularly important for our analysis of gender awareness is the distribution of confessions and the importance of religion. Generally speaking, El Salvador is a very religious country, strongly influenced by the catholic church. The liberation theologian archbishop Óscar Romero, who was murdered in San Salvador on March 24, 1980, is still venerated like a saint, especially after his canonization on October 14, 2018. The country has one of the toughest anti-abortion laws in the allowing to put women behind bars up to 30 years for abortions – which unfortunately are often tragic cases of miscarriages.[20]

We would expect that in the poor and rural district of Morazán religion plays a mayor role in the life of our students and that Catholicism and conservative evangelical religions are even more widespread than in other regions and in the urban center. Religiosity might influence students' basic values and their attitudes towards gender roles or environmental issues. Almost half of the students are catholic, nearly one third are evangelical-protestant, and amazingly, about 17% say they do not officially belong to any church. We also asked our students for the importance of religion in their lives and in their education (figure 3, p. 15).

---

[20] See e.g. article in *The Washington Post*: https://www.washingtonpost.com/world/the_americas/women-serving-decades-long-prison-terms-for-abortion-in-el-salvador-hope-change-is-coming/2018/09/26/0048119e-a62c-11e8-ad6f-080770dcddc2_story.html?noredirect=on&utm_term=.9386172ae6d3; last date accessed: 20.01.2019.
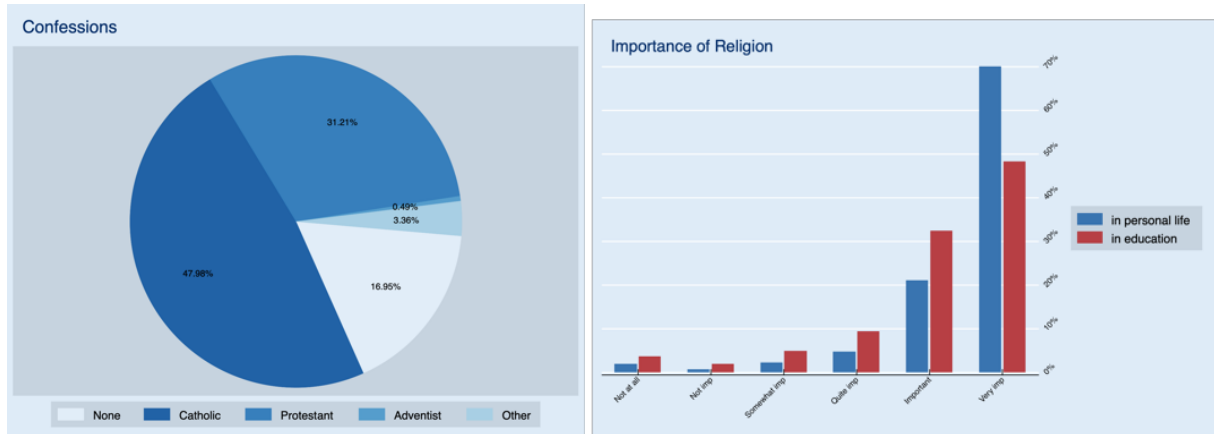
Figure 3: Confessions and Importance of Religion

We can only speculate why over 90% said that religion is important or very important in their lives, while 17% say that they do not belong to any church at all. We guess that for some students the confession question was just too private so they refused to answer. For girls religion plays a greater role than for boys but the difference is not substantial. As we expected, in the rural areas (Rural Center 2, Semi rural and Rural) religion plays a significantly greater role than in the Urban Center and the Rural Center 1. We can assume that these differences are also reflected in attitudes towards gender equality and roles – and possibly even in perceptions on environmental concerns. Somewhat confusing is the low importance of religion in the Rural Center 1, which should be examined more closely. It has the lowest high and the highest low score (table 7, p. 15). We investigated the importance of religion by poverty levels, without finding significant differences (table 8, p. 15).

Table 7: Importance of Religion by Strata

	Urban	RC 1	RC 2	Semi urban	Semi rural	Rural	Total
Not at all	1.66	5.11	0.71	2.48	1.72	0.74	1.87
Not imp	0.74	1.28	0.00	0.31	0.25	0.74	0.55
Somewhat imp	2.39	3.40	0.36	4.04	1.23	1.49	2.10
Quite imp	6.08	9.79	4.63	4.66	2.21	1.98	4.61
Important	89.13	80.43	94.31	88.51	94.59	95.05	90.88

Table 8: Importance of Religion by Poverty Level

	Very poor	Poor	Somewhat poor	Not poor	Total
Not at all	1.75	1.86	1.61	2.72	1.88
Not imp	1.75	0.70	0.25	0.99	0.56
Somewhat imp	0.88	1.17	2.71	1.98	2.16
Quite imp	4.39	6.53	3.90	4.44	4.56
Important	91.23	89.74	91.53	89.88	90.84

2.3. Economic Situation

A brief look at the economic standing of our students reveals an undeniable situation of widespread poverty. Although they come from the comparatively privileged families who can afford to send their children to grammar schools, most of them report being poor.

We measured their economic situation with one direct question and with a set of indirect questions about everyday items students might possess. We asked them to which degree they considered themselves as poor or not-poor on a 5-level scale from *extremely poor* to *not poor*. According to our own experience on the ground, Salvadorans (or probably poor people in general) tend to be too proud to admit extreme poverty. The lowest category gave them the opportunity not to place their cross at the extreme end of the scale while still reporting to be "very poor". As ex-

Figure 4: Overall Poverty

pected, only 1.5% marked this extreme option. For our analysis we combined the two lowest categories. Results are shown in figure 4, p. 16: 19% of our students said they were 'not poor', 5.5% said they were 'extremely' or 'very poor' and more than 75% said they were 'poor' or 'somewhat poor'.

Table 9: Poverty by Stratum

Stratum	Very poor	Poor	Somewhat poor	Not poor	Total
Urban	5.76	17.29	54.28	22.68	100.00
RC 1	4.70	18.80	57.26	19.23	100.00
RC 2	5.84	22.26	53.65	18.25	100.00
Semi urban	3.06	14.98	53.82	28.13	100.00
Semi rural	5.21	20.84	58.81	15.14	100.00
Rural	7.12	26.72	55.47	10.69	100.00
Total	5.39	20.10	55.51	18.99	100.00

Table 10: Poverty by Sex

Sex	Very poor	Poor	Somewhat poor	Not poor	Total
Male	6.00	21.36	53.00	19.63	100.00
Female	4.58	18.94	57.71	18.77	100.00
Total	5.24	20.07	55.52	19.17	100.00

Poverty differed slightly between the geographical strata. As expected, students from rural areas were the least likely to declare themselves as 'not poor'. The richest students, on the other hand, do not come from the urban centers, but from the semi-urban regions.

There is no significant difference in the poverty rate between the sexes. Boys tend to declare themselves somewhat poorer as girls. It is difficult to say whether they perceive their poverty more intensively than girls or whether they actually live in poorer households.

Let us have a look at the *household equipment* that students have in their homes. This is an indirect way to getting a more vivid picture of the economic conditions they live in. We asked for electricity, television, refrigerators, cellphones and smartphone, permanent internet, computers, washing machines and other items like tablets and solar panels. Figure 5, p. 17 shows the distribution of the eight most important household items of facilities on the left and of personal equipment on the right:
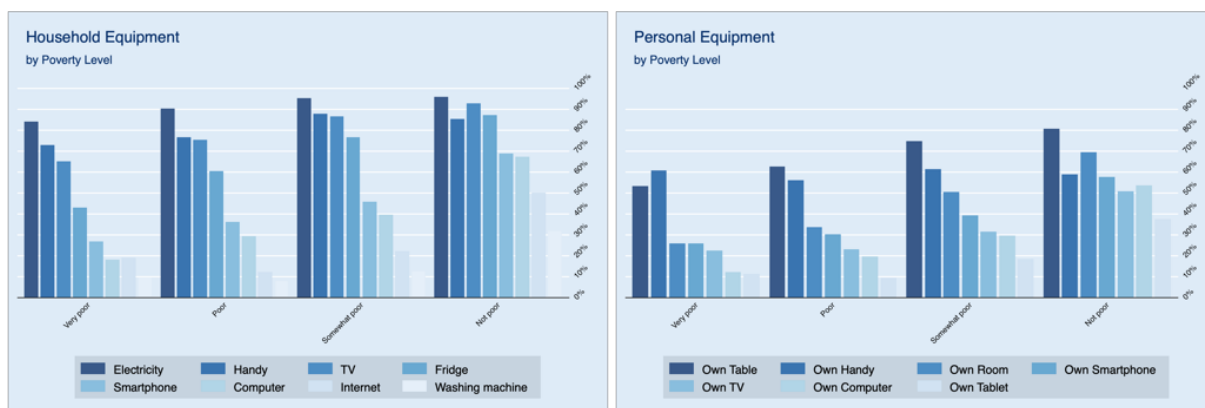
Figure 5: Equipment by Poverty

More than 16 percent of the 'very-poor' households still have no electricity, but above all they lack important sources of information such as permanent internet, smartphones or computers. Still more than 30% of the 'not-poor' say they had no computer or smartphones in their homes and about 50% have no internet.

Although electricity is widespread in Morazán that doesn't mean that households are generously equipped like households in Switzerland with several lamps in every room, heaps of electrical equipment and sockets in every corner. In most poor huts in the countryside or suburbs, electricity means nothing more but a weak bulb on the ceiling of the one room with a tamped clay floor – and this bulb is mostly switched off.

Table 11: Energy for Cooking

Energy - Poverty	Very poor	Poor	Somewhat poor	Not poor	Total
Wood	29.82	21.08	14.02	6.98	14.96
Gas	18.42	18.74	20.95	38.90	23.75
Wood Gas	46.49	59.72	63.09	50.37	59.13
Electricity	1.75	0.23	1.27	2.74	1.36

This explains why most households still cook with either wood or gas. The great need of firewood for cooking may be one of the reasons for the rapid deforestation in El Salvador (c.f. table 2, p. 10). Almost 60% of the households use wood and gas, 24%

only gas and 15% only wood ([11] p. [17]). The gas is not retrieved through a sophisticated gas pipeline network, but from gas cylinders, and the wood is collected in the retreating forests. Electricity is used for cooking by not even 1.5% of the households. The poorer a household the more likely it is to use wood as compared to gas or electricity.[21]

If we look at the *personal equipment* of the students, it is noteworthy that almost half of the poorest students (47%) do not have their own study desk, let alone their own room (75%). Even among the nonpoor, 30% do not have their own room (table [5], p. [17] on the right). Since schools are not generally equipped with study rooms for their students, they have no place to learn concentrated and undisturbed.

Table [6], p. [18], shows the *means of transportation* of households and students. Just 50% of the non-poor households possess a car and merely 28% own a motorcycle. That means that a high proportion of the students and their families depend on public transport which can be a problem in remote areas, especially in times of heavy rainfall. For example in October 2018, all schools throughout the country had to be closed due to rainstorms and floods. Public transportation in Morazán means riding on covered or even open pickups with narrow side benches on which a small proportion of the passengers can sit. The rest is tightly packed in the middle of the loading area, clinging to the struts to avoid falling out of the car in the numerous bends. As adventurous as this may sound to European backpackers, it is very burdensome for the locals who have to endure it on a daily basis, and is particularly hard for old people. Being on the road for two or three hours on such a pickup is exhausting and anything but funny. Most of our students walk to school or take the public bus. The richer they are, the more likely they are to own private means of transport such as motorcycles or cars (table [7], p. [18]).

Figure 6: Means of Transportation
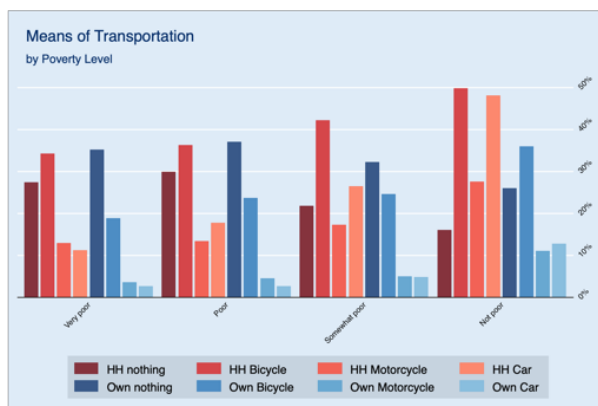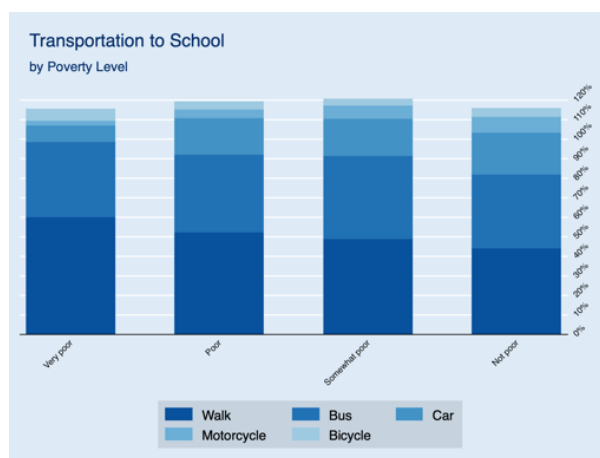

Figure 7: Transportation to School

This concludes our preliminary description of the sample population. We will now look more closely at what the students know and think about our central topics: environment and gender..

---

[21] The sum of the shares per column are slightly below 100% because we excluded negligible categories like 'solar' panels' or 'other kinds of energy'.

3. Theoretical Framework

3.1. Theory of Change

The RdE-project is based on a simple theory of change: There is little awareness of environmental and gender problems among young students in El Salvador. The mechanisms of the climate change, the impact of greenhouse gases, the *El Niño*-phenomenon, the deforestation are little known to the majority of the students (c.f. figures 11 and 12, p. 33). Similarly, the causes and consequences of gender inequality or the catastrophic effects of gender-based violence have failed to reach young people. The specific theory of change that underlies the RdE-project is thus as follows:

1. In a short intervention and with little financial and human resources folk teachers may be able to give the students a better understanding of the complex current societal topics, increase their EnvA and GenA and change their attitudes.

2. Folk teachers will teach alternative best practice behavior and show how to protect the environment and how to eliminate gender discrimination and violence.

3. The change of EnvA and GenA may translate into observable changes in behavior and encourage young people to become socially involved.

4. In the long run changes in environmental and gender behavior will lead to less environmental damage and less gender discrimination and violence.

5. At the macro level, the success of such a project could encourage the introduction of new pedagogical methods and important contents in – if one thinks big – educational policies and official curricula at departmental or even national level. In addition, the folk teacher program, if successful, could be applied by other NGOs and associations to improve educational outcomes and raise awareness for important societal issues.

3.2. Concepts

But let us come back to our more modest research question. The very aim of this study is to find out whether or not the *Consciente* workshops had a causal effect on the EnvA and GenA of the participating students. Therefore, we will now first clarify a few basic concepts and major distinctions to better understand what EnvA and GenA are.[22] What exactly should have been changed by the intervention?

3.2.1. Awareness and Concern

Awareness and Concern are complex social-psychological multi-dimensional, multi-hierarchical concepts. There are a multitude of theories that try to explain what environmental or gender awareness really are. Scholars usually decompose awareness and concern into three components or dimensions on the horizontal level. On a further vertical dimension, a hierarchy of fundamental value systems and basis orientations shaping these horizontal dimensions is assumed (Dunlap, 2017). Let us start with the horizontal dimensions and facets of awareness:

---

[22] For an overview of the current controversial discussion see: Albarracín et al. (2005).

(1) The *cognitive dimension* of awareness consists of the *knowledge* of facts or the *belief* in supposed facts. In terms of EnvA this means knowledge about environmental facts, problems and interdependencies. It means the rational insight that human individuals and societies nowadays endanger their natural environment and the planet as a whole and that human activities have serious impacts on present and future generations.

(2) The *affective-evaluative dimension* of awareness contains an affective-emotional and an evaluative component. For example, the awareness of environmental degradation causes either worries and fears or no emotional reaction at all and it triggers judgments. Similarly, the awareness of systematic exclusion of women, e.g. from the executive floors of companies, may be considered as either bad or natural, and it causes feelings of anger or indifference (Maloney & Ward, 1973; Maloney et al., 1975).

(3) The *conative dimension* of awareness means the individual intentions to act and the willingness to adapt individual behavior to knowledge and attitudes. We know that awareness does not automatically translate into adequate actions, particularly when costs are high (Diekmann & Preisendörfer, 2003). We all know that we often act against our attitudes. A clear awareness of the unpredictable dangers of climate change does not prevent people from frequent flying. That is why we must distinguish clearly between *intentions to act* and actual *observable behavior*.

Let us take a closer look at these dimensions of EnvA and GenA

3.2.2. Cognitive Dimension

We can easily understand that complete cognitive ignorance of ecological problems and gender inequality or violence will have no effects on a person's affects and attitudes, and even less on his intentions to act and his real behavior. On the other hand the greater the environmental expertise, the more likely affective-evaluative reactions and intentional and behavioral adaptions will occur. Knowledge and beliefs are the informational basis of one's attitudes and intentions.

This *cognitive component* not only consists in factual knowledge like that something like climate change exists, by how much temperature will raise until 2100 or that the latest World Conferences on Climate Change took place in France and Poland. Far more important is the rational understanding of the mechanisms of phenomenons like the "Greenhouse effect" or "El Niño", of what driving a car or traveling by plane means for CO_2 emissions, what eating meat has to do with the deforestation of the Amazon rainforests etc. On the other hand, the knowledge dimension of GenA means to be aware of the fact that there is a gender gap, i.e. social inequalities between female and male human beings, that goes far beyond biological differences between sexes. For example, women have fewer professional opportunities and are restricted from participating in the labour market, they earn less for the same work, are excluded from political decisions or exposed to male violence (The World Bank, 2011).

3.2.3. Affective-evaluative Dimension

As human beings we are constantly evaluating the world we live in: people, behavior, objects, products, events, ideas, political opinions etc. and last but not least ourselves. The evaluation engine in our brain is always running. Nowadays, the most obvious expression of attitudes are the likes and dislikes on social media platforms and the inflation of more and more differentiated emoticons that show how we evaluate things and what we feel about them. A definition in a nutshell: an *"attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor"* (emphasis in original, Eagly & Chaiken, 1993, 1). The theoretical answers to the question what attitudes really are oscillate between two opposite positions:

(1) Attitudes are a kind of psychic entities that are fairly constant over time. They are mental 'things' that individuals are aware of, i.e. they keep a set of stable evaluations in memory and can activate them when required. If so, researchers can simply ask their individuals of interest about their attitudes and rely on their responses. The only major problem the researcher has to confront is the problem of *dishonesty*: Individuals may not want to reveal their true opinion and give strategically dishonest and wrong answers. Perhaps they do not trust the researcher and fear harm of any kind, or they are ashamed, timid, indifferent, unwilling or unable to answer, or they want to please the researcher and answer what they consider *socially desirable*.

Therefore, the answers will often be biased.[23] These problems are hard to address. However, following this 'reification concept' they can be solved in principle. You just have to use the right tricks and ask the right questions in the right way. As a consequence, if you solve the problem of strategic responses and social desirability, you can methodically rely on self-reported attitudes collected e.g. through questionnaires. That is why "[t]hroughout the social and behavioral sciences, the dominant method of attitude measurement is the collection of explicit self-reports" (Wittenbrink & Schwarz, 2007, 1). But unfortunately human beings are not that simple.

(2) On the other end of the theoretical continuum attitudes are conceived as largely unstable, fluid, spontaneous, unconsciously generated and, above all, context-dependent ad-hoc judgements. They are no fixed and thus questionable psychological entities. Rather, they are fluid processes. The ad-hoc-judgements strongly depend on deeper psychic and even physiological processes and structures, which are largely beyond the conscious influence and decision of the individual.[24] Attitudes and judgements are highly dependent on the context, the communicative situations and the cognitive processes of interpretation, understanding and judgement formation. As a result, researchers will never be able to 'clean' self-reported attitudes from the noise of the situation in which they ask their questions, since there are no such things like context-independent attitudes. The noise of the situation is the breeding ground on which attitudes and judgements grow. For example, most people would deny that they are racists when asked in a scientific questionnaire – even if they mistrust and discriminate people of other races in real life. Similarly, if people are primed before being asked a question, they tend to give a different answer (Kahneman, 2011).

We have to distinguish attitudes from *affects* and *emotions*. "Regardless of the origin

---

[23] For an early review see: DeMaio (1984); for a more recent contribution: Krumpal (2013)).
[24] For a review of this critique see: Tourangeau et al. (2000)

of attitudes, the term *attitudes* is reserved for evaluative tendencies, which can both be inferred from and have an influence on belief, affect, and overt behavior. (...) Thus, affect, beliefs, and behaviors are seen as interacting with attitudes rather than as being their parts" (Albarracín et al., 2005, 5). It is easy to understand that affects and attitudes are not the same even though affects have a strong impact on our attitudes and overt actions. However, we often observe attitudinal-emotional ambivalence. You may be afraid of bees, but you know very well how useful they are for pollination. So you have a positive rational attitude towards their protection and at the same time you feel negative emotions if you just imagine them buzzing around you. You may be disgusted by Donald Trump's oddities and still vote for him because you think his policy is really great. Affects and emotions are strong fundamental experiences and they accompany and guide our cognition, attitudes and behavior. Affects arise spontaneously. Depending on arousal and valence the reactions on affects can be reflex-like – fear accelerates the heartbeat – or well-considered, but they obviously influence our attitudes.

Two other issues that should not go unmentioned are that (1) attitudes are no stand-alone mental entities or processes. Rather, they depend on each other and form a more or less heterogeneous and inherently contradictory structure of attitudes (Ostrom, 1989). Changing one means changing others. (2) Attitudes and attitude structures are not stable over time. Mostly, they change slowly, but sometimes in one fell swoop, depending on the intensity of new experiences (c.f. Ajzen & Fishbein (2005) on attitude changes after '9/11').

There are several methodological responses to these difficulties:

1. You may stick to the 'thing'-theory of attitudes and refine the method of measurement of explicit self-reports. That means you have to address the problem of unwillingness, inability and social desirability. For example, you can address the problem of distrust by simply assuring anonymity and confidentiality – that's the way we chose in our study, hence with little success. Another potential solution is the use of sophisticated methods like "Randomized Response Techniques RRT" (Warner, 1965) or "Unmatched Count Techniques UCR" (Droitcour et al. (1991)).[25] But this comes at the risk of complexity of the methods undermining the very goal they were developed for: Respondents do not understand how they work and distrust even more.

2. Another response is the indirect measurement of attitudes in order to minimize the respondents awareness of what is measured and thereby lowering the probability of biased responses (c.f. Ferguson & Bargh (2007)).

3. Another method is to simply bypass the possibility of strategic responding by using measurable physiological and neurological reactions (Ito & Cacioppo, 2007).

4. Recently, new possibilities for the measurement of attitudes have emerged. Analyzing big data generated by social media, telecom companies or online shops and profiling individuals and their attitudes and actions simply based on their click-behavior will lead to more objective results.

---

[25] For an experimental comparison of the two techniques see: Coutts & Jann (2011)

3.2.4. Conative Dimension

For practical reasons, researchers often rely on individuals' self-reports when attempting to measure behaviors. But self-reported behavior is biased due to socio-psychological mechanisms such as wishful thinking, self-perception and self-idealization, social desirability and peer group pressure. We constantly observe inconsistencies between attitudes and behaviors. Our declared intentions to act are often, if not most of the time, far away from our real actions. So attitudes and self-declared intentions are poor predictors of actual behavior (Ajzen & Fishbein, 2005, 173-221). In Europe every halfway informed person knows that the raise of the no-frills airlines have led to an enormous growth of air traffic and emissions of greenhouse gases. In spite of the well-known threads related to climate change, growth outlooks of aviation remain overwhelming.

I just want to point to one important theory that tries to explain these attitude-behavior inconsistencies. The "Low-cost Hypothesis" "predicts that the strength of effects of environmental concern on environmental behavior diminishes with increasing behavioral costs" (Diekmann & Preisendörfer, 2003, 441). In low-cost situations the cost difference between ecological and non-ecological action alternatives is small, while it is high in high-cost situations. The lower the cost differences, the higher the impact of ecological attitudes on behavior. "In the extreme of indifference between two alternatives, attitudes make the difference." (Diekmann & Preisendörfer, 2003, 467) We conscientiously recycle coffee cream lids because disposing of them correctly or incorrectly costs almost the same. On the other hand, we have great problems to refrain from flying to Mallorca because going by train or refraining entirely from going there is far too expensive, ecological awareness or not. "To avoid cognitive dissonance and to maintain positive self-esteem, individuals downgrade or eliminate environmental aspects in high-cost situations as a relevant decisions criterion" (Diekmann & Preisendörfer, 2003, 444). We can probably assume that similar mechanisms also play a role in the field of gender equality. A man may easily stand back in trivial situations. But when there is more at stake, he will hardly refrain from taking advantage of his favored position within society.

3.2.5. Basic Orientations and Value Systems

Our beliefs, affects, attitudes, intentions to act and behaviors are socially constructed and depend strongly on our life experiences and the social environment in which we grew up and live in. On the other hand, scholars discuss the influence of genes and physiology (e.g. Olson et al. (2001)) that seem to determine our basic orientations to a certain extend. We can easily imagine that life experiences are shaped by our body's physiology, its health, strength and resistance, its size and attractiveness etc. and that these experiences in turn determine our fundamental values and attitudes structures. Hence, it is plausible that beliefs, attitudes and behavior are strongly shaped by an underlying value system that we acquire through socialization ("socialization hypothesis") and that, to a certain degree, has a biological basis.

One of the most popular theories based on a hierarchical concept of fundamental value systems and overlying attitudes and intentions is Ronald Inglehart's "theory of value shift". It has its roots in Maslow's theory of hierarchically structured human needs and the human socialization approach formulated by Karl Mannheim. The key argument is the following: Generations that have been socialized under material scarcity –

like people in developing countries or those who had to experience World War II – tend to develop *materialistic values* ("scarcity hypothesis") or as Inglehart (1990, 68) puts it: "One places the greatest subjective value to those things that are in relatively short supply". The essential interest of the poor is thus to satisfy material needs. To possess things is happiness. On the other hand, people, who grew up under affluent conditions of economic and physical security – e.g. western post-war generations – experience a shift from materialist to *post-materialist values* such as freedom of speech, quality of life, political participation, personal fulfillment ("affluence hypothesis"). The mechanism underlying this divergent value formation is *socialization*: "To a large extent, one's basic values reflect the conditions of that prevailed during one's pre-adult years" (*socialization hypothesis)* (Inglehart, 1990, 68). In later papers he distinguished post-materialist environmental concern in wealthy societies as an *aesthetic* issue from the materialist concern caused by environmental threats in poor societies suggesting that environmental concern in poor country is a reflex to immediate local environmental threats (Inglehart, 1995).

What follows from this theory? In a nutshell we interfere that EnvA as well as GenA may be a "higher-order, quality-of-life value" (Dunlap & York, 2008, 532) common in affluent countries with a large share of people holding post-materialist values and that the value shift is the main cause of the burgeoning of environmental and gender awareness. Put it differently, post-materialist values may be a pre-requisite of EnvA as well as GenA. People living in developing countries are mostly struggling to meet their basic needs and thus are driven by materialist values. Therefore, the majority of the population can be hypothesized to lack environmental concern and to care little about gender equality.

But there are critics as well: Dunlap & York (2008) showed that this assumption may not be warranted. Analyzing the results of the "Health of the planet HPO" survey from 1992 they found that: "[o]f the 14 different measures of environmental concern employed, 11 were found to be significantly correlated with national affluence – but surprisingly, seven were *negatively* related, indicating higher levels of concern among residents of poor nations! (...) Overall, the HOP results suggested that citizens' concern for environmental quality has become a worldwide phenomenon, *not* limited to wealthy nations." (Dunlap & York, 2008, 534). Nevertheless, Inglehart's theory is the theoretical heart of the large-scale World Values Survey.[26] Another criticism: Inglehart's theory fits perfectly well into the mainstream of modernization theories that postulates a certain kind of historical determinism in development trajectories from poor, underdeveloped, traditional and materialist societies to wealthy, developed, modern and post-materialist societies (Kößler, 2016).

To sum up our theoretical review: Attitudes are multi-dimensional and hierarchically structured more or less stable evaluations and judgements of all kinds of entities. They depend on socialization and, to a certain degree, on biology. They are highly influenced by situations and our emotional reactions and therefore not fixed and unchangeable. They are structured by deeply rooted and more stable value systems. Attitudes are no mental things but rather mental processes. While they do influence our behavior, we normally have to live with attitude-behavior inconsistencies.

What can we conclude for our investigation? Due to the low intensity of the inter-

---

vention and the low affective reactions it will probably cause, we can not expect sudden changes in attitudes, intentions or even behavior, and even less fundamental changes in value systems. So, if there are changes they will probably be small.

4. Research Design, Data and Methods

With these theoretical preliminary considerations we can now start with the presentation of our investigation. In this section we discuss the research design, the sampling, the instrument, the dataset, the instrument, the scores and indices we creates, the total survey error framework, the dataset and the the methods that are used for the analysis.

4.1. Research Design

4.1.1. Randomized Controlled Trial RCT

To resolve the *counterfactual problem* – i.e. the fact that we can never observe a unit simultaneously *under treatment* and *not under treatment* and, as a consequence, we can not calculate the *causal effect* of a treatment – we chose a "Randomized Controlled Trial Design" (RCT) which is "the gold standard of impact evaluation" (Gertler et al., 2016, 64). The core idea of randomization is to eliminate any kind of selection bias and to ensure that the treatment and control group do not differ systematically. So, schools, classes and students were assigned randomly to the two experimental groups with each unit having the same chance to be in one of the treatment groups (c.f. chapter 4.2, p. 26).
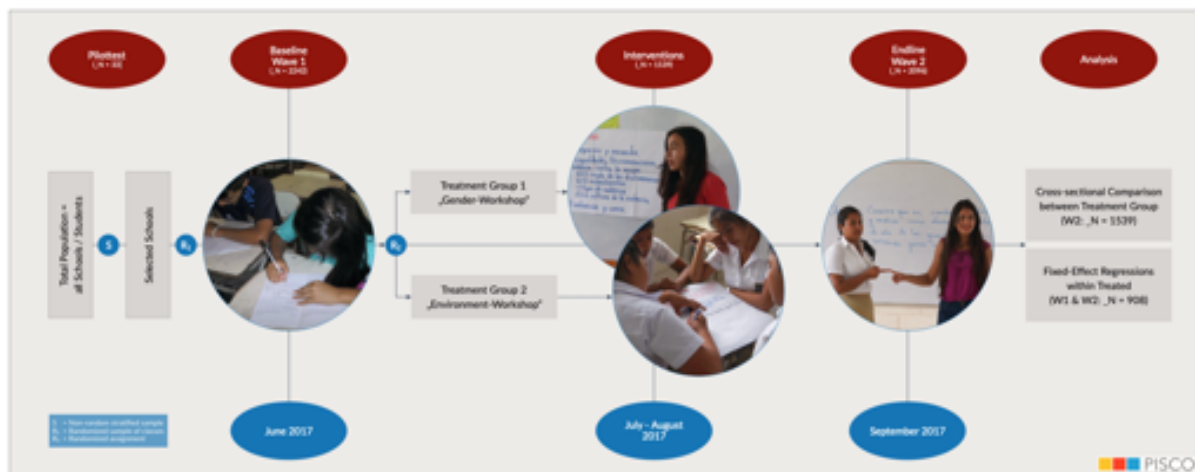
Figure 8: Design and timetable of PISCO and the RdE-program

4.1.2. Timetable

In June 2017 we realized our baseline survey (wave 1 = W1) with 2242 students. From mid-July up to end of August 2017 *Consciente* and its folk teacher team executed the workshops in 98 classes in 13 schools all over Morazán with officially 2848 students. The endline survey (wave 2 = W2) with 2096 respondents took place in September 2017 about four weeks after the treatment. *Consciente* tried to keep an equal time interval

between the treatments and W2, which in most cases worked quite well. So we could assume a homogenous weakening of the learning outcomes and attitudinal changes and did not have to control for different effect mitigation due to time-delayed surveys.

4.1.3. Explorative Interviews

Before we started to develop the design and the instrument we carried out a few structured oral and written interviews to get a look-and-feel impression of the Salvadoran everyday life. The interviewees were the managers of the folk teachers program and two female folk teachers.[27] Unfortunately, no explorative interviews with young people and ordinary teachers from different regions could be conducted. This turned out to be a mistake. We developed the questionnaire on the basis of the knowledge, enthusiasm and high expectations of the program managers, mixed up with our own high expectations and western misjudgments regarding the situation and survey-taking ability of the young people in Morazán. Therefore, the questionnaire overshot the target by far.

4.1.4. Pilot Test

In early June 2017, we realized a small pilot test in an urban school of Gotera with 33 participants. This pilot test was done in a hurry and far too late since W1 started one week later. The main reason for the delay was the belated definition of the actual workshop program, to which the questionnaire, of course, should have been largely adapted. The most important finding from the pilot test was that the questionnaire was far too detailed and the students were largely overwhelmed by it. There were only a few days left to shorten the questionnaire by some extent.

This experience shows that the conduction of an extensive pilot test is indispensable because it can significantly improve the quality of the data. A failure to do so can do serious damage to any scientific study, but is particularly problematic when an investigation is conducted in a setting widely unfamiliar to the researcher.

4.2. Sampling

Let us have a closer look at our data now. For financial and administrative reasons it was clear from the outset that only about half of all students in Morazán would attend a workshop. Hence, it was to decide how to select students, classes and schools to live up to scientific standards. Therefore, a complex, multi-stage sampling procedure had to be developed.

The *total population* of eligible units embraced all students from all *public* high schools in Morazán.[28] The total population was distributed among 29 schools and a total of 212 classes. Our *basic population* comprised 5323 officially registered students according to class lists provided by the "Ministry of Education" MINED in San Francisco Gotera. These lists served as *sampling framework*.[29]

---

[27]   The transcripts can be found in the *06_HumanSubjects*-folder of the documentation.
[28]   Three private schools in the department with 92 students were excluded from the beginning because they were outside the direct control of the Ministry of Education the project was coordinated with.
[29]   We use the well defined terminology of Groves (2009).

4.2.1. Clusters

In our study setting, class-level randomization had several advantages over individual level randomization: (1) We did not have to conduct both *Consciente* workshops in all schools, which was organizationally and financially impossible anyway; (2) we did not have to tear apart classes so that some students of the same class would have been assigned to a treatment, others not; (3) we avoided uncontrollable spillover effects and (4) a lot of envy and trouble among students and teachers. As a consequence, we randomly sampled *schools* and, in a second step, we randomly draw classes until we exceeded the 50% of all students of each selected school. A lightly different procedure was chosen for the three large schools.

4.2.2. Stratification

It can be assumed that the environmental degradation (pollution, waste, traffic, industry, etc.) and thus also EnvA differs between cities on the one hand and villages or remote hamlets on the other hand. The surroundings of the capital San Francisco Gotera are deforested to a large extent, in the most remote villages the huts are under green banana trees at the border of the last bits of jungle. Similarly, it can be assumed that gender and sexual issues are differently present in urban than in rural areas, e.g. because access to the mass media and the Internet is much easier in the center than in remote areas.[30] That is why it would have made sense to stratify our population with regard to geographical characteristics. But geography alone is not sufficient. Poor people may think different than rich people (value shift), students in small schools may be different from students in large schools.[31] Marginalized people may differ from not-marginalized people. So we looked for an appropriate criterium for stratification. We've come across the Salvadoran government's official index of marginality.

4.2.3. Marginality Index

In 2005 the governmental "Fondo de Inversión Social para el Desarollo Local de El Salvador FISDL" published an investigation on poverty in El Salvador. In the "Mapa de Pobreza" FISDL created marginalization indices and maps of all the 262 municipalities (*municipios*) of the country (FISDL, 2005, 23).[32] The most important finding of the investigation: Morazán turned out to be one of the poorest departments of El Salvador.[33] The FISDL marginality index is based on data of the national household surveys

---

[30] Moreover, it can be argued that the political attitude in rural areas is more conservative-right and in urban areas liberal-left. However, the political map of the last elections does not confirm this. https://www.google.com/maps/d/viewer?mid=1csfPEvAb22q6JexTN5qRaiwtJJ1rZyDH&ll=13.767464551453227%2C-88.11419902110282&z=11; last date accessed: 26.11.2018. This may be due to the fact that the left-wing guerrilla had one of its strongholds in the mountains of Morazán during the civil war.

[31] The biggest school in San Francisco Gotera, the "Instituto 14 de julio de 1975", had 1243 students, the smallest "Centro Escolar de San Fernando" in the remote village San Fernando only 15.

[32] http://www.fisdl.gob.sv/temas-543/mapa-de-pobreza; last date accessed: 07.01.2019.

[33] "Como resultados tenemos que hay 32 municipios en pobreza extrema severa, 68 municipios en pobreza extrema alta, 82 municipios en pobreza extrema moderada y 80 en pobreza extrema baja" (FISDL, 2005, 6).

"Encuesta de Hogares de Propósitos Múltiples EHPM" of 2001 through 2004 using six variables: (1) health, (2) education, (3) household size and occupancy, (4) water connection, (5) sanitary facilities, (6) electricity connection. Though more than a decade old, this is the most recent data available at the municipality level. We used the FISDL data as *marginality index* (fig. 9 and 10, p. 28).[34]
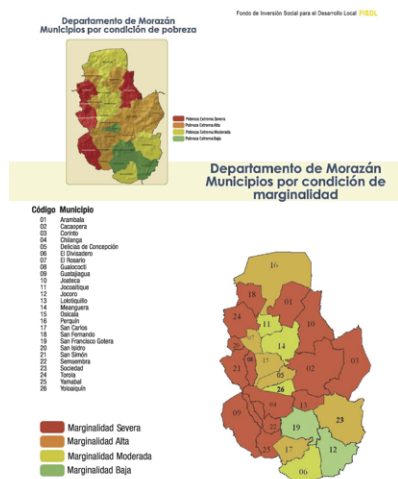
Figure 9: FISDL Marginality Index
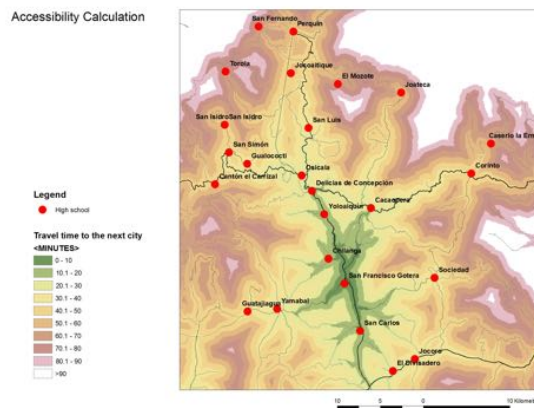
Figure 10: Accessibility Index

We sorted the remaining schools by marginality and divided them into three groups of roughly equal in size. Schools with a marginality index below 30 were placed in stratum 6 (*rural*) schools below 38.5 points in stratum 5 (*semi-rural*), and schools below 44.5 in stratum 4 (*semi-urban*). The large schools, all with marginality indices above 44.5 points, were each treated as a separate stratum (*urban, rural center 1, rural center 2*). We sampled half of the classes from them.

Once the 13 schools and 98 classes had been selected for study participation, we conducted a random assignment of classes to either of the two workshops.

4.2.4. Sample

Tables 12, p. 29 shows the results of our sampling procedure: In total, we selected a *gross sample* of 2848 students from our *sampling framework* of 5323 students. 2242 or 78.7% (*completion rate*) of the gross sample or 42% of the basic population participated in W1. 2096 or 73.6% of the gross sample and 39% of the basic population participated in W2. Table 13, p. 30, shows the distribution of strata, schools, classes and respondents per wave.

---

[34]   We tried out an alternative measure. We calculated the travel time from each school to downtown San Francisco Gotera. The idea was to combine the socio-economic with geodetic data. Our *accessibility index* was based on the coordinates of the schools and the distance and road quality to San Francisco Gotera. However, the two indices did not coincide to a large extent. So we did not use the *accessibility index*. Instead, we relied entirely on the FISDL marginality index.

Table 12: Sample

	Basic Pop.	Gross Sample	Perc.	Cum.
Stratum 1: Urban	1289	645	50.15	22.64
Stratum 2: Rural Center 1	575	297	51.65	10.43
Stratum 3: Rural Center 2	688	345	50.14	12.11
Stratum 4: Semi-urban	1077	538	49.95	18.90
Stratum 5: Semi-rural	1081	539	49.86	18.93
Stratum 6: Rural	903	484	53.59	16.99
Total	5323	2848	53.50	100.00

	Respondents	% of Basic Pop.	% of Gross Sample	
Respondents W1	2242	42.12	78.72	
Respondents W2	2096	39.38	73.60	
Compliers	1539	29.91	54.04	
Defiers	557	10.46	19.55	
Identified Compliers W1/W2	908	17.06	31.88	

4.3. The Instrument

As discussed, EnvA and GenA are multi-dimensional and hierarchically structured concepts. The concept of "attitudes" is latent, meaning that we can not measure it directly. We used a questionnaire as measurement device. Other self-report methods such as face-to-face interviews might have minimized problems related to discipline, concentration or functional illiteracy, but would have been far too expensive. Implicit techniques such as unobtrusive observations, priming or response competition measures have been excluded for financial reasons, too (Milfont & Duckitt, 2010). Since most people and schools in El Salvador do not have access to computers, a comprehensive online survey was ruled out from the outset. A media mix of online surveys in the San Francisco Gotera and paper surveys in the periphery has not been considered either for statistical reasons.

We have, of course, used well-established item batteries to develop our survey questions. But, we soon found out that many questions were hardly applicable to young high school students in a poor Latin American country. The "World Value Survey" is one of the most prominent international surveys on values and attitudes. However, it only contains few items on environmental concern and they are too general or to "western" to be apt for our purpose. Let us e.g. have a brief look at questions on environmental issues: What can a Salvadoran high school student answer to the question: "Are you member in an environmental organization?" (V30) or "Do you trust in environmental organizations?" (V122) when there no such organizations around? Or what sense does an the item make like "Have you participated in a demonstration for some environmental cause?" when we address teenagers of a poor district where such demonstrations rarely – if ever – occur? Some questions aim at the "willingness to pay WTP" for eco-taxes or more expensive

Table 13: Schools, Students and Classes

Stratum	School	_N W1	_N W2	Classes
1. Urban	1. INGO	558	529	20
2. Rural Center 1	2. INSEM	242	229	12
3. Rural Center 2	3. INO	289	242	11
4. Semi urban	4. C.E. Naciones Unidas	98	105	5
	5. Inst. Prof. Jose L. Lopez	235	181	10
5. Semi rural	6. C.E. Juarez Aleman	47	47	3
	7. Inst. de Yamabal	110	114	5
	8. Inst. de Corinto	255	247	10
6. Rural	9. C.E. Caserio Mozote	34	30	2
	10. C.E. Gen. M. J. Arce	72	53	4
	11. Inst. Nac. San Isidro	91	93	5
	12. C.E. Canton el Carrizal	80	73	4
	13. Inst. de Guatajiagua	131	153	7
Total W1 and W2		2242	2096	98

eco-goods. But such questions can hardly be asked to young people who do not have a coin in their pockets. What can a teenager say about his preferences of environmental protection versus economic growth (V81)? The same problem we encountered with other standard surveys like *GESIS*[35] or the *Environmental Attitude Inventory* by Milfont & Duckitt (2010). Most of the questions were simply not applicable. Hence, we decided to develop our own question catalogue which, in the end, turned out to be still far too Western.

4.3.1. The Questionnaire

The questions of our survey were grouped in 12 chapters. In the following overview we report the groups and the most important items.[36]

(01) Paradata (*PD_*variables): *date of the survey*, *school* (2),[37] *grade* and *class* (3). Other PD-variables were filled in by the data collectors during data entry and for consistency checks: *collectors' code*, *collectors' name*, the *workshop topic* hold in the corresponding class, *folk teachers' name* and *date* of the workshops;

(02) Socio-economic characteristics (*SE_*variables): *sex* (4), *birthday* (5), *confession* (10), *place of residence* (16);

(03) Socio-economic household characteristics: *number of persons* (14), *equipment* (26 - 29), *degree of poverty* (25);

---

[35] GESIS: https://www.gesis.org/home/
[36] The complete questionnaires can be found in the *07_ Questionnaires*-folder of the documentation.
[37] Question numbers in brackets.

(04) Socio-economic characteristics of parents such as *education*, *literacy level*, *profession* (17 - 24);

(05) Marginality: *time to school* (30); *time to the center* (31);

(06) *Env*ironmental *K*nowledge questions (*EnvK*_variables: 33 - 42);

(07) *Env*ironmental *A*ttitudes questions (*EnvA*_variables: 43 - 49);

(08) *Env*ironmental *B*ehavior questions (*EnvB*_variables: 50 - 58);

(09) *Gen*der *K*nowledge questions (*GenK*_variables: 59 - 66);

(10) *Gen*der *A*ttitudes questions (*GenA*_variables: 67 - 76);

(11) *Env*ironmental *B*ehavior questions (*GenB*_variables: 77 - 79 (only for girls); 80 - 82 (only for boys));

(12) Feedback (*FB*_variables: 84 - 88);

The question groups 06 to 11 were identical in W1 and W2, the SE_variables were dropped in W2 with the exception of "sex" and "siblings". The FB_variables were supplemented by final feedback questions in W2. At the request of the Ministers of Education, we further included a block of evaluation questions about the students' school and their teachers in W2, but this is not relevant for this study.

4.3.2. Scales

For questions about attitudes, a 6+(1)-level Likert scale was used in the pilot test, i.e. six levels for the assessment of the question and one level for the option "No opinion / No tengo opinión". The +1-option was dropped after the pilot because more than 30 percent of the respondents consistently checked just this option. In retrospect, we probably would retain the +1-option. While respondents can be forced into giving an answer, this does not guarantee an honest answer. What is worse, this procedure makes it impossible for the researcher to differentiate between content-related answers and no-opinion answers, as the latter are hidden in fake content responses and there is no way to isolate them. Moreover, we lost the possibility to measure eventually lower or higher "No opinion"-frequencies after the treatment due to changes in motivation and consciousness. We would also have reduced the frequency of response patterns, too. But we decided differently.

A middle position was also omitted. A prior survey in Morazán by Martina Jakob (Jakob, 2016) showed clearly that the +1-option is disproportionately often used, possibly due to embarrassment or convenience. Above all, young people of Morazán may have never formed an opinion on many questions of our survey. We thought that leaving out a center would force the students to think and show at least a tendency.

4.4. Scores and Indices

For the analysis, we had to reduce the large number of items considerably. For this purpose, we created *knowledge scores* from the Env*K*- and Gen*K*-vars, *attitudes* and *behavior indices* from the Env*A*-, Env*B*-, Gen*A*-, Gen*B*-variables. Our procedure to detect structures in the data and to reduce the number of variables was as follows:

1. First, we summarized items exclusively on the basis of *content considerations*, i.e. we compressed groups of items based on one same underlying concept such as "Belief in classical gender roles", "Skepticism towards Climate Change", "Bad or Good Environmental Behavior'" or "Acceptance of gender violence".

2. In a second step we computed *covariance* matrices and pairwise *correlations coefficients* of related variables.

3. Third, we reviewed the composition of our provisional indices using *confirmatory factor analysis*. If the factor analysis showed deferring results about supposed relations between items we either dropped an item from the index or we gave preference to our content considerations. The latter was done for the following reason: Even though factor analysis can detect unexpected data structures and support content-related considerations, it does not "think" and "understand", but only calculates. Hence, we can not entirely rely on its results.

4. In a last step the indices were re-scaled to values between 0 and 1, so that coefficients can be read as percentage points.

Following this procedure, we created a set of scores and indices as shown in tables 14 and 15 (p. 34, 35).

4.4.1. Confirmatory Factor Analysis

We used the principal factor analysis method to confirm our content considerations. It's the more suitable method for our purposes because it assumes correlated factors and the cummularity is not just assumed to be 1. Hence, it does not decompose the total of variance, but only the common part of the variance. If the cummularity is below 1, the variance is not completely explained. Uniqueness (= 1 - cummularity) should be as low as possible. We used the Kaiser-Meyer-Olkin measure of sampling adequacy MSA to inform our decision wether we could use an item or not. Only items with a minimum *kmu*-value of above 0.6 (mediocre) were considered. Most included items had values above 0.7 (middling) or 0.8 (meritorious). For every block of items we plotted scree-, score- and loading-plots for visualization and further analysis. Factors with Eigenvalues below 1 were usually dropped. You find one set of plots of our index building procedure in the annex, p. 71.[38]

4.4.2. Scores

The environmental knowledge score was built using the EnvK-variables where students had to answer the following questions: "Have you ever heard of climate change?" (EnvK_Greenhouse_SC, 1 point), "Do you know which gas has the greatest influence on climate change?" (EnvK_Gas_SC, 1 point), "What do you think? Are the following natural resources renewable?" (EnvK_CC_SC, 1.5 point), "Do you think there is a relationship between climate change and the following actions?" (EnvK_CC_SC, 1.5 point), "Do you know what the '3 Rs' mean?" (EnvK_CC_R_SC, 1 point), "Do you think the

---

[38] All the other tables and graphs you can reproduce by running the STATA-do-file PISCO_1_04.do in the 09_Analysis-archive.
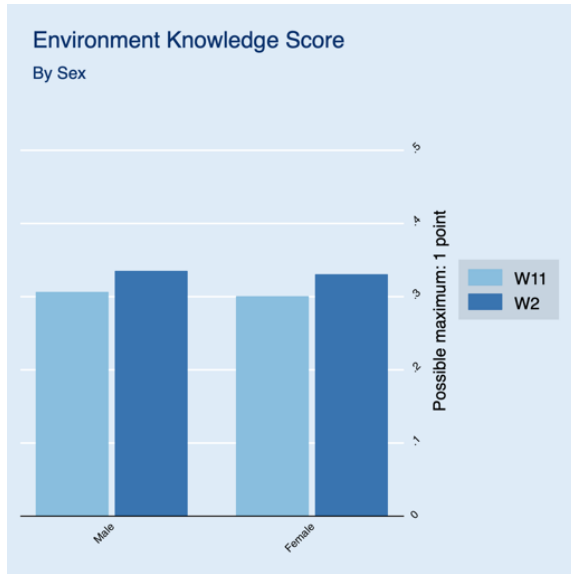
Figure 11: Environment Scores
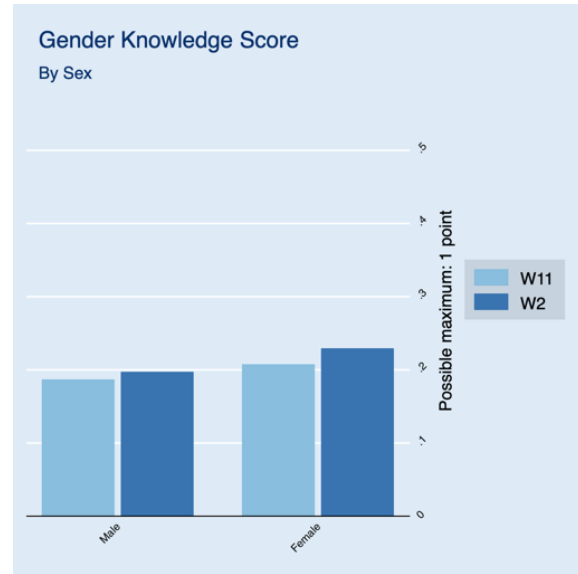
Figure 12: Environment Scores

forest surface in El Salvador has changed in recent years?" (EnvK_Selva_SC, 1 point). The gender knowledge score was formed analogically on the basis of GenK-variables.

For correct answers we awarded 1 point each, for half correct 0.5, for approximately correct 0.25 points and for wrong or missing answers 0 points. The highest possible score in environment knowledge was 8 points and 7 points in gender knowledge. We standardized the scores to values between 0 and 1 to make them comparable.

Open questions were not evaluated manually, but automatically using program code to keep the results as objective as possible. The frighteningly frequent spelling mistakes were corrected as far as possible before running the computer evaluation. Certain keywords had to appear in an answer to be considered correct. For example: In the question for the most influential greenhouse gas the words "CO_2", "carbon dioxide" or "accumulation of greenhouse gases" had to be named. If "ozone layer", "acid rain" or something similar were named we gave zero points. Another example: The keywords "woman" and "violence" had to appear in the question of the law for women protection. The correct name of the pretty famous and modern law is "Ley Especial Integral para una Vida Libre de Violencia para las Mujeres (Special Comprehensive Law for a Violence-Free Life for Women)".[39]

The text analysis program code could still be refined, but for our purposes the results were satisfactory.[40]

The following figures 11 and 12, p. 33, show how little the students really know and how small the changes between W1 and W2 were. Note that the highest possible score was 1, the y-scale shown only reaches up to 0.5 points. For example: 31.71% said they never heard about the climate change. Only 16.47% said they could explain it in a few words and only 0.94% gave a clearly correct answer. 7.22% knew that CO_2 was the most important greenhouse gas. 26.60%, 37.82& and 47% respectively said that petroleum, coal and gas were renewable sources of energy. Table 14 (p. 34) shows the environmental and gender knowledge scores in W1 and W2. We can observe a modest rise of both

---

[39]   See: http://evaw-global-database.unwomen.org/; last date accessed: 12.01.2019
[40]   See: STATA-do-file *PISCO_1_05.do* in the 09_Analysis-folder.

scores but nobody reached the maximum score of 1. We will test if these improvements are significant later in the result chapter 5, p. 44.

Table 14: Scores Wave 1 and 2

	Obs	Mean	Std.Dev.	Min.	Max.
SC_EnvK_W1	2242	0.3025	0.1118	0.00	0.78
SC_EnvK_W2	2096	0.3311	0.1182	0.03	0.81
SC_GenK_W1	2242	0.1959	0.1030	0.00	0.54
SC_GenK_W2	2096	0.2126	0.1116	0.00	0.54

4.4.3. Indices

We formed the following indices on attitudes and behavior:

- Environment:

  - Environmental problems exposure index (IND_Env_exposure)
  - General environmental worry index (IND_EnvA_gen_Worry)
  - Climate change concern index (IND_EnvA_CC_Concern)
  - Skepticism towards environmental problems index (IND_EnvA_Scepticism)
  - Environmental problems trivialization index (IND_EnvA_Trivialisation)
  - Indifference towards climate change index (IND_EnvA_CC_Indifference)
  - Green consumerism index (IND_EnvB_Behavior1)
  - Waste management index (IND_EnvB_Behavior2) and
  - Eco-commitment index (IND_EnvB_Commitment).

- Gender:

  - Female gender problems exposure index (IND_Gen_exposure_w)
  - Machismo index (IND_GenA_machismo)
  - Attitude towards equality index (IND_GenA_equality)
  - Attitude towards classic gender roles index (IND_GenA_classic)
  - Gender violence trivialization index (IND_GenA_violence)
  - Gender violence naturalization index (IND_GenA_violence2)
  - Misogynistic male behavior index (IND_GenB_behavior_m)
  - Female emancipation index (IND_GenB_behavior_w)
  - Female subordination index (IND_GenB_behavior2_w).

We standardized the indices to values between 0 and 1 to make them comparable. These scores and indices will be the basis of our analysis of treatment effects.[41]

---

[41] Table *All_Items.pdf* in *07_Questionnaires*-folder of the documentation shows the composition and the underlying concepts of all indices in detail.

Table 15: Indices W1

	Obs	Mean	Std.Dev.	Min.	Max.
Environmental Problems Exposure	2100	0.8032	0.2052	0	1
General Environmental Worries	1814	0.8368	0.2193	0	1
Climate Change Concern	1787	0.8148	0.2023	0	1
Ecological Scepticism	1286	0.4413	0.2547	0	1
Ecological Trivialisation	1471	0.5351	0.2610	0	1
Ecological Indifference	1709	0.3168	0.2882	0	1
Green Consumerism	2092	0.4777	0.2405	0	1
Waste Management	1911	0.7748	0.1563	0	1
Eco-committment	2040	0.7380	0.2656	0	1
Female Gender Problems Exposure	1051	0.4614	0.3640	0	1
Machismo	1953	0.2970	0.2446	0	1
Gender Equality	1886	0.7284	0.1858	0	1
Classical Gender Roles	1996	0.4954	0.2539	0	1
Gender Violence Trivialization	2007	0.2588	0.2586	0	1
Gender Violence Naturalization	2029	0.2844	0.2301	0	1
Misogynistic Male Behavior	856	0.8321	0.2103	0	1
Female Emancipation Index	1088	0.5681	0.2367	0	1
Female Subordination Index	1010	0.9128	0.1637	0	1

4.5. Total Survey Error Framework

Groves (2009, 40) condenses the error problem of surveys into two characteristics that a survey must fulfill: "1) Answers people give must accurately describe characteristics of the respondents. 2) The subset of persons participating in the survey must have characteristics similar to those of a larger population." These two general requirements give rise to two potential sources of errors: "Measurement errors'" and errors of "nonobservation". With regard to the measurement the author discusses *validity*, *measurement errors* and *processing errors*; with regard to the representation, he differentiates between *coverage errors*, *sampling errors*, *nonresponse errors*, and *adjustment errors*, i.e. post-survey adjustments: weighting, imputation (Groves, 2009, 48). The main threat to the validity of an RTC experiment are *non-response*, *non-attendance*, *non-compliance* and *attrition* (Glennerster & Takavarasha, 2013).

So let us examine the most important types of error that may have affected our survey.

4.5.1. Representativity

Thanks to our well-defined target population we did not systematically exclude parts of the basic population (apart from the negligible number of private school students). But our sampling frame, i.e. the students lists of the MINED, was far from being reliable.

We have no idea of the magnitude of our *under-coverage rate* and how many *ineligible units* we have in our frame population, i.e. students who simply do not exist. So we can not pinpoint our *coverage errors* exactly. As we have already seen in table 12, p. 29, we observed a loss of 606 of students (-21.28%) in W1 and of 752 students (-26.40%) in W2. If these losses are random, we will not have a selection problem. If they are systematic, we will loose representativity of our sample. What could account for the absence of these students during the surveys? Different explanations can be put forward:

(1) One part of the problem may be caused by administrative failure. Possibly the registers of the MINED were simply not up-to-date. You would expect that the provision of a complete and up-to-date school, class and students lists would not be a major problem for the Ministry of Education. Our western expectation is that all data is centrally recorded on computers. But far from it: Here we are dealing with a problem of state institutions that are still under construction and have limited funds at their disposal. Therefore, there is no central school administration software, neither in the government nor in the schools out there. They have only laboriously compiled and mostly outdated paper lists to manage their schools. So data is not recorded systematically, and its updating lags far behind the constant changes in the school. Even though *Consciente* tried to verify the figures by countless phone calls with the school principals, the information could not be updated satisfactorily. Many school principals did not want to answer because they distrusted the unknown and perhaps unauthorized callers, even though the MINED had given the authorization and had informed their principals in advance. But orders do not spread as easily as we would expect. Some school principals even seemed to boycott the strange project of the MINED, be it for party political reasons, be it that they simply wanted no one to interfere in their everyday school life. Administrative errors do not compromise the representativity of our sample, because we can reasonably expect list errors to be random rather than systematic.[42]

(2) On the other hand, the most obvious reason for the difference does indeed affect sample representativity: students may simply skip school.[43] We can assume that truancy rate is high, but we can not assume, that the truants are exactly like the attending students. Perhaps the absence rate has something to do with poverty and the distance to school – or it is simply lack of interest. As a consequence, the descriptive statistics of our sample population may be biased due to the imperfect representativeness of the sample. Possibly, this also affects the estimates of our treatment effects. If the truants are the weaker students, i.e. they might have learnt less in the workshops and our treatment effects would be biased upwards. Conversely, if the truants are the better students who believe they can afford truancy the treatment effects might be biased downwards.

In summary: Even though our randomized sampling procedure eliminated *systematic sampling errors* our effects have to interpreted with caution if we want to make statements about the basic population.

What about the severe problems of unit-non-response and attrition, of non-compliance

---

[42] Another, admittedly nasty explanation might be that school principals could systematically indicate higher student numbers in order to receive more financial governmental contributions for their schools (and themselves). But these suspicions are far from being confirmed.

[43] End of August 2017 I attended some of the workshops to get a personal impression as to how they work. I was amazed at the coming and going of the students. When I asked some of the late students where they had been, they replied: "At the kiosk."

and item-non-response? Let us look at this more closely to have get an idea of the *total survey error* of our study (Groves, 2009).

4.5.2. Unit-non-response and Attrition

As discussed, we can not distinguish between unit-non-response and non-sample-related failures. We have no adjusted net sample size. Our unit-non-response-rate is at the most 21.28 or 26.40% in W1 and W2 respectively, depending on wether we take the official lists to be accurate. For unknown reasons, we lost 146 units between W1 and W2. Our attrition rate is thus –6.5%. Maybe we're just dealing with another group of truants or some students did not show up in W2 because they knew they had to fill out another boring survey.

4.5.3. Non-compliance

The main goal of the randomized assignment to treatment is to eliminate sampling variability and selection bias, i.e. to make sure that control and treatment groups are not systematically different. Table 16, p. 38, shows that 1539 students or 73.43% of all respondents in W2 said they participated in a workshop, 557 said they did not or they did not answer the question, so they can not be taken as treated. As a consequence our *non-compliance rate* is 26.57%. We can not simply assume, that the drop-out has been random. Rather, it is very likely that the groups of compliers and defiers are systematically different in important characteristics. For instance, the rate of male compliers was far below that of female compliers (71.34/76.10) although the ratio of male to female students did not vary a lot between the waves (W1: 46.18/53.82; W2: 45.74/54.26). The boys seem to be more reluctant to participate and they may have a higher tendency to skip school when they are not interested.

For our analysis that means that the groups of compliers and defiers are systematically different. But due to our random design this will not affect our estimators of treatment effects because we can assume that the defiers are homogenous in both assignment groups.

4.5.4. Individual Identification W1 and W2

In order to maintain the credibility of anonymity and thus to ensure openness and honesty of answers, we did not use any obvious tools for identification like ID numbers, names, addresses etc. in the questionnaire. Anonymity was important to make sure that students would answer unusual, personal and even intimate questions honestly. We also did not control class lists and participation during the surveys and the workshops. The effort would have been very laborious, expensive and prone to error. Instead, we erroneously assumed that the identification of individuals in W1 and W2 could easily be ensured by comparing *class*, *sex*, *birthday* and the number of *siblings*. But that was a momentous mistake.

We have been able to identify just 1,126 (53.72%) individuals across the two waves, 908 of whom participated in a workshop, whereas 218 said they did not. In total, 970 (46.28%) students could not be identified, 631 of whom participated in a workshop and 339 of whom did not. Table 17, p. 39, shows the participation rates in the different workshops. 55.45% of the survey students participated in the environmental workshop,

44.57% in the gender workshop. We have to assume that the group of identified individuals differs considerably from the group of non-identifiable individuals. They may be more conscientious, careful, attentive, concentrated, inquisitive, servile etc. Hence, all our findings based on this group suffer from limited generalizability.

In 46.28% of the observations identification did not work for various reasons:

1. Data entry errors: We have to assume a not negligible number of data entry errors because our control and replication possibilities were limited;[44]

2. Incorrect answers: Many students might have distrusted the guaranteed anonymity. Perhaps they feared identification possibilities and gave false answers on crucial items such as their date of birth intentionally. Quite frequently our surveyors reported problems of mistrust and refusal despite all assurances of anonymity and the exclusive scientific use of the data in faraway Switzerland. This phenomenon may be interpreted as a lack of social capital like trust in Salvadorian society;

3. Another reason for making identification impossible may be the unusual indiscretion and impertinence of some questions, above all on sex and religion that many complained about in the feedback questions. So probably some have given wrong birthday dates to exclude identification;

4. Some of the false answers may simply result from the silliness, indifference and lack of concentration of pubescent teenagers.

5. Sometimes they might have failed to recall their exact birthday or counted the number of older siblings in different ways, e.g. once with and once without half–brothers and -sisters.

Table 16: Treatment Compliance

ID / Workshops	No	Yes	Total
Not identified	339	631	970
	60.86	41.00	46.28
	34.95	65.05	100.00
Identified	218	908	1126
	39.14	59.00	53.72
	19.36	80.64	100.00
Total	557	1539	2096
	100.00	100.00	100.00
	26.57	73.43	100.00

---

[44] e.g. 177 observations were entered with the same date of birth (01.01.1970), an irreversible recording error noticed too late. We had a couple of data entry clerks who were not sufficiently careful.

Table 17: Compliers

ID / Workshops	Environment WS	Gender WS	Total
Not identified	348	283	631
	55.15	44.85	100.00
Identified	505	403	908
	55.62	44.38	100.00
Total	853	686	1539
	55.43	44.57	100.00

4.6. Data Quality

We had to pay special attention to cleaning up the data. To prevent data collection errors at source, we urged data collectors to control questionnaires before collection. While, in theory, they should have reviewed students' surveys for missing or inconsistent answers, this was not regularly done in practice for several reasons: (1) Data collectors were often too unexperienced and too young to have sufficient authority in a setting of chaos and lack of discipline that is widespread in classrooms. They were not sufficiently prepared and many of them were not aware of the importance of their task. (2) On the other hand, teachers and even school principals often did not support the data collection process. Some were suspicious, others felt controlled and monitored, and unfortunately a few deliberately boycotted a program that came from the wrong government. (3) Students often simply refused for distrust and for simple adolescent silliness and renunciation. But unfortunately the lack of concentration and even widespread functional illiteracy, i.e. their inability to understand more complex texts, questions and tasks, might have had an influence, too. (4) Finally, the financial means limited the possibilities to improve data collection – e.g. by relying on more experienced data collectors or by cross-checking data recording and engaging more people. For this reason, we encountered considerable shortcomings in data collection and quality and a high proportion of unanswered questions, missings, inconsistent answers and uniform response patterns.

4.6.1. Outliers

We looked for *outliers* in all items, but most items were designed in a way that made outliers impossible, i.e. students could not exaggerate their answers. Where we found outliers, we dropped extreme values such as "28 brothers and sisters" which is unlikely to be true even in a country like El Salvador. Some of the students answered the questions about extinct animals in El Salvador with "dinosaurs" which was funny but the item was useless anyway and we dropped it.

Much more important was the analysis of inconsistent response patterns, which posed a serious thread to our data quality.

4.6.2. Response Patterns

There are three response characteristics that cast doubt upon the seriousness of the respondent and the quality of the data he or she generated: (1) *uniform response vectors*, i.e. a person responds most questions with the same category, e.g. the respondent puts all

crosses on the right, the middle or on the left side of a block of questions; (2) *inconsistent responses*: persons that show conflicting responses, i.e. they give the same values to reversed coded items; (3) *extreme categories*, i.e. persons that endorse a high number of rare or extreme categories.

A general *trend towards approval or rejection* and *response patterns* were observed in about 10 percent of the observations. Such reaction patterns are an indication of lack of attention and fatigue, carelessness and disinterest. The length and complexity of the instrument have certainly reinforced these trends. However, the patterns were rarely observed throughout a person's questionnaire but rather within one block of questions or on a specific page. This suggests ups and downs in concentration and interest. So we eliminated the patterns by block overwriting bad answers with missings.

We checked for inconsistent responses and suspicious response patterns and eliminated patterns that revealed the respondent did not really read and answer the questions. Uniform response patterns like 111111, 222222, 555555, 123456, 654321 in a 6-question-block were replaced with missings *before* we reversed item polarization. With this procedure we eliminated uniform response patterns and inconsistent answers at the same time. Of course this procedure results in higher item-non-response rates. But the means of the items shifted in the right direction and we improved standard deviations. We demonstrate these effects on just one item, the *EnvA_worry_03*-variable ("The gravity of the ecological problems are exaggerated by politicians and ecologists."). The number of observations decreased from 1735 to 1598, the mean from 3.63409 to 3.472534 and the standard deviation improved from 1.553086 to 1.510147.

This cleaning process could, of course, be significantly improved and refined, but the effort was disproportionate to the yield.

We had to analyze the group of the pattern-responders more in detail because they probably differ from those who answered all the questions carefully. To control the frequency and structure of the response patterns we generated an index of patterns (*IND_Patterns*). We counted the number of question blocks a respondents could have answered with a pattern. The lowest possible score was 0 blocks, the highest 15. We reduced this index to four categories (0 = 0 blocks, 1 = 1 − 5, 2 = 6 − 10, 3 = 11 − 15). As we can see in table 18, p. 41, the frequency of suspicious response patterns increased in W2. In W1 76.49% of the respondents showed no suspicious response patterns at all whereas in W2 this share decreased to 71.14%. Although this decline is not very pronounced, it suggests that some students might have been tired and bored of answering the same questionnaire again.

As we expected, we found better response quality among female than among male students (table 19, p. 41). Girls seem to have taken the survey more seriously than boys which is consistent with our own field observations and the feedback from our surveyors.

There were considerable differences in pattern occurencies between the strata with highest 0-proportions in the semi-rural stratum and lowest in the urban stratum. The big school *Instituto 14 de Julio de 1875* in San Francisco Gotera (INGO) seems to have had the biggest problem with discipline and resistance (table 20, p. 42).

Astonishingly there are pattern differences between the groups of self-declared poverty, too. The poorest had the lowest 0-pattern-rates and the highest 11 − 15-rates. This could be a hint that the poorest students had the biggest problems with the questionnaire because they had the worst reading skills and lowest concentration. Another explanation

Table 18: Patterns

Patterns Index	W1	W2	Total
0 blocks	1715	1491	3206
	76.49	71.14	73.91
1 - 5 blocks	428	459	887
	19.09	21.90	20.45
6 - 10 blocks	77	117	194
	3.43	5.58	4.47
11 - 15 blocks	22	29	51
	0.98	1.38	1.18
Total	2242	2096	4338
	100.00	100.00	100.00

might be, that they showed stronger tendencies to approve and not to contradict, whereas the rich might have had more problems with discipline and obedience. Nevertheless, with 73.30% of 0-pattern-responses the rich were still far better than the poorest with 64.10%.

Table 19: Patterns by Wave & Sex

Patterns	Male	Female	Total	Male	Female	Total
		W1			W2	
0 blocks	738	939	1677	642	819	1461
	73.14	79.85	76.75	68.66	73.85	71.48
1 - 5 blocks	212	201	413	211	229	440
	21.01	17.09	18.90	22.57	20.65	21.53
6 - 10 blocks	45	29	74	71	43	114
	4.46	2.47	3.39	7.59	3.88	5.58
11 - 15 blocks	14	7	21	11	18	29
	1.39	0.60	0.96	1.18	1.62	1.42
Total	1009	1176	2185	935	1109	2044
	100.00	100.00	100.00	100.00	100.00	100.00

In general, the response pattern detection gave us information about the question at which school, in which class and with which group of students the instrument worked better or worse. Of course, we lost a lot of item-response. However, the following factor analyses shows that the data still does contain valuable information, i.e. that the items that belong together in content loaded on the same factors. We can therefore assume that the pattern elimination significantly improved the quality of the data possibly providing more reliable results. However, this further limits the representativity of our data as unit- and item-non-response are far from random.

Table 20: Patterns by Stratum

Patterns	Urban	RC 1	RC 2	Semi urban	Semi rural	Rural	Total
0 blocks	722	384	405	407	674	614	3206
	66.42	81.53	76.27	65.75	82.20	75.80	73.91
1 - 5 blocks	275	75	99	160	124	154	887
	25.30	15.92	18.64	25.85	15.12	19.01	20.45
6 - 10 blocks	71	8	21	40	18	36	194
	6.53	1.70	3.95	6.46	2.20	4.44	4.47
11 - 15 blocks	19	4	6	12	4	6	51
	1.75	0.85	1.13	1.94	0.49	0.74	1.18
Total	1087	471	531	619	820	810	4338
	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 21: Patterns by Grades of Poverty

Patterns	Very poor	Poor	Somewhat poor	Not poor	Total
0 blocks	75	336	954	302	1667
	64.10	77.06	79.24	73.30	76.86
1 - 5 blocks	32	89	205	84	410
	27.35	20.41	17.03	20.39	18.90
6 - 10 blocks	7	6	33	24	70
	5.98	1.38	2.74	5.83	3.23
11 - 15 blocks	3	5	12	2	22
	2.56	1.15	1.00	0.49	1.01
Total	117	436	1204	412	2169
	100.00	100.00	100.00	100.00	100.00

4.6.3. Missings Management

We defined globals in STATA to impute missings and manage missings values. The significance of the missing codes is shown in table 22, p. 43. The advantage of steering failures by globals is the comfortable adaptability later on.

Table 22: Missings

Global Name	Missing	Comment	Variable type
DKnn	.a	"I don't know"	num-vars
DKns	.a I don't know	"I don't know"	string-vars
Nan	.b	"No answer"	num-vars
Nas	.b no answer	"No answer"	string-vars
Nopn	.c	"No opinion / no idea"	num-vars
Nops	.c no opinion, no idea	"No opinion / idea"	string-vars
SMn	.d	"System missing"	num-vars
SMs	.d system missing	"System missing"	string-vars
BAn	.e	"Bad answer"	num-vars patterns
BAs	.e bad answer	"Bad answer"	string-vars patterns
NoPn	.x	"Not asked"	num-vars
NoPs	.x not asked	"Not asked"	string-vars
NoPan	.y	"o Workshop"	num-vars
NoPas	.y no workshop	"No Workshop"	string-vars
NoAppln	.z	"Not applicable"	num-vars
NoAppls	.z not applicable	"Not applicable"	string-vars

4.6.4. Weights

The sample was stratified by strata and schools. It is not self-weighted. We chose a disproportional sampling approach in which rural and small grammar schools were over-represented compared to urban grammar schools. This was done in order to achieve a sufficiently high number of cases in small communities to be able to detect potential differences between centers and the rural periphery. To maintain representativity on order to produce unbiased estimates of population statistics, we should have used sampling weights. But as discussed, representativity was already lost. The difference between official school lists and respondents in the survey were high and we had no reliable basis for a serious calculation of weights. Therefore, we have waived the calculation of weights.

4.7. Data Analysis of ATT and ITT

We can use 1539 students for the *cross-sectional measurement* of the *average treatment effect of the treated ATT*. Expecting weak effects we will start our analysis with the compliers only using mean value comparisons of the treatment groups with *t-tests* and *OLS-regression*. As we have seen, the defiers most probably differ from the compliers. As a consequence, we can not assume unconfoundedness and under usual conditions the ATT would be biased – probably upwards because compliers might be the more conscientious

students. Now, the advantage of our design comes into play: The composition of the dropout groups will be similar in both treatment groups, meaning that the bias should be the same in both groups. As a consequence, we can assume that, in our setting, ATT-estimators will not be biased.

908 identified students can be used for longitudinal modeling of the ATT. We will compute fixed-effect-regressions (FE), i.e. within effects between of W1 and W2.

The ATT shows us the treatment effect on the compliers, but it does not show us the effects of the program as a whole which is the core question of project evaluations: If a large share of selected students does not participate in the workshops, the latter will clearly be less effective. Hence, we should base our evaluation not only on the treated, but on the initial treatment assignment to compute the *intention to treat effect ITT*. We will start with the estimation of the ATT effects. If they are too weak is does not make much sense to estimate the ITT because they are supposed to be even weaker.

5. Results

As we have seen, the study-design opens two paths of analysis to isolate the causal effect of the treatments:

1. We can compare the outcomes, i.e. the *differences* of EnvA and GenA *between* the treatment groups T1 and T2 in W2. Remember that in our design the treatment groups act as control group for each other. Thus, mean value comparisons with t-tests or OLS regressions can be used to measure the ATT. We can compute the ITT including the defiers, too. As we expect very weak effects, we concentrate on the ATT.

2. We can compare the *changes* of outcomes of the identified individuals between W1 and W2 using fixed-effect-regressions (FE).

3. A third option is to compare different groups of the sample: (1) those who participated in both W1 and W2 *and* in T1 or T2, (2) those who participated in both W1 and W2 without any T, and (3) those who participated only in W2. This procedures can teach us more about the effects of the instrument, although if we have to face a high selection bias. As we will see, the results are astonishing.

Let us get a short overview of our cross-sectional W2 sample again (table 23, p. 45): 1539 (73.4%) individuals of the 2096 respondents said they participated in T1 or T2, 557 (26.6%) said, they did not (defiers). 908 individuals (43.32%) individuals could be identified in W1 and W2, while 1188 (56.68%) individuals could not. For the cross-sectional comparison between the T1 and T2 we can work with the 1539 compliers to estimate the *Average Treatment Effects on the Treated (*ATT*)*. The estimator will not be biased when the defiers in both groups do not differ systematically. There are few reasons to believe that the non-complier groups differ between the two treatments because the students did not know in advance which workshops they would attend.

Table 23: Participants

Workshops	1	2	_missing_b	_missing_y	Total
Not identified	348	283	90	249	970
	40.80	41.25	73.17	57.37	46.28
Identified	505	403	33	185	1126
	59.20	58.75	26.83	42.63	53.72
Total	853	686	123	434	2096
	100.00	100.00	100.00	100.00	100.00

5.1. Cross-sectional Mean Value Analysis

5.1.1. Scores

We first computed a mean value comparison of our knowledge scores on environmental and gender issues, T1 being the reference group. The participants of T1 showed a 1.7 percentage points better result in their environmental knowledge score and an almost 1 percentage points lower gender knowledge score than T2. The first result was significant on the 1%-level, the second was not far from reaching the 5% significance level.[45]

Table 24: T-Test on Scores

	Coef.	Std. Err.	$Pr(|T| > |t|)$	$Pr(T > t)$	Obs.
SC_EnvK	0.0174	0.0061	0.0046**	0.0023**	1539
SC_GenK	-0.0099	0.0056	0.0798	0.9601	1539

* p<0.05, ** p<0.01, *** p<0.001

How can we interpret these results? Students might have really learned what they were supposed to in the different workshops, even though the difference is very small.

5.1.2. Indices

When comparing our attitudes indices, we do not find significant results except in one index: Those who participated in the gender workshop seem to be a little bit more sceptic towards ecological problems (table 25 p. 46). The effect is weak but it suggests that – as expected – the skepticism about the existence of climate and environmental problems has decreased in the environment group in comparison to the gender group. The trivialization index shows, albeit not significant, differences in the same direction. All others indices show high p-values and are far from significant. If we look at the indices of self-declared environmental and gender behavior, we do not find any significant results either. So, we have to acknowledge that the workshops largely failed to produce the intended changes of attitudes and intentions to act.

---

[45] The results are somewhat difficult to interpret, because the coefficients indicate the difference from T2 to T1. A minus therefore means higher T2 values and lower T1 values and vice versa.

We controlled all item blocks and every single item separately. The result was: Most items were far from showing significant t-test results. By way of example, we report the outputs of just two item blocks. Table 27, p. 47 shows the t-test results of the environmental worry and the gender violence variables.

Table 25: T-Test on Attitudes and Behavior Indices

	Coef.	Std. Err.	Pr(|T| > |t|)	Pr(T > t)	Obs.
Attitudes:					
General Environmental Worries	-0.0014	0.0121	0.9052	0.5474	1335
Climate Change Concern	0.0051	0.0112	0.6468	0.3234	1331
Ecological Scepticism	0.0341	0.0149	0.0223*	0.0112*	1092
Ecological Trivialisation	0.0224	0.0145	0.1231	0.0616	1166
Ecological Indifference	-0.0141	0.0156	0.3651	0.8175	1273
Machismo	-0.0099	0.0131	0.4473	0.7763	1369
Gender Equality	0.0080	0.0105	0.4506	0.2253	1340
Classical Gender Roles	0.0039	0.0143	0.7850	0.3925	1360
Gender Violence 1	-0.0013	0.0132	0.9230	0.5385	1425
Gender Violence 2	-0.0058	0.0118	0.6242	0.6879	1429
Behavior:					
Green Consumerism	-0.0164	0.0120	0.1716	0.9142	1447
Waste Management	-0.0143	0.0087	0.1006	0.9497	1297
Eco-committment	0.0095	0.0144	0.5091	0.2546	1459
Misogynistic Male Behavior	0.0133	0.0187	0.4776	0.2388	603
Female Emancipation Index	0.0160	0.0166	0.3349	0.1674	811
Female Subordination Index	0.0055	0.0107	0.6042	0.3021	767

* p<0.05, ** p<0.01, *** p<0.001

Table 26: T-Test on Environmental Variables

	Coef.	Std. Err.	Pr(|T| > |t|)	Pr(T > t)	Obs.
General Environmental Worries:					
EnvA_worry_01	-0.0508	0.0667	0.4468	0.7766	1364
EnvA_worry_02	0.0161	0.0660	0.8072	0.4036	1357
EnvA_worry_03	-0.1010	0.0877	0.2501	0.8750	1214
EnvA_worry_04	-0.0156	0.0728	0.8305	0.5848	1289
EnvA_worry_05	0.0612	0.0809	0.4494	0.2247	1318
EnvA_worry_06	-0.0445	0.0618	0.4716	0.7642	1350
EnvA_worry_07	-0.0592	0.0678	0.3831	0.8085	1319
Worries about Climate Change:					
EnvA_CC_01	-0.1581	0.0814	0.0523	0.9738	1250
EnvA_CC_02	0.2491	0.0862	0.0039**	0.0020**	1403
EnvA_CC_03	0.0283	0.0653	0.6650	0.3325	1418
EnvA_CC_04	0.0883	0.0640	0.1680	0.0840	1424
EnvA_CC_05	-0.0508	0.0625	0.4164	0.7918	1367
EnvA_CC_06	0.0320	0.0871	0.7132	0.3566	1216
EnvA_CC_07	0.0840	0.0886	0.3430	0.1715	1350
EnvA_CC_08	0.1177	0.0794	0.1383	0.0691	1335
EnvA_CC_09	0.0341	0.0869	0.6951	0.3476	1360
EnvA_CC_10	0.0770	0.0915	0.4003	0.2002	1311

* p<0.05, ** p<0.01, *** p<0.001

Table 27: T-Test on Significant Items

	Coef.	Std. Err.	Pr(|T| > |t|)	Pr(T > t)	Obs.
Belief in Natural Causes of CC	-0.2491	0.0862	0.0039**	0.9980**	1403
Rich Countries are responsable	0.1865	0.0802	0.0202*	0.0101*	1479
Small Countries have to adapt	-0.1723	0.0791	0.0295*	0.9852*	1459
Garbage-conscious Shopping	-0.1497	0.0740	0.0433*	0.9783*	1475
Proper Waste Disposal	-0.1793	0.0685	0.0089**	0.9955**	1475
Equality of Womens' Earnings	-0.1625	0.0499	0.0011**	0.9994**	1281
Equality of Women in Politics	-0.1195	0.0555	0.0316*	0.9842*	1149
Equality of Womens' Fortunes	-0.1088	0.0520	0.0367*	0.9816*	1166
Woman Role: Household	0.1832	0.0922	0.0472*	0.0236*	1436
Discussions on Gender Issues	0.1951	0.0956	0.0416*	0.0208*	805

* p<0.05, ** p<0.01, *** p<0.001

However, there are a couple of items that show significant t-test results. These results are quite interesting (table 27, p. 47). Surprisingly, the belief in natural causes of climate change is significantly higher rather than lower in T2 (EnvA_CC_02). Similarly, students in T2 believe less in the sole responsibility of richer countries and the political possibilities of smaller countries. Also quite counterintuitive are the better scores of the T2 group in garbage-conscious shopping and proper waste disposal. We would have expected results in the opposite direction. On the other hand, the differences in knowledge about gender equality are as expected: The T2 group knows more than the T1 group about distribution inequality in El Salvador. Moreover, they belief less frequently that the right place for women is the household. In summary, some of the significant items show effects in the expected direction, but the results should be interpreted with care. With such a large number of items, it is to be expected that some of them are of purely random significance.

5.2.   Cross-sectional Regressions

So far, the t-test only showed us if the means our items, scores and indices differ significantly between T1 and T2 or not. They do not tell us anything about potential heterogeneity with regard to certain subgroups of the population. It is possible that girls reacted differently to the gender workshop than boys, that urban students reacted differently to the environmental workshops than the rural students. Unfortunately, we could use only the few control variables we collected in W2 such as sex and marginalization (stratum) to show the treatment effect differences between boys and girls and between the regions. Other control variables that could have had an impact like poverty level or education of parents were not available in W2 and could thus not be used in our cross-sectional analysis.

We estimated OLS regression models of our scores. Remember that the reference category are male students of stratum 1 who participated in T1. In model01 we see the significantly lower environment knowledge score of the gender workshop participants, we already know from the t-test. Moreover, women and men started at the same level in the environmental knowledge score, but women were significantly better informed in gender issues from the outset (model05) the interaction effect not beeing significant. Similarly, there are no differences in knowledge scores between the strata neither before nor after the treatment.[46]

---

[46]   Note that we show interaction effects only, but we computed all effects.

Table 28: Environmental Knowledge Score Regression

	Model01	Model02	Model03
Treatment	-0.0174*	-0.0084	-0.0170
	(0.0184)	(0.3756)	(0.4238)
Sex:			
Women		-0.0021	-0.0048
		(0.7765)	(0.4640)
T2#Women		-0.0163	-0.0127
		(0.0762)	(0.1719)
Strata:			
T2#Rural Center 1			0.0201
			(0.3252)
T2#Rural Center 2			0.0373
			(0.4259)
T2#Semi urban			0.0320
			(0.1658)
T2#Semi rural			0.0033
			(0.8930)
T2#Rural			-0.0264
			(0.2379)
N	1539	1511	1511
R^2	0.005	0.008	0.031
adj. R^2	0.005	0.006	0.023

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 29: Gender Knowledge Score Regression

	Model04	Model05	Model06
Treatment	0.0099	0.0108	0.0070
	(0.3837)	(0.1785)	(0.7497)
Sex:			
Women		0.0334**	0.0298**
		(0.0035)	(0.0029)
T2#Women		0.0004	0.0045
		(0.9690)	(0.5648)
Strata:			
T2#Rural Center 1			0.0104
			(0.7584)
T2#Rural Center 2			-0.0034
			(0.9282)
T2#Semi urban			0.0002
			(0.9950)
T2#Semi rural			-0.0096
			(0.6431)
T2#Rural			0.0101
			(0.6350)
N	1539	1511	1511
R^2	0.002	0.025	0.046
adj. R^2	0.001	0.023	0.038

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All other index regressions did not show significant results for the interaction effects. We show just one table, the gender equality index regressions (table 30, p. 51). We can see that women started on a significantly different level: They were, as one would expect, more in favor of gender equality than their male peers. Yet, that doesn't mean that they changed their attitudes after the workshops significantly more than their male fellow students. Indeed, the interaction term does not show significant coefficients. Moreover, there are some differences in the starting levels between the strata (not reported) regarding gender equality attitudes: In the rural center 1 they are somewhat higher than in the reference category of the urban school in Gotera. As to the respective interactions, the coefficient plots (figures 13, p. 52, and 14, p. 52) show that none of the index coefficients are significant.

Table 30: Gender Equality Index Regression

	Model07	Model08	Model09
Treatment	-0.0080	-0.0057	-0.0197
	(0.5484)	(0.7533)	(0.5690)
Sex:			
Women		0.0797***	0.0745***
		(0.0000)	(0.0001)
T2#Women		0.0016	0.0128
		(0.9301)	(0.5036)
Strata:			
T2#Rural Center 1			0.0193
			(0.6494)
T2#Rural Center 2			-0.0007
			(0.9925)
T2#Semi urban			0.0420
			(0.1788)
T2#Semi rural			-0.0095
			(0.7946)
T2#Rural			0.0028
			(0.9407)
N	1340	1315	1315
R^2	0.000	0.044	0.071
adj. R^2	-0.000	0.042	0.062

p-values in parentheses

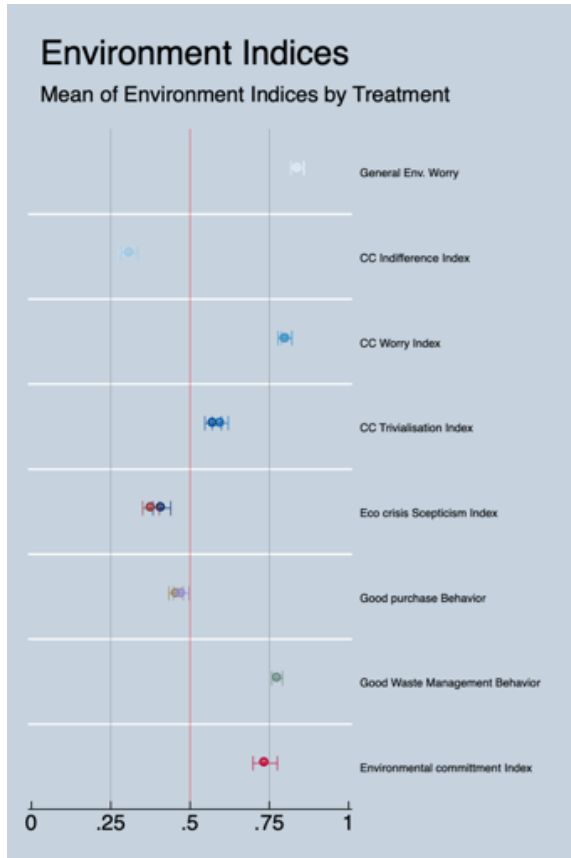* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
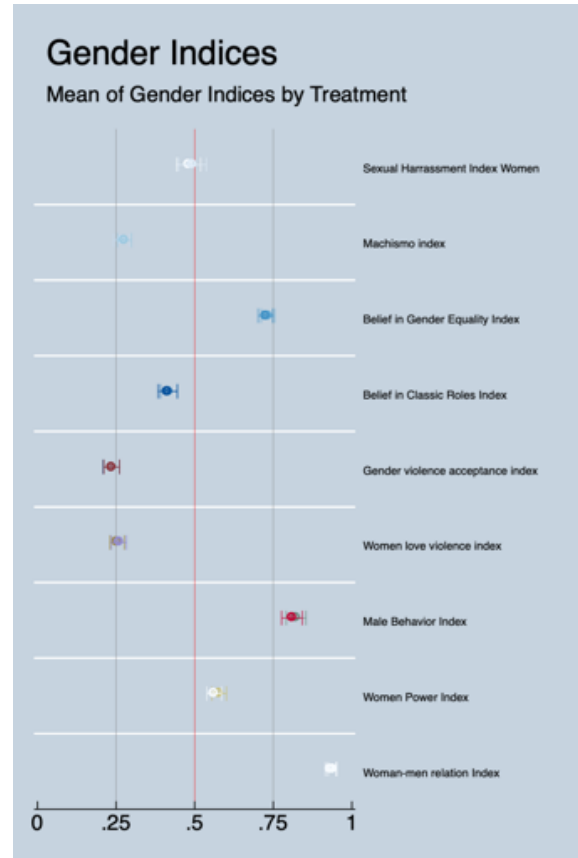
Figure 13: Environment Indices

Figure 14: Gender Indices

5.3. Fixed Effect Regressions

So let us have a look at the identified individuals. To compare the within-changes between W1 and W2 we needed to identify individuals. We were able to identify 1126 individuals of whom 908 (80.64%) participated in a treatment, 218 (19.36%) did not (table 16, p. 38). We are analyzing these individuals with fixed effect regressions.

We can observe that the *wave effect* is highly significant in almost all models, indices and items, but that the treatment effects, i.e. the difference between T1 and T2 is virtually never significant except the environmental knowledge score (tables 31 and 32, p. 54). The results seem to be robust as they confirm the results of the cross-sectional analysis. We interact the treatments with stratum and poverty in the environmental knowledge score and with confession, region and poverty in the gender knowledge score.[47]

Clearly, students changed their answering behavior between W1 and W2 regardless of treatment participation. This wave-effect can be interpreted as learning effect of the whole implementation, i.e. the treatments *and* the two surveys. This assumption is supported by the results presented in the following section 5.4, p. 55.

---

[47]   We computed FE-regressions for all indices and a couple of items without finding any treatment effects but substantial wave-effects; see: STATA-file PISCO_3_3.do

Table 31: FE-Regressions Environment Knowledge Scores

	FE_Model_10	FE_Model_11	FE_Model_12
T2	-0.0231*	-0.0161	0.0089
	(0.0187)	(0.1801)	(0.7964)
W2	0.0470***	0.0470***	0.0454***
	(0.0000)	(0.0000)	(0.0000)
Strata:			
T2#R Center 1		0.0038	0.0152
		(0.8671)	(0.4825)
T2#R Center 2		0.0080	0.0160
		(0.7159)	(0.4853)
T2#Semi urban		-0.0115	-0.0025
		(0.5847)	(0.9086)
T2#Semi rural		-0.0188	-0.0147
		(0.2329)	(0.3676)
T2#Rural		-0.0119	-0.0051
		(0.4079)	(0.7587)
Poverty:			
T2#Poor			-0.0345
			(0.3502)
T2#Little poor			-0.0392
			(0.1957)
T2#Not poor			-0.0166
			(0.6240)
N	1816	1816	1776
R^2	0.089	0.091	0.088
adj. R^2	0.088	0.087	0.082

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 32: FE-Regressions Gender Knowledge Scores

	FE_Model_13	FE_Model_14	FE_Model_15	FE_Model_16
T2	-0.0004	-0.0078	-0.0125	-0.0037
	(0.9545)	(0.4874)	(0.4307)	(0.8830)
W2	0.0195***	0.0190***	0.0190***	0.0184***
	(0.0000)	(0.0000)	(0.0000)	(0.0001)
Confessions:				
T2#Catholic		0.0071	0.0046	0.0049
		(0.5695)	(0.7382)	(0.7316)
T2#Protestant		0.0148	0.0096	0.0069
		(0.3180)	(0.5703)	(0.6881)
T2#Adventist		-0.0112	-0.0005	0.0007
		(0.7088)	(0.9890)	(0.9842)
T2#Other		0.0152	0.0174	0.0174
		(0.7360)	(0.7124)	(0.7007)
Strata:				
T2#R Center 1			0.0012	0.0064
			(0.9363)	(0.7035)
T2#R Center 2			0.0351*	0.0415*
			(0.0282)	(0.0132)
T2#Semi urban			-0.0120	-0.0085
			(0.4936)	(0.6222)
T2#Semi rural			0.0090	0.0127
			(0.5194)	(0.3550)
T2#Rural			0.0188	0.0256
			(0.2768)	(0.1439)
Poverty:				
T2#Poor				-0.0269
				(0.2676)
T2#Little poor				-0.0111
				(0.6420)
T2#Not poor				-0.0047
				(0.8315)
N	1816	1812	1812	1772
R^2	0.038	0.038	0.045	0.046
adj. R^2	0.037	0.035	0.039	0.038

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.4. Instrument Effects

We tried to find out if the large surveys themselves could have had any effect on the outcomes. Thanks to the test-retest design, we were able to measure the learning effects of the questionnaire itself and analyze the possibility that participation in the survey, independently of the workshops, triggered reflections on both topics and changed the response behavior (Diekmann, 2014, 252). We created the variable *Treatment3* to isolate the effects of survey participation as a sort of treatment. The "least treated" students, i.e. those who neither participate in W1 nor in any of the workshops, served as a reference group. The results are shown in tables 33 and 35, p. 56 and 57.

Many of our indices and all our scores show significant or even highly significant differences for the group of the *most treated*, i.e. the students who filled out survey W1 and W2 *and* participated in T1 or T2 – compared to those who only attended the second wave, but neither a workshop nor the first wave. Surprisingly, the group that participated in both surveys, but not in any of the workshops (*+W1 −T +W2*), also showed significantly higher Env-knowledge-scores and some higher gender indices than the reference group. How can these results be interpreted?

It would be premature to claim that these effects are due to learning outcomes, changes of attitudes and improvements of self-declared behavior. First, large parts of the effect might be attributable to a change in responsiveness to social desirability – i.e. students might have figured out how they were expected to answer by the second wave and tried to live up to it. Similarly, learning effects might have occurred in the handling the surveys rather than in actual opinions: During the second wave, students might have been better-prepared for the intensity of the questionnaire, the unusual nature of the questions, the impressive seriousness of the whole event and perhaps by respect for this unknown actor behind the whole study, the University of Bern, which was so interested in their opinion. A second factor might have been self-selection: Those who participated diligently in all the surveys and the workshops might also be those with greater interest in the two topics of the workshops.

Bearing these restrictions in mind, we can conclude that the "treatment effect of the instrument" might have been stronger than the participation in a particular workshop. Interestingly, this effect is particularly pronounced for those indiviuals that also participated in either of the workshops.

Table 33: Instrument Effects on Environment Indices

	IND01	IND02	IND03	IND04	IND05	IND06	IND07
+W1 - T +W2	0.043	0.013	-0.009	-0.053	-0.004	-0.013	0.045
	(0.090)	(0.353)	(0.755)	(0.157)	(0.772)	(0.436)	(0.177)
- W1 +T1 +W2	0.010	-0.000	-0.056	-0.012	-0.009	0.003	0.076
	(0.752)	(0.981)	(0.184)	(0.794)	(0.817)	(0.914)	(0.099)
- W1 +T2 +W2	0.007	-0.010	-0.070	0.037	0.047	-0.021	-0.009
	(0.850)	(0.832)	(0.138)	(0.501)	(0.221)	(0.325)	(0.884)
Most treated:							
+W1 +T1 +W2	0.116***	0.081***	-0.041	-0.108***	0.012	0.015	0.167***
	(0.000)	(0.000)	(0.101)	(0.000)	(0.218)	(0.296)	(0.000)
+W1 +T2 +W2	0.120***	0.075***	-0.077**	-0.096***	0.026	0.032	0.164***
	(0.001)	(0.001)	(0.002)	(0.000)	(0.063)	(0.068)	(0.000)
N	1773	1745	1423	1683	1915	1695	1950
R^2	0.035	0.025	0.012	0.020	0.003	0.011	0.049
adj. R^2	0.032	0.023	0.009	0.017	0.001	0.008	0.047

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 34: Instrument Effects on Gender Indices

	IND08	IND09	IND10	IND11	IND12	IND13	IND14
+W1 - T +W2	-0.005	0.011	-0.036	-0.055*	0.011	-0.015	-0.046**
	(0.866)	(0.533)	(0.336)	(0.031)	(0.486)	(0.589)	(0.005)
- W1 +T1 +W2	-0.012	0.010	-0.070*	-0.075*	0.031	-0.012	0.033
	(0.453)	(0.690)	(0.037)	(0.020)	(0.064)	(0.785)	(0.118)
- W1 +T2 +W2	0.012	-0.005	-0.072	-0.025	0.014	-0.169*	0.013
	(0.812)	(0.911)	(0.127)	(0.473)	(0.692)	(0.013)	(0.830)
Most treated:							
+W1 +T1 +W2	-0.084*	0.072**	-0.073***	-0.116***	-0.035*	0.030	0.009
	(0.011)	(0.007)	(0.001)	(0.000)	(0.011)	(0.203)	(0.584)
+W1 +T2 +W2	-0.073**	0.065**	-0.077*	-0.117***	-0.026	0.027	-0.010
	(0.005)	(0.005)	(0.034)	(0.000)	(0.180)	(0.303)	(0.491)
N	1791	1763	1795	1896	1904	813	1029
R^2	0.021	0.022	0.008	0.022	0.008	0.022	0.008
adj. R^2	0.018	0.019	0.005	0.020	0.005	0.016	0.003

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 35: Instrument Effects on Scores

	SC_EnvK	SC_GenK
+W1 -T +W2	0.019*	-0.007
	(0.045)	(0.333)
-W1 +T1 +W2	0.029	-0.002
	(0.068)	(0.834)
-W1 +T2 +W2	0.015	0.005
	(0.260)	(0.761)
Most treated:		
+W1 +T1 +W2	0.066***	0.043***
	(0.000)	(0.001)
+W1 +T2 +W2	0.047***	0.051***
	(0.000)	(0.000)
N	2096	2096
R^2	0.036	0.047
adj. R^2	0.034	0.045

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6. Discussion, Recommendations and Conclusion

In the final discussion I will review the shortcomings of PISCO and the RdE-program and make some recommendations for future projects.

6.1. Shortcomings

How can we explain our results? There are four possible answers to this question:

(1) the workshops did not work, so there was little or no effect; (2) the workshops worked, but the effects were too small to be detected; (3) the instrument did not work, so we can not determine if the workshops worked or not; (4) the management of the research implementation did not work, so we can not say anything neither about the quality of the workshops nor the quality of the instrument.

All these answers may be partially correct and they concern both the RdE-project and the research-project PISCO and its implementation. Let us have a look at the most important shortcomings:

1. The RdE-workshops did not work.

   ▪ Workshop intensity: The intensity of the workshop was most probably far too low. We can not expect great changes in attitudes after three hours of group work and games. This does not mean that the workshops did not address important issues and that they did not touch the young people. Given the low intensity of he workshops, it is surprising that they produced some – though

modest – improvements in knowledge. However, if attitudes or behaviors are to be changed, a more prolonged intervention might be more promising.

That's why, the NGO should think about how to intensify the treatments, e.g. by realizing several half- or even full-day workshops over several months or a one-week workshop.

- Folk teacher training: The folk teachers were trained in a couple of crash courses and weekend-workshops. Although their motivation was overwhelming, their content knowledge and pedagogical skills might have been incomplete. The NGO can, of course, work continuously on this matter, e.g. by intensifying their training.

- Folk teachers' age: Discipline is a major problem in the classrooms of El Salvador. Basically, respect is more likely to be paid towards grey-haired men. As a consequence, students probably perceived the folk teachers as peers, not as respectable teachers. The youth of the folk teachers may have been an important strength, but at the same a weakness of the program.

  Nevertheless, the NGO should think about how to improve discipline, respect and seriousness during the workshops. Perhaps the mere presence of the class teacher would be sufficient. But even when I visited a couple of workshops and grabbed some photos students still were fidgety despite the presence of a gray-haired white man. And it did not look much different in the normal classes we visited, where older teachers were teaching.

  Another approach would be to realize the RdE-program only with people who are really interested in the topics and participate voluntarily. We could expect better learning outcomes and considerable spillover effects.

- Information: The idea of the NGO was the implementation of new pedagogical methods based on discussion and reflection rather than mere information transfer. The students were supposed to work out the problems by themselves and form their own opinion about them, although students might have actually only reproduced the bit of gender and environment awareness they already had. However, this important participative component may have been beneficially complemented by a transfer of important facts and figures. For example, the NGO might supplement the workshops with a concise information block. From a scientific point of view, this would also considerably facilitate the task of measuring learning outcomes.

2. The workshop worked, but the effect was too small to be detected.

- Sample size: If the intensity of the workshop was low that doesn't mean that the did not have any effect. Everything has some, though negligible, effect on something. Either the instrument was not sensitive enough, the measurement was not accurate enough or the sample size was too small. To answer the last point precisely, we should have run a power analysis.

- Power analysis answers the question what minimum effect size we could have measured with our given sample or what sample size we would have needed to detect a desired magnitude of effect. We assumed that the large number

of students in our frame sample would make power analysis superfluous. As a result, we simply do not know how large the minimum detectable effect actually was. Although we do not believe that the results would have changed much with a larger sample, a power analysis should still be part of every study.

3. The research instrument did not work.

The partial instrument failure might have several causes. First, the instrument was not suitable for this particular target group.

- Alien topics: The discourse of environmental and gender problems have not yet arrived in the consciousness of most people, and even less so of the young, who are cut off from most sources of information. Other problems such as poverty, health problems, insecurity are at the forefront of everyday life. Environment and gender problems are "esthetic" problems for them. Possibly, most students had never been asked for their opinions and attitudes on environmental and gender issues previous to the survey application. Many of them may have never really thought or heard about these kinds of societal problems. As a result, they have not yet had the opportunity to form an attitude or opinion at all. Consequently, they may have been hopelessly overwhelmed with the numerous questions about alien topics.

- Literacy: Many students may have had problems to read such a large amount of demanding questions. Even at the high school levels, the literacy skills of many students can be assumed to be insufficiently developed to cope with the required task.[48] This problem opens up a wide field for further important research.

- Concentration and discipline: The majority of the students had never participated in a survey in their life. The mere completion of the questionnaire was a challenge. It was difficult for most students to sit quietly for one hour and concentrate on their work. According to our own on-site observation and interviews, the discipline in the classrooms is generally lousy. So, some students gave up quickly and started distracting others. Many reacted with inconsistent response patterns and uniform response vectors or simply with non-response.

- Truancy: We underestimated school truancy by far. It seems to be quite usual not to go school or not to enter the classroom when students do not feel like it or have better things to do. As a consequence, we had to put up with high rates of non-compliance and attrition.

Foreign topics, poor literacy, lack of concentration and discipline as well as truancy had a severely detrimental impact on the quality of the data. As a consequence, we observed much noise and many suspicious response patterns in the data.

Second, the instrument was overcharged by far.

- Comprehensiveness: The questionnaire included far too many questions. It may have overstrained the respondents' capacities and willingness. More than

---

[48] The assumption is not unfounded: The letters the scholarship holders of *Consciente* – usually students with top grades – have to write to their godparents annually are of worrying linguistic quality.

50% of the students in W1 and more than 60% in W2 said it took them more than 45 minutes to complete the questionnaire. This may have had (1) a negative effect on the quality of the data and (2) the instrument-effect superimposed the effect of the treatments. Therefore, we might have done better with, say, 20 percent of the questions and assumining a maximum concentration span of 15 to 20 minutes.

- Adequacy: We should have adapted the questions much more diligently (1) to the cultural context and the level of the students and (2) to the contents of the workshops. Incomprehensible and inappropriate questions may have reduced students' willingness and motivation to answer faithfully.

- Pretest: Due to organizational problems and delays, our pretest was carried out in a great hurry. The coordination between the researchers and project managers in Switzerland and the team of project managers and implementers was a little messy. The workshop contents were not entirely defined at the beginning of the pretest. Only immediately before the intervention did the teaching contents become known, meaning that no adaptations of the instrument to the exact contents of the workshops were possible.

- Preliminary investigation: Every scientific project that wants to explore and quantify "things" like values, attitudes and intentions to act by the means of a survey must have a fairly clear and precise idea of the interviewees' world, their everyday lives, their stock of knowledge, their living habits, their value system and basic attitudes, their scope of action and ways of acting. Without this context knowledge, it is simply not possible to ask meaningful questions. Therefore, it would be appropriate to use comprehensive qualitative pre-investigation in order to explore and experience the world respondents live in (c.f. Dannecker & Englert (2014) or Denzin & Lincoln (2005)). We should have started the project adapting exploratory observational methods like participatory observation, guided interviews (Slezak, 2014) and ethnographic interviews (Flick, 2017, 193ff). Without these preliminary investigations, it is impossible to *understand* the foreign cultural context and to design and realize well-adapted development projects and research instruments.

Third, we made some methical mistakes on the scientific level:

- Identification: Contrary to our expectations, identification of individuals in W1 and W2 proved to be a major problem. We should have developed a reliable method to identify students without arousing their mistrust (e.g. an inconspicuous bar-code or ID-numbers on the questionnaires and personal delivery to each student). However, the organizational challenge posed by such a procedure would have been substantial and possibly unsurmountable.

- Questionnaire composition: If we had known about the low identification rate in advance, we could have changed the composition of the questionnaires. It would have been more useful to ask the socio-economic questions *after* the treatments. We could then have included more control variables into our cross-sectional analysis. Instead, we could have asked the school evaluation questions for MINED in W1.

- Scales: We observed, that yes-no-scales worked pretty well, but items with a 6-level Likert-scale did not. Salvadoran students may not be used to this kind of questions. Maybe they are not prepared to think in so many nuances and to give subtle answers. They are probably more used to black & white-, yes & no-, right & wrong-thinking. This problem should be examined more in detail. So, maybe 4-level scales would have worked better.

  We omitted the "I don't know"-category consequently. Perhaps that was not the best way of nudging the students to take a stand. Rather, it mixed honest answers with empty answers without us being able to separate them.

4. The management of PISCO did not work.

   If the management of a research program fails, then even the best treatments and the best instruments will be useless.Therefore, implementation errors endanger the success of the whole research project.

   - Coordination: It is in the nature of things that communication and cooperation across the ocean and across the cultural gap is difficult. So the project management and coordination between Switzerland and El Salvador was often affected by cultural differences and misunderstandings. For example, time management and planning methods differ greatly. The project management from Switzerland was certainly too unsystematic, inconsistent and sometimes to restrained. The lack of leadership, especially when it came to meeting task lists, timetables and deadlines, caused many problems. However, decent leadership is a difficult task in the development environment where paternalistic attitudes have to be avoided.

     Hence, much more attention needs to be paid to the project management. Perhaps the consistent use of a project management tool would make sense.

   - Information and marketing: Although we could count on the official support of MINED, the data collection team often failed to encounter sufficient support from directors and ordinary teachers. Often the latter were unaware or had simply forgotten that a survey was to take place. So they were absent, the students were not around or the whole school was closed due to an event. Hence, previous information visits (rather than mere telephone calls) or even a kind of "contract of commitment" including a timetable would be desirable.

     In addition, the project should be better sold in order to inspire all parties involved. The NGO should think about an active marketing approach.

   - Scientific gap: It was difficult to explain to the local program staff what an RCT is and why the random selection of schools and classes and the random assignment of the workshops should be strictly adhered to. The data collectors were happy to fulfill their basic task and come back with the surveys of the day. So data collection did often not meet the scientific requirements. Data registration, too, was not sufficiently stringent. Although most of the team worked hard and carefully, a few of the young data registrators might have taken it "on the lighter shoulder'". After a few hours of typing in data, of course, fatigue and concentration might have interfered with the quality requirements

of the data. Therefore, we have to assume a considerable data entry error rate which can only partly be corrected for.

Due to the limited budget, a systematic control of data and replication was not possible. This problem should definitely be taken into account in future investigations. For further research projects, consistent training of the local staff on topics such as research designs, methods and scientific requirements would be advisable.

- I'd like to make one last comment: I started the project without any *personal* experience in the country. All my information and knowledge was second-hand. I had never *experienced* life in El Salvador. I had never *felt*, not even in the slightest bit, what it is like to be a Salvadorian boy in the highlands of Perquín or a girl in the poor suburbs of San Francisco Gotera. My world and theirs are light years apart from each other. Economist Gérard Roland puts it this way: culture is "the set of values and beliefs people have about how the world (both nature and society) works and the behavioral norms derived from that set of values." (Roland, 2014, 289). The cultural gap between a 61-year old, well-off Swiss researcher and a 18-year old Salvadorian boy or girl who grew up in a hut under banana trees, who experiences chronicle poverty, who never even left his region and doesn't know a lot about the world – this cultural gap could hardly be wider and perhaps it is insurmountable. This is one of the most serious problems in development cooperation and in the design and evaluation of projects – be it interventions or research studies – that really work.

Each implementation error incorporates noise into the data and reduces their reliability and informativeness. Even if scientific data and methods pretend to be of high precision and accuracy, they may not say much about reality. In the light of all these shortcomings, great caution is therefore required when interpreting our results.

In spite of these shortcomings, the project may still have been of great importance. If we did not find many significant treatment effects, the great effort behind the RdE-program and the PISCO investigation was by no means useless. On the contrary, it was very instructive for myself and the whole team: We have made many of the mistakes one can make in a scientific study and discovered some factors that may limit the effectiveness of educational project that aims at raising awareness and changing attitudes and intentions. Moreover, we found significant instrument effects, i.e. the RdE-program in combination with the questionnaires might have triggered a substantial change in the students.

6.2. Side Effects

And don't forget about all the other effects of the program:

The program was a great success in terms of the learning and maturing process that the folk teachers themselves have gone through. The feedback we received from these young people was overwhelming. They feel taken seriously, they learned to take responsibility, to speak in front of a large group of peers and they acquired new contents and pedagogical methods. As folk teacher team members, they feel that they have something to say and that they can make a difference. And they are eager to share what they have learned with their circle of acquaintances and to apply it in their community.

The learning effect for the young NGO and its young employees themselves is also remarkable. Managing such a large project organizationally and logistically is by no means an easy task. All in all, the project can therefore be considered a great success.

6.3. Concluding Remark

Although we did not find most of the expected treatment effects, i.e. significant differences in EnvA and GenA between the treatment groups, we did find a significant effect in one knowledge score and an overall effect of the intervention. If this is interpreted as a causal effect, something substantial was achieved with regards to the students who participated in all three stages of the intervention. If survey exposure made them reflect upon the problems and revise some of their opinions, we can certainly count that as a further success. Moreover, this can also be taken to suggest that with a relatively small intervention a measurable effect can be achieved. The program managers should therefore think carefully about what might have caused such an effect in the surveys and if it could be transferred to the workshops.

For me personally, the project was of inestimable value. When we started, I had no idea of how demanding the realization of such a project would be and how difficult it is to overcome a deep ocean and an even deeper cultural gap. It is hard to imagine a better opportunity to learn so much at all levels about development cooperation and sociological research than the "Red de Educadores Populares'" program and the PISCO project.

Nevertheless, there is still much more to be done.

7. Annex

7.1. Cost-benefit and Cost-effectiveness Analysis

The basic economic question of a cost-benefit and cost-effectiveness analysis (Gertler et al., 2016, 18ff) is the core question of all development projects: What effect have we achieved for every dollar invested? Did we achieve the maximum effect possible or would there have been better alternatives? If we had made a complete cost account of PISCO: How much dollars would PISCO have cost if the entire project had been purchased externally, i.e. if all services and materials had had to be procured at market prices? Did that pay off? Is the effect worth the effort and investment? Would there have been cheaper alternatives to achieve the same or better effects? How high is the cool return on investment without the hot idealism? We can not answer all these question, however important they are. Instead, we create a simple cost calculation and calculate the costs per treated person.

We had a budget of 6000 USD for the evaluation of the RdE-project sponsored by the Institute of Sociology of the University of Bern, Switzerland. This amount was used for flights, surveys development, printing of questionnaires and realization of the survey (transports, catering etc.). The costs for the implementation, on the other hand, were kept comparatively low due to the voluntary work of the folk teachers. The salaries of folk teachers at market prices are not included in the calculation. If we were to include these costs based on normal teachers' salaries ($\approx$ 600 USD per month), we would have to add another 15,000 to 18,000 USD for trainings and workshops. Additionally, the NGO spent about 8000 USD on the implementation and another 2000 USD for internal administration costs like planning, trainer salaries, transports, catering, location rent, administration, organization, accounting etc. (IDB, 2018)

Considering that we did find only small effects, there is no need for a complicated cost and efficiency calculation to find financial key indicators like "USD per standard deviation" etc. Instead, we present a simple calculation of costs per unit treated in table 36, p. 64.[49]

Table 36: Costs and Efficiency Calculation

Costs p. unit	Amount	Details
Total Evaluation Costs	6,000.00 USD	
Total Implementation Costs (IC)	10,000.00 USD	
IC per Class:	102.20 USD	98 selected classes
IC per Potential Student:	4.50 USD	2242 Students of Wave 1
IC per Complier:	6.50 USD	1539 participating students

---

[49] For best practices of cost-benefit and cost-effectiveness analysis, c.f. Dhaliwal et al. (2013) or the informative website of the Inter-American Development Bank IDB:https://www.iadb.org/en/topics/development-effectiveness/evaluation-hub/cost-benefit-cost-effectiveness,17864.html; last date accessed: 18.10.2018

7.2. List of Tables, Figures and References

List of Tables

List of Figures

References

Acemoglu, D. & Robinson, J. A. (2012). *Why nations fail : the origins of power, prosperity and poverty.* New York: Crown Publishers.

Ajzen, I. & Fishbein, M. (2005). *The influence of attitudes on behavior*, (pp. 73–221). Erlbaum: Mahwan N.J.

Albarracín, D., Johnson, B. T., & UZanna, M. P. (2005). *The handbook of attitudes.* Mahwan N.J.: Erlbaum.

Banerjee, A. V. & Duflo, E. (2011). *Poor economics.* New York: Public Affairs.

Campante, F. R. & Chor, D. (2012). Schooling, political participation, and the economy. *Review of Economics and Statistics*, 94(4), 841–859.

Chzhen, Y. (2013). Education and democratisation: Tolerance of diversity, political engagement, and understanding of democracy. *Background paper prepared for the Education for All Global Monitoring Report.*

Coutts, E. & Jann, B. (2011). Sensitive questions in online surveys: Experimental results for the randomized response technique (rrt) and the unmatched count technique (uct). *Sociological Methods Research*, 40(1), 169–193.

Dannecker, P. & Englert, B. (2014). *Qualitative Methoden in der Entwicklungsforschung.* Gesellschaft, Entwicklung, Politik. Wien: Mandelbaum.

DeMaio, T. J. (1984). Social desirability and survey. *Surveying subjective phenomena*, 2, 257–282.

Denzin, N. K. & Lincoln, Y. S. (2005). *The Sage handbook of qualitative research.* London: Sage Publications Ltd.

Dhaliwal, I., Duflo, E., Glennerster, R., & Tulloch, C. (2013). Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education. *Education Policy in Developing Countries*, (pp. 285–338).

Diekmann, A. (2014). *Empirische Sozialforschung, Grundlagen, Methoden, Anwendungen.* Reinbek bei Hamburg: Rowolth, 8th edition.

Diekmann, A. & Preisendörfer, P. (2003). Green and greenback the behavioral effects of environmental attitudes in low-cost and high-cost situations. *Rationality and Society*, 15(4), 441–472.

Droitcour, J., Caspar, R., Hubbard, M., & Ezzati, T. (1991). Measurement errors in surveys (eds. pp biemer, rm groves, le lyberg, na mathiowetz, and s. sudman), chap. the item count technique as a method of indirect questioning: A review of its development and a case study application, 185–210.

Dunlap, R. E. (2017). *A brief history of sociological research on environmental concern.* Routledge: New York.

Dunlap, R. E. & York, R. (2008). The globalization of environmental concern and the limits of the postmaterialist values explanation : Evidence from four multinational surveys. *The Sociological Quarterly*, 49(3), 529–563.

Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes.* Orlando: Harcourt Brace Jovanovich College Publishers.

Evans, D. K. & Popova, A. (2016). What really works to improve learning in developing countries? an analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, 31(2), 242–270.

Ferguson, M. J. & Bargh, J. A. (2007). Beyond the attitude object: Automatic attitudes spring from object-centered-contexts. *Implicit measures of attitudes*, (pp. 216–246).

FISDL (2005). *Mapa de pobreza : política social y focalización*, volume 1. San Salvador: Gobierno de El Salvador.

Flick, U. (2017). *Qualitative Sozialforschung : eine Einführung.* Rowohlts Enzyklopädie. Reinbek bei Hamburg: Rowolth, 8th, completely revised new edition.

Gertler, P., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, Christel, M. J. (2016). *Impact evaluation in practice.* Washington, D.C.: The World Bank, 2nd. revised edition.

Glennerster, R. & Takavarasha, K. (2013). *Running randomized evaluations : a practical guide.* Princeton: Princeton University Press.

Glewwe, P. & Muralidharan, K. (2015). Improving school education outcomes in developing countries : evidence, knowledge gaps, and policy implications. *University of Oxford, Research on Improving Systems of Education (RISE)*.

Gneezy, U. & List, J. (2013). *The why axis: hidden motives and the undiscovered economics of everyday life.* New York: PublicAffairs.

Groves, R. M. (2009). *Survey methodology.* Wiley series in survey methodology. Hoboken, N.J.: Wiley, 2nd edition.

IDB, I. D. B. (2018). Cost-benefit cost-effectiveness. https://www.iadb.org/en/topics/development-effectiveness/evaluation-hub/cost-benefit-cost-effectiveness,17864.html. Last date accessed: 18.10.2018.

Inglehart, R. (1990). *Cultural shift in advanced industrial societies.* Princeton N.J.: Princeton University Press.

Inglehart, R. (1995). Public support for environmental protection : Objective problems and subjective values in 43 societies. *PS: Political Science Politics*, 28(1), 57–72.

Ito, T. A. & Cacioppo, J. T. (2007). *Attitudes as Mental and Neural States of Readiness: Using Physiological Measures to Study Implicit Attitudes*, (pp. 125–158). Guilford Press: New York.

Jakob, L. (2018). Pay or die? – prison or cemetery? a story about the maras – the mafias of the poor in el salvador. `https://edinuniv.maps.arcgis.com/apps/Cascade/index.html?appid=f9d4ebbf7667439dbe1ac292e23203ac`.

Jakob, M. S. (2016). Educational decision-making in rural el salvador. `https://consciente.ch/wordpress_b/wp-content/uploads/2018/01/Higher_Education_Morazan.pdf`.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality Quantity*, 47(4), 2025–2047.

Kößler, R. (2016). *Modernisierungstheorien*, (pp. 27–40). Springer VS: Wiesbaden.

Maloney, M. P. & Ward, M. P. (1973). Ecology : Let's hear from the people : An objective scale for the measurement of ecological attitudes and knowledge. *American psychologist*, 28(7), 583.

Maloney, M. P., Ward, M. P., & Braucht, G. N. (1975). A revised scale for the measurement of ecological attitudes and knowledge. *American psychologist*, 30(7), 787.

Milfont, T. L. & Duckitt, J. (2010). The environmental attitudes inventory : A valid and reliable measure to assess the structure of environmental attitudes. *Journal of environmental psychology*, 30(1), 80–94.

Muralidharan, K. (2017). Field experiments in education in developing countries. *Handbook of Economic Field Experiments*, 2, 323–385.

Olson, J. M., Vernon, P. A., Harris, J. A., & Jang, K. L. (2001). The heritability of attitudes: a study of twins. *Journal of personality and social psychology*, 80(6), 845.

Ostrom, T. M. (1989). Interdependence of attitude theory and measurement. *Attitude structure and function*, (pp. 11–36).

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy,*, 6(1), 65–78.

Putnam, R. D. (2000). *Bowling alone: the collapse and revival of American community*. New York: Simon Schuster.

Roland, G. (2014). *Development economics*. The Pearson series in economics. Boston: Pearson.

Sen, A. (1999). *Development as freedom*. New York: Knopf, 2nd printed edition.

Slezak, G. (2014). *Von Wahrnehmung und Erfahrung*, book section 267, (pp. 267 S.). Mandelbaum: Wien.

Solon, G. (1999). *Intergenerational mobility in the labor market*, volume 3, (pp. 1761–1800). Elsevier: Amsterdam.

The World Bank (2011). *World development report 2012: gender equality and development*. Washington: World Bank Publications.

The World Bank (2018). *Learning to realize education's promise*. World development report. Washington D.C.: The World Bank.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambrigde, UK: Cambridge University Press.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309), 63–69.

Wittenbrink, B. & Schwarz, N. (2007). *Implicit measures of attitudes*. New York: Guilford Press.

Wright, E. O. (2010). *Envisioning real utopias*. London / New York: Verso.

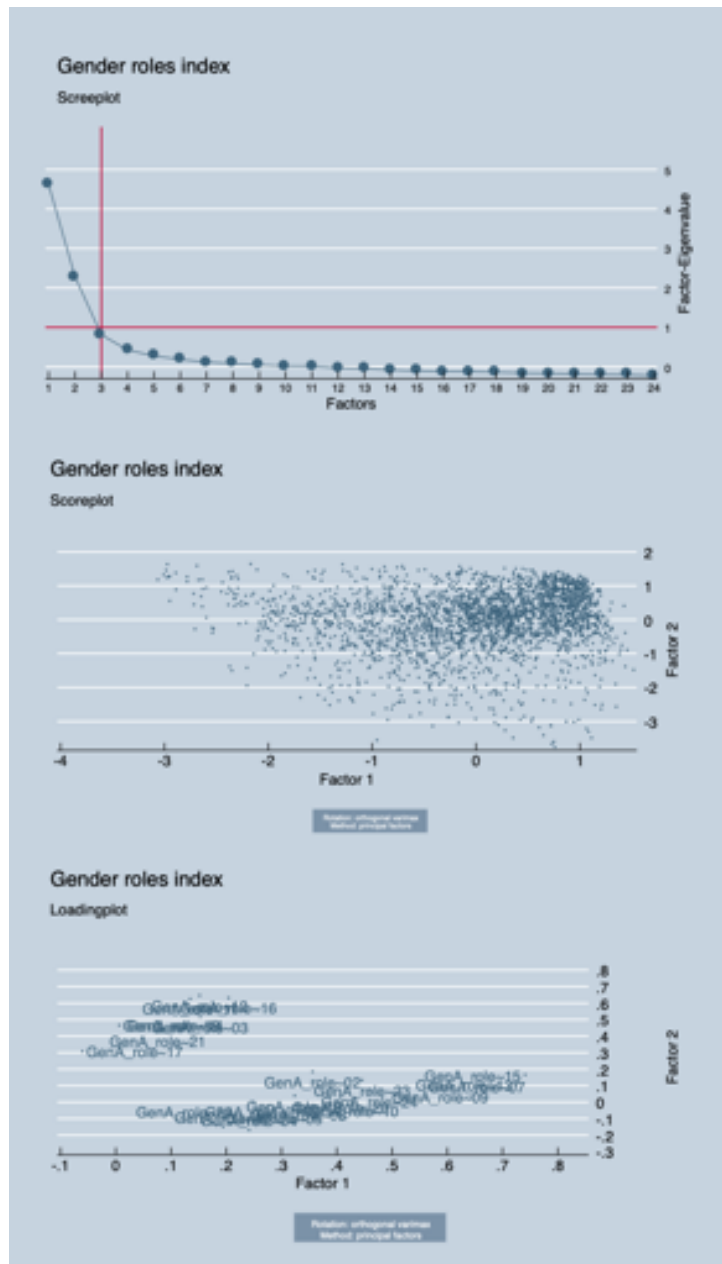## 7.3. Supplementary Tables, Figures and Plots

Figure 15: Factor Analysis Plot for Classic Gender Roles Index

Table 37: T-Test of all Behavior and Attitudes Variables

	Coef	Std. Err.	Pr(|T| > |t|)	Pr(T > t)	Obs.
EnvA_worry_01	-0.0508	0.0667	0.2234	0.4468	1364
EnvA_worry_02	0.0161	0.0660	0.5964	0.8072	1357
EnvA_worry_03	-0.1010	0.0877	0.1250	0.2501	1214
EnvA_worry_04	-0.0156	0.0728	0.4152	0.8305	1289
EnvA_worry_05	0.0612	0.0809	0.7753	0.4494	1318
EnvA_worry_06	-0.0445	0.0618	0.2358	0.4716	1350
EnvA_worry_07	-0.0592	0.0678	0.1915	0.3831	1319
EnvA_CC_01	-0.1581	0.0814	0.0262	0.0523	1250
EnvA_CC_02	0.2491**	0.0862	0.9980**	0.0039**	1403
EnvA_CC_03	0.0283	0.0653	0.6675	0.6650	1418
EnvA_CC_04	0.0883	0.0640	0.9160	0.1680	1424
EnvA_CC_05	-0.0508	0.0625	0.2082	0.4164	1367
EnvA_CC_06	0.0320	0.0871	0.6434	0.7132	1216
EnvA_CC_07	0.0840	0.0886	0.8285	0.3430	1350
EnvA_CC_08	0.1177	0.0794	0.9309	0.1383	1335
EnvA_CC_09	0.0341	0.0869	0.6524	0.6951	1360
EnvA_CC_10	0.0770	0.0915	0.7998	0.4003	1311
EnvA_Defor	-0.0169	0.0597	0.3888	0.7776	1466
EnvA_Change_01	-0.1865*	0.0802	0.0101*	0.0202*	1479
EnvA_Change_02	-0.0716	0.0632	0.1287	0.2574	1469
EnvA_Change_03	0.0738	0.0816	0.8171	0.3658	1468
EnvA_Change_04	-0.0011	0.0763	0.4945	0.9889	1477
EnvA_Change_05	0.0483	0.0730	0.7458	0.5084	1465
EnvA_Change_06	0.0402	0.0848	0.6824	0.6352	1452
EnvA_Change_07	0.0308	0.0830	0.6445	0.7110	1460
EnvA_Change_08	0.0643	0.0849	0.7753	0.4494	1467
EnvA_Change_09	0.1723*	0.0791	0.9852*	0.0295*	1459
EnvA_Change_10	-0.0511	0.0731	0.2425	0.4850	1448
EnvA_Change_11	-0.0142	0.0571	0.4016	0.8032	1483
EnvA_Change_12	0.0463	0.0723	0.7389	0.5223	1427
EnvA_Change_13	0.1282	0.0670	0.9720	0.0561	1425
EnvA_Change_14	0.0261	0.0545	0.6842	0.6317	1425
EnvA_Change_15	0.0249	0.0762	0.6280	0.7440	1425
EnvA_Change_16	-0.0337	0.0421	0.2115	0.4229	1424
EnvA_Change_17	-0.0443	0.0602	0.2310	0.4620	1426

Table 37: T-Test of all Behavior and Attitudes Variables

EnvA_Change_18	-0.0467	0.0669	0.2428	0.4856	1427
EnvA_Change_19	-0.0439	0.0464	0.1721	0.3442	1426
EnvA_Change_20	0.1102	0.0658	0.9529	0.0942	1426
EnvA_aff_01	0.0506	0.0573	0.8113	0.3775	1513
EnvA_aff_02	0.0972	0.0647	0.9333	0.1334	1494
EnvA_aff_03	0.0788	0.0522	0.9343	0.1315	1497
EnvA_aff_04	0.0058	0.0488	0.5475	0.9051	1502
EnvA_aff_05	0.0346	0.0489	0.7604	0.4792	1504
EnvA_aff_06	0.0237	0.0477	0.6904	0.6191	1502
EnvA_aff_07	0.0497	0.0515	0.8327	0.3346	1513
EnvB_consum_01	-0.0520	0.0346	0.0662	0.1324	1491
EnvB_consum_02	0.0404	0.0519	0.7820	0.4360	1476
EnvB_consum_03	0.0256	0.0492	0.6990	0.6021	1472
EnvB_consum_04	-0.0424	0.0540	0.2163	0.4325	1474
EnvB_consum_05	0.0318	0.0601	0.7013	0.5975	1475
EnvB_vegetarian	0.1056*	0.0453	0.9901*	0.0198*	1404
EnvB_behavior_01	-0.0353	0.0760	0.3213	0.6425	1475
EnvB_behavior_02	-0.1067	0.0718	0.0685	0.1371	1474
EnvB_behavior_03	-0.0556	0.0788	0.2401	0.4802	1473
EnvB_behavior_04	0.0225	0.0760	0.6165	0.7670	1464
EnvB_behavior_05	-0.0086	0.0721	0.4527	0.9055	1468
EnvB_behavior_06	-0.1497*	0.0740	0.0217*	0.0433*	1475
EnvB_behavior_07	-0.1793**	0.0685	0.0045**	0.0089**	1475
EnvB_behavior_08	-0.0486	0.0757	0.2606	0.5212	1463
EnvB_energy	0.0017	0.0431	0.5161	0.9677	1491
EnvB_01	0.0927	0.0678	0.9143	0.1714	1435
EnvB_02	0.0965	0.0730	0.9068	0.1865	1425
EnvB_03	0.1356	0.0711	0.9716	0.0568	1425
EnvB_04	0.1007	0.0913	0.8649	0.2702	1429
EnvB_05	-0.0047	0.0523	0.4645	0.9291	1348
EnvB_06	0.0656	0.0588	0.8677	0.2647	1342
EnvB_07	-0.0276	0.0469	0.2778	0.5556	1336
EnvB_08	-0.0436	0.0545	0.2120	0.4239	1505
EnvB_09	-0.0187	0.0267	0.2424	0.4848	1491
EnvB_10	0.0092	0.0370	0.5979	0.8041	1496
EnvB_11	0.0186	0.0385	0.6858	0.6284	1484

Table 37: T-Test of all Behavior and Attitudes Variables

EnvB_12	0.0128	0.0349	0.6431	0.7137	1498
EnvB_13	0.0228	0.0325	0.7581	0.4838	1506
EnvB_disc_01	-0.0205	0.0460	0.3282	0.6565	1504
EnvB_disc_02	-0.0062	0.0496	0.4501	0.9001	1491
EnvB_disc_03	0.0019	0.0516	0.5145	0.9710	1462
EnvB_disc_04	-0.0577	0.0481	0.1154	0.2307	1451
GenA_earnings	-0.1625**	0.0499	0.0006**	0.0011**	1281
GenA_politics	-0.1195*	0.0555	0.0158*	0.0316*	1149
GenA_fortune	-0.1088*	0.0520	0.0184*	0.0367*	1166
GenA_roles_01	0.0376	0.0780	0.6852	0.6295	1452
GenA_roles_02	-0.0448	0.0717	0.2662	0.5324	1441
GenA_roles_03	0.0341	0.0786	0.6678	0.6645	1458
GenA_roles_04	-0.0057	0.0879	0.4741	0.9482	1438
GenA_roles_05	-0.1832*	0.0922	0.0236*	0.0472*	1436
GenA_roles_06	-0.1651	0.0915	0.0357	0.0715	1406
GenA_roles_07	0.0295	0.0772	0.6487	0.7026	1442
GenA_roles_08	0.0453	0.0935	0.6860	0.6279	1437
GenA_roles_09	0.0028	0.0849	0.5130	0.9740	1429
GenA_roles_10	-0.0203	0.0877	0.4085	0.8170	1424
GenA_roles_11	-0.0012	0.0722	0.4935	0.9869	1449
GenA_roles_12	-0.0385	0.0650	0.2766	0.5532	1445
GenA_roles_13	0.0708	0.0763	0.8233	0.3534	1444
GenA_roles_14	0.0795	0.0685	0.8769	0.2462	1433
GenA_roles_15	-0.0102	0.0784	0.4485	0.8969	1444
GenA_roles_16	0.1006	0.0732	0.9152	0.1696	1466
GenA_roles_17	0.0679	0.0902	0.7741	0.4519	1404
GenA_roles_18	0.0389	0.0796	0.6874	0.6253	1412
GenA_roles_19	0.0912	0.0873	0.8520	0.2960	1441
GenA_roles_20	-0.0687	0.0849	0.2094	0.4187	1436
GenA_roles_21	-0.0385	0.0831	0.3217	0.6435	1439
GenA_roles_22	-0.1083	0.0878	0.1088	0.2175	1419
GenA_roles_23	0.0053	0.0798	0.5267	0.9466	1432
GenA_roles_24	0.1117	0.0853	0.9046	0.1908	1425
GenA_dichos_01	0.1074	0.0574	0.9692	0.0617	1486
GenA_dichos_02	0.1077	0.0676	0.9445	0.1111	1480
GenA_dichos_03	0.1054	0.0614	0.9570	0.0860	1460

Table 37: T-Test of all Behavior and Attitudes Variables

GenA_dichos_04	0.0855	0.0677	0.8967	0.2066	1453
GenA_dichos_05	0.0039	0.0648	0.5238	0.9524	1470
GenA_violence_01	0.0837	0.0681	0.8905	0.2190	1496
GenA_violence_02	-0.0118	0.0678	0.4308	0.8616	1479
GenA_violence_03	0.0192	0.0815	0.5932	0.8136	1454
GenA_violence_04	0.0631	0.0919	0.7537	0.4926	1469
GenA_violence_05	0.0451	0.0872	0.6976	0.6048	1474
GenA_violence_06	0.0022	0.0816	0.5106	0.9788	1482
GenA_violence_07	-0.0178	0.0722	0.4025	0.8050	1474
GenA_violence_08	0.0053	0.0751	0.5283	0.9435	1437
GenA_roles_25	-0.1414	0.0834	0.0451	0.0902	1424
GenA_roles_26	-0.1369	0.0786	0.0410	0.0819	1423
GenA_roles_27	0.0501	0.0813	0.7310	0.5379	1405
GenA_roles_28	-0.0380	0.0787	0.3146	0.6291	1413
GenA_roles_29	-0.0456	0.0921	0.3102	0.6204	1423
GenA_roles_30	-0.0521	0.0883	0.2774	0.5548	1418
GenA_roles_31	0.0812	0.0862	0.8268	0.3465	1416
GenA_roles_32	0.0791	0.0846	0.8250	0.3500	1419
GenA_roles_33	-0.0577	0.0743	0.2187	0.4374	1411
GenA_roles_34	-0.0378	0.0827	0.3239	0.6477	1414
GenA_home_ideal	-0.0787	0.0786	0.1585	0.3170	1447
GenA_home_real	-0.1448	0.0897	0.0533	0.1066	1443
GenB_disc_w_01	0.1951*	0.0956	0.9792*	0.0416*	805
GenB_disc_w_02	-0.0345	0.0950	0.3583	0.7167	780
GenB_disc_w_03	0.1630	0.0940	0.9583	0.0833	781
GenA_selfie_w_01	0.0623	0.0677	0.8213	0.3574	807
GenA_selfie_w_02	0.0616	0.1007	0.7294	0.5411	805
GenA_selfie_w_03	0.0654	0.0701	0.8244	0.3512	794
GenA_selfie_w_04	-0.1019	0.1057	0.1676	0.3352	809
GenA_selfie_w_05	0.0928	0.1014	0.8198	0.3605	811
GenA_selfie_w_06	-0.0581	0.1179	0.3111	0.6222	804
GenA_selfie_w_07	0.0729	0.0885	0.7948	0.4103	799
GenA_selfie_w_08	0.0780	0.0437	0.9627	0.0746	786
GenA_selfie_w_09	-0.0093	0.0549	0.4329	0.8659	787
GenA_selfie_w_10	0.0388	0.1183	0.6284	0.7432	798
Gen_affect_w_01	-0.0490	0.0907	0.2948	0.5895	821

Table 37: T-Test of all Behavior and Attitudes Variables

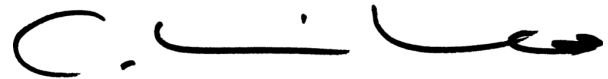|					
Gen_affect_w_02	-0.0498	0.1032	0.3146	0.6293	812
Gen_affect_w_03	-0.0335	0.1120	0.3826	0.7651	800
Gen_affect_w_04	0.0592	0.1039	0.7156	0.5688	807
Gen_affect_w_05	0.0328	0.0983	0.6308	0.7385	809
Gen_affect_w_06	0.1219	0.0952	0.8995	0.2009	813
Gen_affect_w_07	0.0749	0.0982	0.7769	0.4462	807
Gen_affect_w_08	-0.5250	0.5891	0.2019	0.4037	13
GenB_disc_m_01	-0.1374	0.1019	0.0890	0.1780	652
GenB_disc_m_02	-0.1359	0.0962	0.0791	0.1583	619
GenB_disc_m_03	-0.0917	0.1053	0.1921	0.3842	623
GenA_selfie_m_01	0.0413	0.1060	0.6514	0.6971	649
GenA_selfie_m_02	0.1264	0.1163	0.8611	0.2777	646
GenA_selfie_m_03	-0.0306	0.1050	0.3856	0.7711	640
GenA_selfie_m_04	-0.0012	0.1197	0.4961	0.9922	633
GenA_selfie_m_05	0.1026	0.1203	0.8030	0.3940	637
GenA_selfie_m_06	0.1236	0.1197	0.8490	0.3020	634
GenA_selfie_m_07	0.0534	0.1071	0.6909	0.6181	637
GenA_selfie_m_08	-0.0146	0.1109	0.4477	0.8954	640
GenA_selfie_m_09	-0.0718	0.1328	0.2943	0.5887	644
GenA_selfie_m_10	0.1394	0.1200	0.8772	0.2457	652
GenA_selfie_m_11	0.0737	0.1047	0.7591	0.4817	646
GenA_selfie_m_12	0.0683	0.1235	0.7097	0.5806	648
Gen_affected_m_01	-0.0242	0.0882	0.3922	0.7843	646
Gen_affected_m_02	0.0113	0.1028	0.5436	0.9128	647
Gen_affected_m_03	0.0526	0.1079	0.6870	0.6260	644
Gen_affected_m_04	0.0478	0.1038	0.6774	0.6451	646
Gen_affected_m_05	-0.0362	0.1032	0.3630	0.7261	648
Gen_affected_m_06	1.0000	0.6831	0.9062	0.1875	9

* p<0.05, ** p<0.01, *** p<0.001

7.4. Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet.

Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, January 25, 2019