
LONG-TERM OUTCOMES OF TEGV AFTER-SCHOOL PROGRAMS

Findings from a Matched Comparison Group Evaluation
Design

August 2017

Submitted by:

Yasemin Kisbu-Sakarya, PhD
E-mail: ykisbu@ku.edu.tr
www.evalresearchlab.com

Submitted to:

TEGV - Educational Volunteers Foundation of Turkey
Acıbadem Cd. Rauf Paşa Hanı Sk. No:42
34660 Acıbadem, Istanbul, Turkey
Phone: +90 (216) 290 70 00



Acknowledgements

We extend sincere thanks to the many people who contributed to the preparation of this report. In particular, we are indebted to TEGV education program teams and their leaders. We thank the evaluation and data management teams at TEGV for providing helpful information and feedback throughout this evaluation study. We would like to thank former members of TEGV, Suat Kardaş, Gülşah Gürkan, and Gizem Gündüz for the initiation of this evaluation project and for their great effort and help. We would like to thank Mete Meleksoy for his sincere support throughout the project.

Executive Summary

Introduction

Children spend a large proportion of their waking hours out of school. Especially for children living in disadvantaged neighborhoods, the way they spend this time puts them at a greater risk in terms of lower academic achievement and attainment, poor cognitive, social, and emotional development, and risky choices such as delinquency and drug use. **In light of research showing the positive consequences of supervised and organized after-school activities for children, Educational Volunteers Foundation of Turkey's (TEGV) main activities have focused on the implementation of after-school programs for children in poverty as a prominent strategy to promote positive child and youth development in Turkey.**

Review studies indicate that children participating in after-school programs achieve higher academic and developmental gains compared to children who do not participate in the programs, but mostly focus on short-term effects. The goal of this study is to assess the long-term benefits of TEGV programs following a pilot survey study conducted in 2012 in order to investigate the status of TEGV alumni. A total of 92% of surveyed TEGV alumni in that study stated they have graduated from high school, 78% of them claimed that TEGV had a significant influence on their decision to continue school, 77% of them reported that they have a strong sense of self-esteem, and 78% of them reported they have good empathy and communication skills. The current study aims to take the alumni survey one step further and conduct an impact evaluation of TEGV after-school programs to test its effectiveness in the long term, specifically 10-12 years after program entry. **Using a matched comparison group evaluation design, this study assesses whether educational and psychosocial outcomes differ between individuals who received the TEGV after-school programs during their childhood and similar individuals who did not participate in the TEGV program.**

TEGV after-school programs

TEGV designs child-centered learning programs for school-age children. Collectively, TEGV programs aim to support psychomotor/physical, social, emotional, and cognitive development of children through active learning. The evaluated education programs are gathered together under four main topics: math, science, reading, and arts. The contents of the programs vary according to level and grade, lasting for 8-10 weeks.

Methodology

Two strategies were used in order to achieve **a comparison group who would mimic the pre-intervention characteristics of TEGV participants. The first strategy was the use of purposive sampling. The second strategy was the use of propensity score matching analysis.**

TEGV participants: Data was collected from 108 23- to 25-year-old individuals who enrolled in the TEGV after-school program at the age of 12-14 for the first time and participated in the programs between 8-96 weeks.

Non-participants: Data was collected from 1212 23- to 25-year-old individuals living in the same cities and neighborhoods that are socioeconomically similar to the neighborhoods where TEGV programs were given to TEGV participants.

Matched sample: We have matched a TEGV participant up to two non-participants using their pre-intervention characteristics. **Matching yielded a sample size of 216 individuals (80 TEGV participants, 136 non-participants) from the following 13 cities: Istanbul, İzmir, Ankara, Antalya, Batman, Denizli, Diyarbakır, Eskişehir, Gaziantep, Kahramanmaraş, Samsun, Sivas, and Van.**

All outcome analyses were conducted using the matched sample.

Findings and conclusions

Our evidence suggests that TEGV programs had a statistically significant effect on academic outcomes and interpersonal and intrapersonal competencies in the long-term. **TEGV participants had higher years of schooling, life satisfaction, self-concept, interpersonal skills, frequency of reading books, and more positive attitudes toward education, compared to non-participants.** Results also indicate that TEGV programs had especially a **boosting effect on female participants' outcomes,** compared to males.

This evaluation study offers empirical support for the impact of after school programs on academic and psychological development of young adults. We hope our findings will stimulate more interest in after-school practices in Turkey.

Outcome	Treatment Effect Size	
	Female	Male
Education level	0.52*	0.38†
High School GPA	0.14	0.05
Attitudes towards continuing education	0.79***	0.79***
Frequency of reading books	0.43*	0.22
Income	0.15	0.04
Life satisfaction	0.44*	0.47*
Interpersonal competence	0.55**	0.39†
Self-esteem	0.60**	0.51*
Self-efficacy	0.82***	0.19

* Significantly different from zero at the .05 level.

** Significantly different from zero at the .01 level.

*** Significantly different from zero at the .001 level.

† Marginally significantly different from zero at the .10 level

Contents

ACKNOWLEDGEMENTS

SUMMARY

INTRODUCTION

1

Purpose and focus of the impact evaluation

Related studies on outcomes of after-school programs

What works?

 Characteristics of effective after-school programs

Who benefits the most?

TEGV after-school programs

Key evaluation questions

METHODOLOGY AND DATA

11

Study design and analytic approach

Characteristics of study participants

Outcome measures

FINDINGS

17

Comparing outcomes for participants and nonparticipants

 Academic outcomes, employment, and life satisfaction effect estimates

 Interpersonal and intrapersonal competency, and social responsibility estimates

Subgroup Analyses

UNANSWERED QUESTIONS

23

LIMITATIONS

25

REFERENCES

26

APPENDIX

30

 Covariate balance after matching

Page left intentionally blank for double-sided printing

INTRODUCTION

Purpose and focus of the impact evaluation

Children spend a large proportion of their waking hours out of school. Especially for children living in disadvantaged neighborhoods, the way they spend this time puts them at a greater risk in terms of academic achievement and attainment, cognitive development, social and emotional consequences, problem behaviors, and risky choices, such as delinquency and drug use. In light of research showing the positive consequences of supervised and organized after-school activities for children, Educational Volunteers Foundation of Turkey's (TEGV) main activities have focused on the implementation of after-school programs for children in poverty as a prominent strategy to promote positive child and youth development in Turkey.

TEGV after-school programs strive to improve the physical, social, emotional, and cognitive development of children using a child-centered learning approach. To date, TEGV has worked with thousands of volunteers to serve over two million children in its activity centers located in disadvantaged neighborhoods across the country. Guided by an evidence-based approach, all implemented TEGV programs undergo an evaluation phase to assess if the targeted short-term outcomes are achieved. The goal of this study is to assess the long-term benefits of TEGV programs following a pilot survey study conducted in 2012 in order to investigate the status of TEGV alumni in the long term. A total of 92% of surveyed TEGV alumni in that study stated they have graduated from high school, 78% of them claimed that TEGV had a significant influence on their decision to continue school, 77% of them reported that they have a strong sense of self-esteem, and 78% of them reported they have good empathy and communication skills. The current study aims to take the alumni survey one step further and conduct an impact evaluation of TEGV after-school programs to test its effectiveness in the long term, specifically 10-12 years after program entry. **Using a matched comparison group evaluation design, this study assesses whether educational and psycho-social outcomes differ between individuals who received the TEGV after-school programs during their childhood and similar individuals who did not participate in the TEGV program.**

A second aim of this evaluation study is to build the capacity of the field to learn from this research and use this data to further strengthen after school practices in Turkey.

Despite the educational reform efforts in Turkey, there is still major concern regarding how to make children stay in school, acquire necessary knowledge and skills, have fewer problem behaviors and better psycho-social adjustment, successfully join the labor force, and become engaged members of the society. Well-implemented programs and policies are recognized as the main tools to create a positive impact on children's academic, social, and other outcomes. However, the key to program success is the dissemination of well-implemented programs that are shown to be beneficial for the program participants and eventually for society at large. Not all evaluation studies show that implemented after-school programs are effective in improving the targeted outcomes. This has led to a useful discussion on the responsible allocation of resources when deciding which social and educational programs to deliver. As similar efforts will lead to emerging knowledge about what works on improving child and youth outcomes, the highest quality after-school programming should achieve a positive impact on a larger scale.

Related studies on outcomes of after-school programs

Existing studies on after-school programs is mostly focused on programs' short-term effects on the cognitive and social development of children. Review studies indicate that children participating in after-school programs achieve higher academic and developmental gains compared to children who do not participate in the programs (Durlak and Weissberg 2007; Lauer et al. 2006; Redd et al. 2002; Granger 2008). Several studies found that attending after-school programs has led to improved outcomes related to academic performance (Miller, 2003), such as better grades (Baker & Witt, 1996; Brooks, 1995; Cardenas, 1992; Carlisi, 1996; Hamilton & Klein, 1998; Hamilton, Le & Klein, 1999; Schinke et al., 1992), higher homework completion rates (Carlisi, 1996; Johnson et al., 1999), improved achievement test results (Hamilton & Klein, 1998; Hamilton et al., 1999; Huang, 2001; Huang et al., 2000; Johnson et al., 1999), reduced drop-out (Jones & Offord, 1989), and lower grade retention (Hamilton et al., 1999). In addition to better academic performance, studies also showed that students gain

attitudes, skills, and behaviors that would benefit school success. For example, Vandell and Pierce (1999) examined children who lived in low-income, high-crime neighborhoods and found that higher levels of attendance in formal after-school programs led to improved work habits. Other studies showed that attending after-school programs led to positive behavior in school (Baker & Gribbons, 1998; Johnson et al., 1999; Posner & Vandell, 1994), positive attitudes toward school (Brooks, 1995; Huang et al., 2000; Schinke et al., 1998; Schlegel, 2003), and reduced absenteeism from school (Huang, Gribbons, Kim, Lee, & Baker, 2000; Johnson et al., 1999; Schinke, Cole & Poulin, 1998; Vandell & Pierce, 1999).

With regard to intra- and interpersonal gains, students attending after-school programs had increased self-efficacy (Campbell et al., 1995), better emotional adjustment (Marshall et al., 1997), improved conflict resolution skills (Carlisi, 1996; Posner & Vandell, 1994; Rodriguez et al., 1999; Vandell & Pierce, 1997; Warren et al., 2002), and greater feelings of belonging in the program or community (Schlegel, 2003). Some studies also investigated the effects of after-school programs on reducing negative behaviors. Jones and Offord (1989), who examined 5- to 15-year-old children attending a skill-development program in a public housing complex in Canada found that children attending the program had significantly reduced rates of antisocial behavior outside home and school.

What works?

Several systematic reviews and meta-analysis¹ studies point out the positive results of after-school programs (Durlak & Weissberg, 2010; Lauer et al., 2006). Yet there are also studies that did not find significant gains for children in after-school care. A meta-analysis study conducted by Zief and Lauver (2006) reports that out of five impact evaluation studies, only one showed a significant difference between children in the after-school program and control groups in terms of academic, social, and emotional outcomes. **While these mixed results about the effectiveness of after-school programs can be explained by the study inclusion criteria**

¹ A meta-analysis study, by using a statistical approach, combines the results from several selected studies that investigated a common outcome and develops a single conclusion with greater power.

of these review papers, they also point out the variation in effects due to program characteristics and quality.

An extensive review of after-school programs that are successful in influencing academic performance; social, emotional, health, and well-being outcomes; and prevention of negative behaviors, such as reduction in juvenile crime, violent behaviors, and drug use, suggests that critical factors for a successful after-school program include having appropriate structure and supervision, well-prepared staff, intentional programming, and strong partnerships with families, other community institutions, and schools (Little et al., 2008). These factors contribute to access to and sustained participation in the program. Higher implementation quality on program management and climate is also found to be associated with positive experiences of middle school students who attended an enhanced after-school program with prevention modules (Cross et al., 2010).

There is further quantitative evidence showing the association between program characteristics and positive child and youth outcomes. A meta-analysis by Beckett et al. (2001) shows that three program characteristics are significantly associated with improved outcomes: variety of activities, flexibility of programming, and emotional climate. A following meta-analysis study investigates the specific links between program quality factors, mathematics and reading test scores (Leos-Urbel, 2015). Results show that the presence of a supportive environment in an after-school program is significantly associated with higher scores on standardized mathematics and reading test scores in grades 4-5; additionally, the presence of a supportive environment and structured interactions are significantly associated with higher reading scores in grades 6-8.

In 2007, Durlak and Weissberg wanted to investigate the effects of after-school programs on improving academic, personal, and social skills. They included all evaluation studies of after school programs that operated during the school year and targeted the development of at least one personal or social skill into their meta-analysis study. Even though their meta-analysis results showed significant positive effects on seven of eight outcomes, there was a striking

variation in the effects. Authors categorized the after-school programs into two groups according to program characteristics. They named the after-school programs that focused on specific social and personal skills, employed a sequenced learning approach, and had students actively involved as SAFE (Sequenced, Active, Focused, and Explicit) programs. The remarkable result was that only the SAFE after-school programs had significant positive effects on seven of the eight outcomes, while the non-SAFE programs had no effects.

Table 1. Positive effects of after-school programs (from Durlak & Weissberg, 2007)

Effects	SAFE programs	Other programs
Feelings and attitudes		
Child self-perceptions	✓	ns.
School bonding	✓	ns.
Behavioral adjustment		
Positive social behaviors	✓	ns.
Reduced problem behaviors	✓	ns.
Reduced drug use	✓	ns.
School performance		
Achievement test scores	✓	ns.
School grades	✓	ns.
School attendance	ns.	ns.

✓ denotes a positive significant effect was found; ns denotes no significant effect was found.

Who benefits the most?

Not all students in an after-school program are influenced to the same degree. Previous research shows that program effects are usually higher for those who are in greater need for the program at baseline. Since disadvantaged children do not have much access to programs or opportunities fostering their cognitive, social, and emotional development during their out of school time, an after-school program is more likely to boost their performance and well-being by offering a good quality, structured, supervised, and safe learning environment. Studies show that after-school programs help the students with initially low achievement test scores, high absence rates, and low language skills to close the gap with non-program students (Huang et al., 2000; Posner & Vandell, 1994). One can expect higher program effects for female or immigrant children who have no quality activity choices during out-of-school time as well.

Box 1: Characteristics of effective after-school programs

High quality programming is critical to achieve successful outcomes. Based on an extensive review of successful after-school programs, Miller advocates the following list as the characteristics of high quality programming (directly taken from Miller, 2003):

- Authentic curricula that is age-appropriate and provides engaging, skill-building, hands on activities geared toward the particular goals of the program
- High quality content
- Clear rules and expectations with consistent consequences
- Flexibility that allows participants to choose activities that interest them or choose the approach they use to achieve goals
- Youth valued as resources, including having a voice in determining program content and connecting to larger community
- Low adult-to-youth ratios
- Sufficient well trained and compensated staff
- Staff who work in the program over a long period of time and are able to build relationships with the participants
- Staff who understand the developmental tasks faced by young people and are able to help them accomplish these tasks
- Staff who engage in frequent, positive interactions with participants
- Staff who understand the cultural, racial, ethnic, and class backgrounds of participants and are able to support healthy identity development in a diverse group of young people
- Staff with high expectations of all youth
- A role for participants in creating, implementing and reflecting on the program
- Strong, sustainable administration
- A full-time coordinator
- Clarity of mission and goals
- On-going self-assessment and evaluation
- Adequate funding without constant threat of loss
- Managers who are skilled administrators, inspiring leaders, and connected to community resources
- Support of the school principal (especially if school-based)
- Involvement of parents
- Access to appropriate space for program and storage
- Connections to community partners and infrastructure for programming, training, technical assistance

TEGV after-school programs

TEGV designs child-centered learning programs for children aged 6-16. Children can sign up, at no cost, to participate in TEGV's various after-school programs implemented by volunteers who receive extensive training before they meet the children.

The programs take place in TEGV activity centers located across Turkey. There are three types of activity centers TEGV uses to deliver the programs: education parks, education units and mobile learning units which are called Fireflies. Today, TEGV has a total of 10 education parks, two in Istanbul and one in each of the following cities: Ankara, Antalya, Eskişehir, Gaziantep, Izmir, Samsun, Şanlıurfa, and Van. Education parks are built on spacious grounds allocated by the local authorities, and they have all the spatial and technological capacities to support a versatile education. The parks have 10-20 decares of open space, and 1,200 square meters of indoor space. Each year, about 3,500 children attend activities at the education parks. There are soccer pitches, basketball courts, etc. in the outdoor areas, and the indoor areas have 10 activity rooms, 2 computer rooms, and a library. Education units are on a smaller scale. They are set up in cities and municipalities where education opportunities are limited. The units are established in sites that are allocated by individuals or local authorities. The education units have about 250-300 square meters of indoor space and serve around 700 children each year. The education units have 4 activity rooms, with each designed for a different purpose along with 1 technology room and 1 library.

Collectively, TEGV programs aim to support psychomotor/physical, social, emotional, and cognitive development of children through active learning. The structured programs and elective classes strive to improve children's knowledge and skills in the following nine areas:

- Academic
- Verbal
- Cognitive
- Communication
- Health

- Arts
- Sports
- Technology
- Civic skills

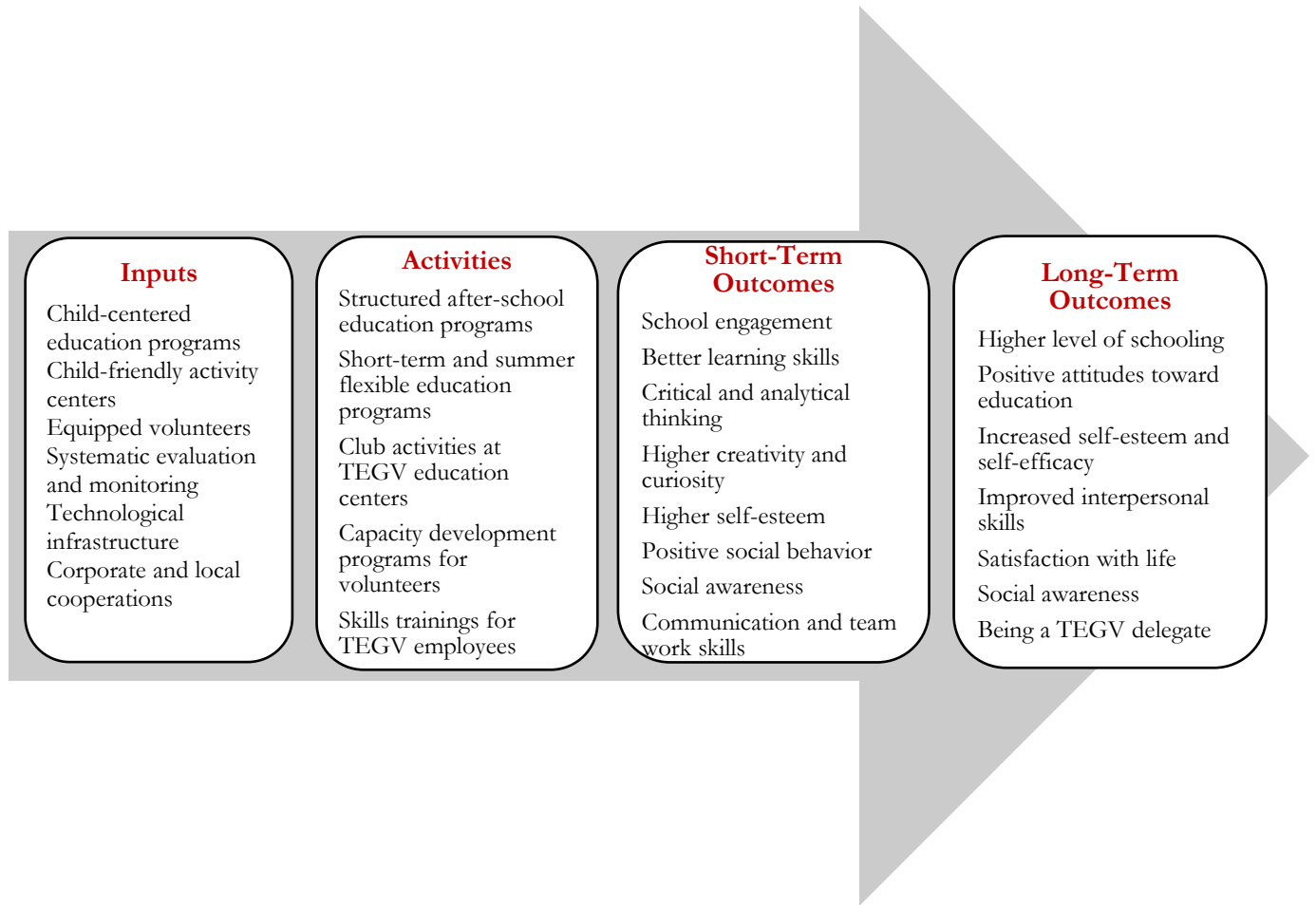


Figure 1. TEGV after-school programs logic model for children outcomes

The structured education programs are gathered together under five main topics: math, science, reading, arts and informatics. The contents of the programs vary according to level and grade, lasting for 8-10 weeks. Children are allowed to enroll in many programs consecutively. The math program contributes to the development of children’s positive attitudes toward

mathematics and reduced mathematics anxiety. Children learn mathematics with thinking practices, station works, cubes, colorful images, examples from daily life, geometric objects, and stories. The science program aims to change children's attitudes toward science in a positive way and improve their knowledge about scientific methods and cause-and-effect relationships. Children conduct lab experiments to acquire basic scientific knowledge and skills. The creative reading program is based on interdisciplinary applications that aim to develop children's creative thinking skills. The arts program improves children's life skills, such as creativeness, teamwork, problem solving, self-esteem, communication, and taking responsibility. It includes 2 or 3 dimensional techniques, like painting, sculpture, print art, and waste material utilization. Children also learn about famous artists. The informatics program aims to help children build 21st Century digital skills, focusing on coding, online safety, cognitive skills, and computational thinking.

Key evaluation questions

In this study, we focus on whether individuals who attended TEGV after-school programs had better educational and psycho-social outcomes than their matched counterparts who did not attend TEGV programs, 10-12 years later. In particular, we focus on two key questions:

1. Did students who attended the TEGV programs achieve better outcomes than would likely have been the case had they not attended the TEGV programs? Specifically, this study aims to discover if they show improved outcomes in the following areas:
 - a. Educational attainment
 - b. High school GPA
 - c. Attitudes toward continuing education
 - d. Frequency of reading
 - e. Life satisfaction
 - f. Social responsibility
 - g. Interpersonal competence
 - h. Self-esteem and self-efficacy

2. Were the impacts of attending TEGV after-school programs similar for all program participants who attended these programs, or did different gender subgroups of program participants differentially benefit from attending these programs?

Although income is not an outcome targeted by TEGV programs, we have also examined whether the program had any impact on income or employment status of participants due to the policy relevance of these variables.

METHODOLOGY AND DATA

Study design and analytic approach

A rigorous impact evaluation requires the presence of a control group in order to estimate what would have happened in the absence of the intervention. Ideally, children who are eligible to attend the TEGV program would be randomly assigned to either the control group or the intervention group (TEGV group). Then, the intervention effect would be estimated by comparing the outcome scores of the two groups (i.e., randomized controlled trial design). In this study, there were no children assigned to a control group at the beginning of the intervention. Therefore, in order to have a comparison group of people that would allow us to estimate the effect of attending TEGV programs, we have used a matched comparison group design.

Two strategies were used in order to achieve a comparison group who would mimic the pre-intervention characteristics of TEGV participants. The first strategy was the use of purposive sampling. Data for the comparison group was collected from the same age group individuals living in the same cities and neighborhoods that are socioeconomically similar to the neighborhoods where TEGV programs were offered.

The second strategy was the use of propensity score matching analysis. Propensity score matching is a quasi-experimental method to match a participant to a non-participant with similar pre-intervention characteristics. The propensity score is defined as the probability of receiving a treatment based on a particular set of measured covariates. It is expected that balance on observed pre-treatment covariates will be achieved through matching on the estimated propensity of selecting the treatment. The assumption of propensity score matching is that outcomes are independent of program participation conditional on this particular set of observable characteristics.

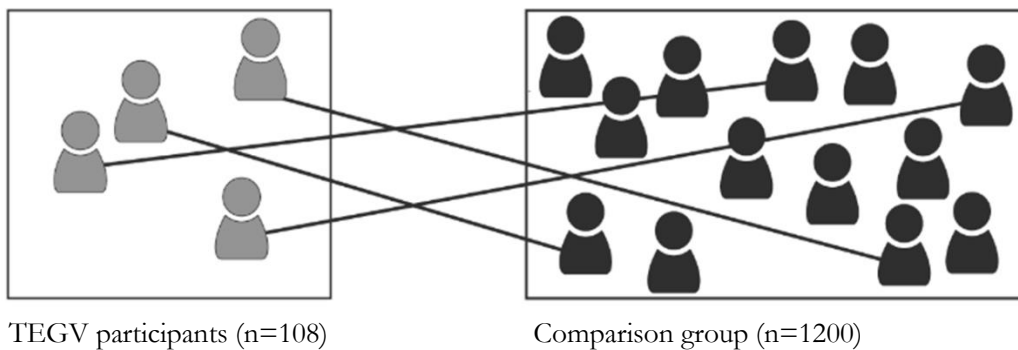
For every individual in the TEGV group, we have found up to two matching individuals among the comparison group who are as similar as possible in terms of observable pre-intervention

characteristics. The propensity score matching method was conducted to match the TEGV participants to the comparison group individuals on the following variables that could influence children’s probability of enrolling in the TEGV program and the outcome variables:

- a. Age
- b. Gender
- c. City
- d. Number of siblings
- e. Mother’s education level
- f. Father’s education level
- g. Grade repetition in primary school

As a result, with two groups of comparable individuals (i.e., TEGV participants vs. comparison group), the long-term effect of attending a TEGV program, if any, could be estimated more accurately using a non-experimental design (more details on the balance of covariates after matching are provided in the Appendix).

Figure 2. Illustration of propensity score matching



Characteristics of study participants

TEGV participants. Data was collected from 23- to 25-year-old individuals who enrolled in the TEGV after-school program at the age of 12-14 for the first time. As children were

allowed to enroll in TEGV programs (except for the informatics and Fireflies programs which did not exist at that time) for as many consecutive semesters as they wished, we have limited the number of participated weeks to between 8-96 (that is, up to two years) for the study sample (mean weeks of participation=38.11, SD=20.43). Participants were from 20 cities across Turkey. In total, data from 108 TEGV participants were collected (56.5% females).

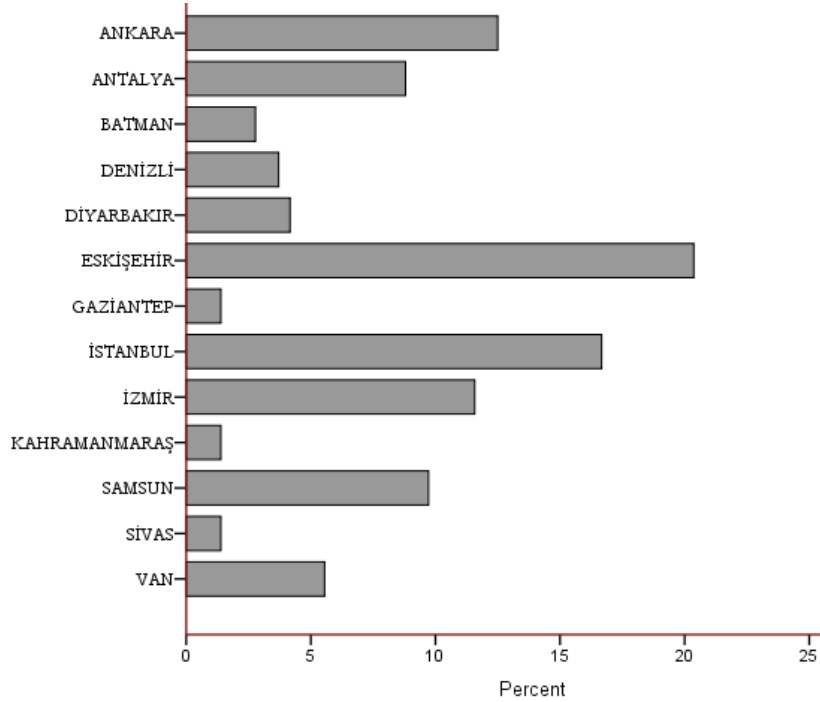
Non-participants. Data was collected from 23- to 25-year-old individuals living in the same 20 cities and neighborhoods that are socioeconomically similar to the neighborhoods where TEGV programs were given. In total, data from 1,212 non-participants were collected (48.2% females).

Matched sample. We have matched a TEGV participant up to two non-participants using the covariates listed above. Propensity score matching yielded a sample size of 216 individuals (80 TEGV participants, 136 non-participants) from the following 13 cities: Istanbul, İzmir, Ankara, Antalya, Batman, Denizli, Diyarbakır, Eskişehir, Gaziantep, Kahramanmaraş, Samsun, Sivas, and Van. All outcome analyses were conducted using the matched sample.

Table 2. Demographic characteristics of the final matched sample

Is a female	59.3 %
Age (mean)	24
Married	12 %
Is a parent	8.2 %
Has high school diploma	88.3 %
Owns a computer	83.8 %
Mother has high school diploma	19 %
Father has high school diploma	30,1 %
Number of siblings (mean)	2.48

Figure 3. Distribution of the matched sample across 13 cities



Outcome measures

Academic outcomes, employment, and life satisfaction

To assess the longer-term impacts of TEGV program on academic achievement, attitudes towards education, employment, and life satisfaction, we used the following outcome measures.

Education level. Education level of respondents are categorized as follows: 1=Not graduated from primary school, 2= not graduated from high school, 3=some high school, 4=high school graduate, 5=some college or more.

High School GPA. High school GPA were self-reported by the respondents using the following scale: 1=50-62, 2=63-71, 3=72-80, 4=81-90, 5=91-100.

Has a full-time job with social security benefits. A binary indicator equal to 1 if an individual reported having a full-time job with social security benefits.

Frequency of reading books. A single question asked respondents how frequently they engaged in reading books on a 1 to 5 scale, with 1 representing never and 5 representing every day.

Attitudes toward continuous education. We have used the short form of Adults' Attitudes toward Continuous Education scale (Blunt & Yang, 2002; Darkenwald et al., 1986). The scale contains nine statements about importance of continuing education, enjoyment of learning activities, and intrinsic value of continuing education using a 5-point Likert scale (1=Disagree Strongly, 5=Agree Strongly). A sample item is "Continuing education helps people make better use of their lives". Higher scores on this measure indicate more positive attitudes. The scale has satisfactory internal consistency (reliability coefficient $\alpha = .76$)

Income. A single question asked respondents to report their monthly earnings on a 1 to 5 scale (1 = less than 500 TL, 2 = 500 TL - 1000 TL, 3 = 1000 TL - 2000 TL, 4 = 2000 TL - 3000 TL, 5 = more than 3000 TL).

Life Satisfaction. Life satisfaction was measured using the Diener et al.'s (1985) Satisfaction with Life Scale. The scale contains five statements about general life satisfaction (i.e., subjective well-being) using a 7-point Likert scale (1=Disagree Strongly, 7=Agree Strongly). Higher scores on this measure indicate more life satisfaction. The scale has strong internal consistency (reliability coefficient $\alpha = .87$)

Interpersonal and intrapersonal competency, and social responsibility

To assess the longer-term impacts of TEGV program on social responsibility, interpersonal competence, and self-constructs, we used the following outcome measures.

Interpersonal Competence. Interpersonal Competence Questionnaire (Buhrmester et al., 1988) that measures the following five skills was used: initiating relationships, emotional support, asserting influence, self-disclosure, and conflict management was used. The scale includes 25 items rated on a 5-point Likert-type scale (1=Disagree Strongly, 5=Agree Strongly). The scale has good internal consistency (coefficient alpha=.86).

Self-esteem. Self-esteem, defined as relatively stable feelings of overall self-worth, was assessed using the Rosenberg Self-Esteem Scale. The scale consists of 10 items that refer to self-respect and self-acceptance rated on a 4-point Likert-type scale (1=totally disagree, 4=totally agree). Sample item is “I feel that I'm a person of worth, at least on an equal plane with others”. The scale has good internal consistency (coefficient alpha=.79).

Self-efficacy. Self-efficacy, defined as one's belief in one's ability to succeed in specific situations or accomplish a task, was assessed using the general self-efficacy subscale of Sherer et al.'s (1982) Self-Efficacy Scale. The scale consisted of 16 items using a 5-point Likert scale (1=Disagree Strongly, 5=Agree Strongly). Sample items are “when I set important goals for myself, I rarely achieve them,” and “when I decide to do something, I go right to work on it.” The reliability of the scale is good (coefficient alpha = .84).

Has volunteered for an NGO. A binary indicator equal to 1 if an individual reported having worked in a non-governmental organization as a volunteer.

FINDINGS

Comparing Outcomes for Participants and Nonparticipants

Academic outcomes, employment, and life satisfaction effect estimates

Our evidence suggests that the TEGV program had a statistically significant effect on educational attainment. TEGV participants achieved a higher level of schooling compared to non-participants (Table 3). However, among those who completed high school, respondents in both research groups reported similar levels of high school grade point average (GPA), with an average graduation GPA of 2.91 for TEGV group participants and a GPA of 2.80 for the comparison group on a 1 to 5 scale.

We found that the TEGV program had a large and statistically significant effect on attitudes toward continuing education. TEGV participants had more positive attitudes toward education through life, higher enjoyment of learning activities, and higher intrinsic values for continuing education. TEGV participants also reported higher frequency of reading books compared to non-participants.

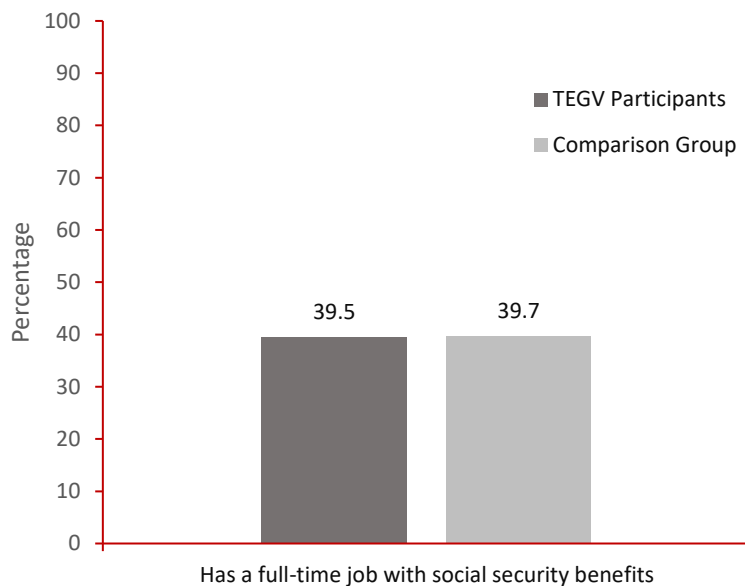
We have also examined whether the program had any impact on income and employment status of participants. Although income is not a long-term outcome targeted by TEGV programs, it may be relevant to policy, as evidence shows that higher levels of schooling is related to higher income. Our results indicate that even though TEGV participants had higher levels of schooling, they did not significantly differ in terms of employment status: 39.5% of TEGV participants reported having a full-time job with social security benefits, compared to 39.7% percent of individuals in the comparison group (Figure 4). Additionally, among those with a full-time job, respondents in both groups did not significantly differ in terms of monthly earnings. However, our findings suggest a significant effect of TEGV programs on life satisfaction, a targeted long-term outcome of TEGV. TEGV participants were more satisfied with their life compared to non-participants (Table 3).

Table 3. Effects on academic achievement, attitudes toward education, employment, and life satisfaction

Measure	Scale	TEGV Participants	Comparison Group	Difference	Effect size ²
Education level	1-5	4.84	4.51	0.33**	0.43
High School GPA	1-5	2.91	2.80	0.11	0.11
Attitudes towards continuing education	1-5	4.24	3.72	0.52***	0.76
Frequency of reading books	1-5	3.94	3.50	0.44*	0.33
Income	1-5	3.19	3.13	0.06	0.08
Life satisfaction	1-7	5.06	4.35	0.70***	0.46

* Significantly different from zero at the .05 level.
 ** Significantly different from zero at the .01 level.
 *** Significantly different from zero at the .001 level.

Figure 4. Percentage of respondents who have a full-time job with social security benefits among individuals who attended TEGV programs and similar individuals who did not attend TEGV programs.



² According to Cohen’s criteria, effects of size 0.20, 0.50, and 0.80 are considered “small”, “medium”, and “large”, respectively.

Interpersonal and intrapersonal competency, and social responsibility effect estimates

We found positive impact estimates with regard to intrapersonal competencies. TEGV participants had higher self-esteem compared to non-participants, with average scores of 3.51 and 3.22, respectively, on a 1 to 5 general self-esteem scale (Table 4). Similarly, TEGV participants had significantly higher generalized self-efficacy scores compared to non-participants. Both estimates had medium effect sizes.

TEGV participants also had significantly higher interpersonal competence scores, which comprised of skills such as initiating relationships, emotional support, asserting influence, self-disclosure, and conflict management, compared to non-participants. This effect was large in magnitude.

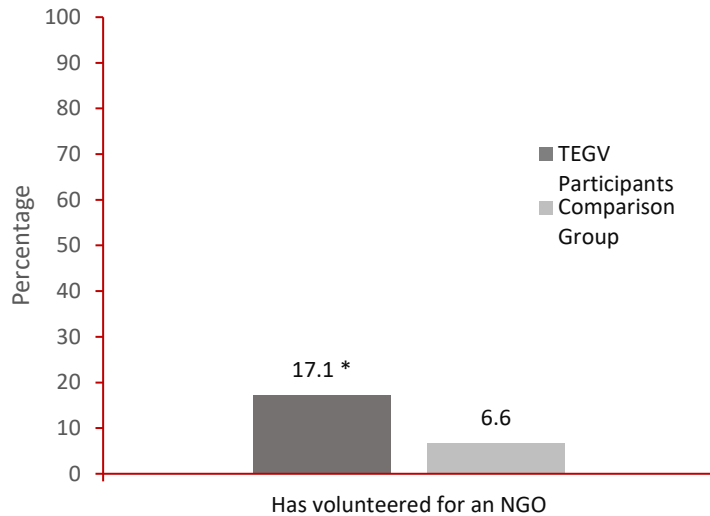
TEGV programs also aim for increased social responsibility in the long term. We have examined that this targeted outcome was realized. Our results indicate that 17.1% of TEGV participants engaged in voluntary work for a non-governmental organization compared to 6.6% of the respondents in the comparison group (Figure 5).

Table 4. Effects on interpersonal and intrapersonal competency

Measure	Scale	TEGV Participants	Comparison Group	Difference	Effect size
Interpersonal competence	1-5	3.82	3.55	0.27***	0.76
Self-esteem	1-5	3.51	3.22	0.29***	0.56
Self-efficacy	1-5	4.05	3.72	0.33***	0.52

*** Significantly different from zero at the .001 level.

Figure 5. Percentage of respondents who have volunteered for an NGO among individuals who attended TEGV programs and similar individuals who did not attend TEGV programs.



* Difference between TEGV participants and comparison group is significant at the 0.05 confidence level.

Box 2: The snowball effect

One of the long-term aims of TEGV programs is to make the participants TEGV delegates. The current survey shows that this aim is being realized and helps TEGV programs to become widespread in two forms: by increasing the number of future TEGV participants and increasing the number of TEGV volunteers.

- 93.1% of surveyed TEGV participants recommend TEGV programs to children and families they know.
- 66.7% of TEGV participants have facilitated a child to participate in one of TEGV's programs.
- 28.2% of TEGV participants have later become TEGV volunteers.
- 91.1% of TEGV participants recommend becoming a TEGV volunteer to individuals they know.
- 39.6% of TEGV participants have facilitated a person to be a volunteer for TEGV.

Subgroup analyses

We have examined whether TEGV's after-school programs were more effective for certain gender groups. Evidence shows that the estimated effects of attending TEGV programs on some of the targeted outcomes varied across males and females. Results indicate that for some outcome variables, the program effects were more salient for female respondents (Table 5). The effect of attending TEGV programs on the completed years of schooling was larger for females, compared to males. Also, the effect size for reading books was 0.43 for females and 0.22, though nonsignificant, for males. A noticeable result was the effect of TEGV programs on female participants' self-efficacy. Results indicate that TEGV programs had a boosting effect on female participants' self-efficacy compared to a nonsignificant effect for males.

Table 5. Program effects by gender groups

Measure	Treatment Effect Size	
	Female	Male
Education level	0.52*	0.38†
High School GPA	0.14	0.05
Attitudes towards continuing education	0.79***	0.79***
Frequency of reading books	0.43*	0.22
Income	0.15	0.04
Life satisfaction	0.44*	0.47*
Interpersonal competence	0.55**	0.39†
Self-esteem	0.60**	0.51*
Self-efficacy	0.82***	0.19

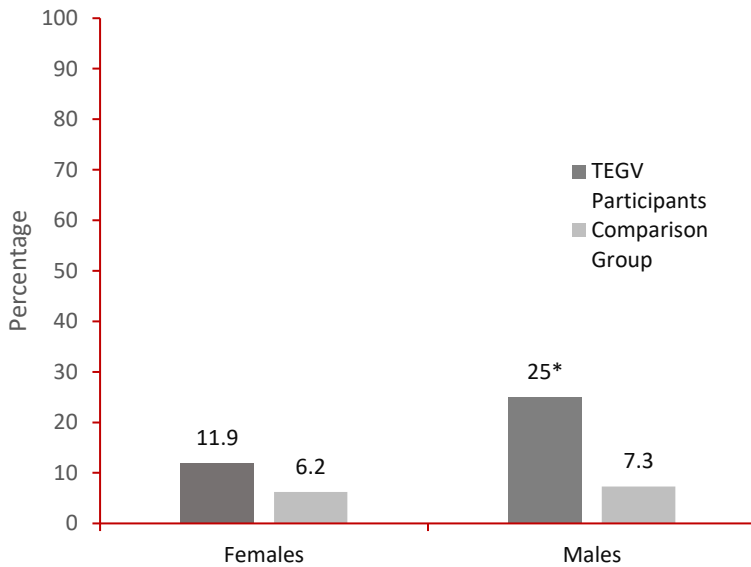
* Significantly different from zero at the .05 level.

** Significantly different from zero at the .01 level.

*** Significantly different from zero at the .001 level.

† Marginally significantly different from zero at the .10 level.

Figure 6. Percentage of respondents who have volunteered for an NGO among individuals who attended TEGV programs and similar individuals who did not attend TEGV programs by gender groups.



* Difference between TEGV participants and comparison group is significant at the 0.05 confidence level.

Figure 7. Percentage of respondents who have a full-time job with social security benefits among individuals who attended TEGV programs and similar individuals who did not attend TEGV programs by gender groups.



UNANSWERED QUESTIONS

The current report investigated challenging but critical and policy-relevant questions regarding the long-term gains of after-school programs. Yet there are important unanswered questions that would be helpful for the implementation of future programs.

Dosage Effect

Due to low sample size, we did not have the opportunity to investigate the dosage effect.

Dosage reflects the amount of participation of children in the TEGV programs. It may be related to attendance level of a student or the duration of the program.

Studies show that program effect sizes vary as a result of the number of program hours a child received in an after-school program. For instance, according to a comprehensive meta-analysis study (Lauer et al., 2006), the highest effect size for out-of-school time programs happens when the programs last between 44-84 hours for

TEGV students have the opportunity to participate in more than one structured program consecutively and attend further elective classes. One important future research question would be to investigate the relationship between the improvement in the targeted outcomes and amount of exposure to TEGV programs. Such a study would also inform decision makers on the opportunity cost of reaching more children (that is output) versus creating a higher and long-lasting impact for the participating children.

the reading outcome and 46-100 hours for the mathematics outcome. Other studies also show that greater intensity and duration of the programs are associated with better outcomes, and students benefit from the program if they only receive a sufficient dosage (McComb & Scott-Little, 2003). **Future studies at TEGV can test the effect of number of program hours received on improving the targeted outcomes. And with the help of modern causal inference methods, we are now able to obtain dosage estimates by establishing causality and moving further from correlational results (Kluve et al., 2012).**

For Whom?

The subgroup analyses in this study were limited to gender groups, yet there may still be significant variation on program effectiveness across different groups of participants. It is possible that program effects would be stronger for children who were relatively low on academic achievement and psycho-social skills at baseline. In this study, we also did not learn about the potential benefits of after-school programs for children from middle and high economic statuses. Determining for whom the after-school programs work or do not work (that is, by which variables the program effect is moderated) is crucial for revising the program according to different groups or deciding which groups to recruit if a targeted strategy is adopted. From the policy makers' perspective, this knowledge would guide the funding and implementation strategies on the use of after-school programs to promote positive child and youth development in Turkey.

How?

Another unanswered question is through which mechanisms has the TEGV after-school program achieved its effects? For example, we have found that TEGV participants had higher life satisfaction in the long term compared to non-participants. However, we do not know which element of TEGV programs resulted in this outcome. Although one can reflect on how the program led to changes in the outcome variable, modern statistical methods now allow us to quantitatively test these indirect effects (aka mediated effects) and obtain individual effect sizes for each component of an intervention (Kisbu-Sakarya, MacKinnon, O'Rourke, 2015; MacKinnon, 2008). **By opening their black box and objectively testing their mediating mechanisms, after-school programs may be revised via selecting effective components and deleting counterproductive components.**

Are Certain Programs, Approaches, or Delivery Features More Effective Than Others?

In this report, we were not able to test the distinct effects of different after-school programs offered by TEGV (such as the science program or the arts program) due to the ex post facto nature of the study. **Future rigorous impact evaluation studies (preferably using a**

randomized controlled design) can obtain unbiased effect estimates for each of TEGV's after-school programs. These impact evaluation studies can also incorporate the testing of different learning approaches or delivery features (such as the use of online training for a specific program or module) by using factorial experiments in order to optimize program effectiveness (Collins et al., 2009a, 2009b).

Another advantage of conducting rigorous impact evaluations is that once the impact estimates are obtained, they can be combined with program cost information to compute a cost benefit balance for a program or compare various program alternatives targeting the same outcome (such as reduced school dropout rates) in terms of cost effectiveness. In combination with program optimization studies to increase program effectiveness, this cost-related information would help decision makers in the allocation of the limited budget for the dissemination of successful programs.

LIMITATIONS

One should note that the results of the current report should be interpreted with caution due to some study limitations. As there was no children assigned to a control group at the beginning of the study (that is when TEGV participants enrolled in the program), we have used a matched comparison group to estimate the program effects. However, the variables we have used in matching were rather limited. Therefore, there may still exist confounders that may bias the impact estimates. Another concern is that we were limited to self-reporting of respondents when measuring some hard-to-remember numeric variables, such as high school GPA, because we did not have access to administrative data at the individual level. A further challenge was reaching the TEGV alumni who had participated in the program 10-12 years ago. We could only reach participants for whom we had a valid contact address or phone number. Therefore, one cannot know if those we were able to contact have different characteristics compared to unreached TEGV alumni, which may bring a potential bias for the impact estimates.

References

- Baker, D., & Witt, P. A. (1996). Evaluation of the impact of two after-school programs. *Journal of Park and Recreation Administration, 14*, 23–44.
- Baker, E. L. & Gribbons, B. (1998). *Evaluating the Long-term Impact of After School Programs: Applying New Methodologies to Assess the Effects of LA's BEST on Student Performance*. Los Angeles: University of California.
- Blunt, A., & Yang, B. (2002). Factor Structure of the Adult Attitudes Toward Adult and Continuing Education Scale and its Capacity to Predict Participation Behavior: Evidence for Adoption of a Revised Scale. *Adult Education Quarterly, 52*(4), 299-314.
- Brooks, P. E. (1995). *Longitudinal Study of LA's BEST After School Education and Enrichment Program, 1992-1994*. Los Angeles.
- Buhrmester, D., Furman, W., Wittenberg, W.T., & Reis, H.T. (1988). Five domains of interpersonal competence in peer relationships. *Journal of Personality and Social Psychology, 55*(6), 991-1008.
- Campbell, P. B., Storo, J. & Acerbo, K. (1995). *Math, Science, Sports, and Empowerment: Girls Incorporated Replication and Expansion of the Eureka! Model, Executive Summary*. Groton, MA: Campbell-Kibler Associates.
- Cardenas, J.A. (1992). The Coca-Cola Valued Youth Program: Dropout prevention strategies for at-risk students, *Texas Researcher* (Vol. 3, pp. 111-130).
- Carlisi, A. M. (1996). *The 3:00 Project Program Evaluation*. Decatur, GA: Georgia School-Age Care Association.
- Collins, L. M., Chakraborty, B., Murphy, S. A., & Strecher, V. (2009a). Comparison of a phased experimental approach and a single randomized clinical trial for developing multicomponent behavioral interventions. *Clinical Trials, 6*(1), 5-15.
- Collins, L. M., Dziak, J. J., & Li, R. (2009b). Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs. *Psychological Methods, 14*, 202-224.
- Cross, A. B., Gottfredson, D. C., Wilson, D. M., Rorie, M., & Connell, N. (2010). Implementation Quality and Positive Experiences in After-School Programs. *American Journal of Community Psychology, 45*(3-4), 370-380.

- Darkenwald, G.G., & Hayes, E.R. (1986). Adults' attitudes towards continuing education. Paper presented at the National Conference of the American Association for Adult and Continuing Education.
- Darkenwald, G. G., & Hayes, E. R. (1988). Assessment of adult attitudes toward continuing education. *International Journal of Lifelong Education*, 7(3), 197-204.
- Diener, E.D., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction with life scale. *Journal of Personality Assessment*, 49, 71-75.
- Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Collaborative for Academic, Social, and Emotional Learning.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology*, 45, 294-309.
- Granger, R. C. (2008). After-School Programs and Academics: Implications for Policy, Practice, and Research. Social Policy Report. Volume 22, Number 2. *Society for Research in Child Development*.
- Hamilton, L. S. & Klein, S. P. (1998). Achievement test score gains among participants in the Foundations School Age Enrichment Program (pp. 1-16). Unpublished manuscript.
- Hamilton, L. S., Le,V. & Klein, S. P. (1999). *Foundations School-Age Enrichment Program: Evaluation of Student Achievement*. Santa Monica, CA: Rand Education.
- Huang, D. (2001). An after-school evaluation system for middle and high school programs, *National Association of Secondary School Principals Bulletin* (Vol. 85, pp. 45-61).
- Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, E. L. (2000). *A Decade of Results: The Impact of the LA's Best After School Enrichment Program on Subsequent Student Achievement and Performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation.
- Johnson, L. J., Zorn, D., Williams, J., & Smith, J. (1999). *1998-99 School Year Program Evaluation: Urban School Initiative School Age Child Care Expansion*. Cincinnati, OH: University of Cincinnati.
- Jones, M. B. & Offord, D. R. (1989). Reduction of antisocial behavior in poor children by non-school skill development, *Journal of Child Psychology and Psychiatry*, 30, 737-750.
- Kisbu-Sakarya, Y., MacKinnon, D.P. & O'Rourke, H. (2015). Statistical models of mediation for drug program evaluation. In L. M. Scheier (Ed.), *Handbook of drug abuse*

- prevention research, intervention strategies, and practice. Washington, DC: American Psychological Association.
- Kluve, J., Schneider, H., Uhlendorff, A., & Zhao, Z. (2012). Evaluating continuous training programmes by using the generalized propensity score. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *175*(2), 587-617.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, *76*(2), 275-313.
- Leos-Urbel, J. (2015). What works after school? The relationship between after-school program quality, program attendance, and academic outcomes. *Youth & Society*, *47*(5), 684-706
- Little, P., Wimer, C., & Weiss, H. B. (2008). After school programs in the 21st century: Their potential and what it takes to achieve it. *Issues and opportunities in out-of-school time evaluation*, *10*(1-12).
- MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.
- Marshall, N., Coll, C. G., Marx, F., McCartney, K., Keefe, N., & Ruh, J. (1997). After-school time and children's behavioral adjustment, *Merrill-Palmer Quarterly*, *43*, 497-514.
- McComb, E. M., & Scott-Little, C. (2003). A review of research on participant outcomes in after-school programs: Implications for school counselors. *ERIC Digest*. Greensboro, NC: ERIC Clearinghouse on Counseling and Student Services.
- Miller, B. M. (2003). *Critical Hours: Afterschool Programs and Educational Success*. Nellie Mae Education Foundation.
- Posner, J. K. & Vandell, D. L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development*, *65*, 440-456.
- Posner, J. K. & Vandell, D. L. (1999). After-school activities and the development of low income urban children: A longitudinal study, *Developmental Psychology*, *35*, 868-879.
- Redd, Z., Cochran, S., Hair, E., & Moore, K. (2002). *Academic achievement programs and youth development: A synthesis*. Washington, DC: Child Trends.
- Rodriguez, E., Hirschl, T. A., Mead, J. P., & Groggin, S. E. (1999). *Understanding the Difference 4-H Clubs Make in the Lives of New York Youth: How 4-H Contributes to Positive Youth Development*. Ithaca, NY: Cornell University.

- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Schinke, S. P., Cole, K. C. & Poulin, S. R. (1998). Research report: Thirty month data and process findings: Evaluation on educational enhancement program of Boys and Girls Clubs of America (pp. 1-31).
- Schinke, S. P., Orlandi, M.A. & Cole, K. C. (1992). Boys and girls clubs in public housing developments: Prevention services for youth at risk, *Journal of Community Psychology*, 118-128.
- Schlegel, C. (2003). Citizen Schools: Evaluation summary. In B. M. Miller (Ed.). Boston, MA.
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, 51(2), 663-671.
- Vandell, D. L. & Pierce, K. M. (1999). *Can After-school Programs Benefit Children Who Live in High-crime Neighborhoods?* Paper presented at the Biennial meeting of the Society for Research in Child Development, Albuquerque, NM.
- Warren, C., Feist, M. & Nevarez, N. (2002). *A Place to Grow: Evaluation of the New York City Beacons*. New York City: Academy for Educational Development.
- Zief, S. G., & Lauver, S. (2006). *The impacts of after-school programs on student outcomes: A systematic review for the Campbell Collaboration*.

APPENDIX: Covariate balance after matching

After matching, we have examined the balance of all observed covariates. Results indicated that matching improved overall balance. No covariate exhibited a large imbalance ($d > .25$) after matching.

Figure A. Absolute standardized differences in means before and after matching.

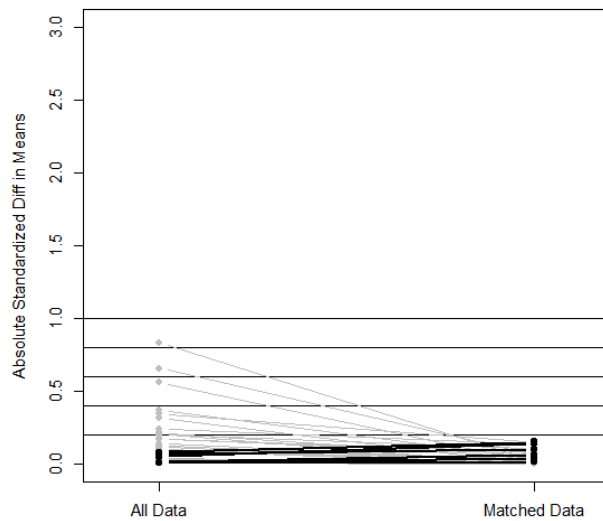
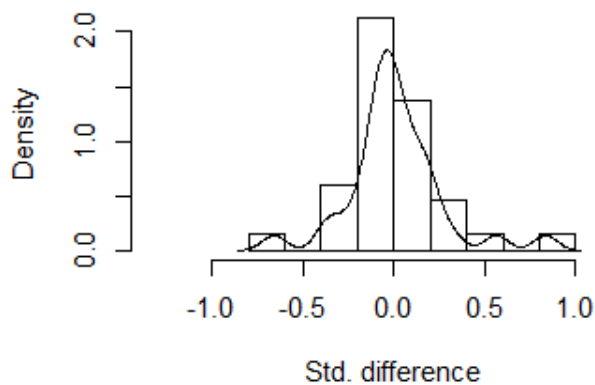


Figure B. Histograms with overlaid kernel density estimates of standardized difference before and after matching.

Standardized differences before matching:



Standardized differences after matching:

