Impact Evaluation: Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt, 2011–2012

Final Report

August 2013

Submitted to:

U.S. Department of Labor
Office of Child Labor, Forced Labor, and Human Trafficking
Frances Perkins Building
200 Constitution Avenue NW
Washington, DC 20210

Submitted by:

11785 Beltsville Drive
Calverton, MD 20705

ACKNOWLEDGMENTS

This report is based on evaluation research undertaken by ICF International (ICF) under its contract with the U.S. Department of Labor's (USDOL) Office of Child Labor, Forced Labor, and Human Trafficking (OCFT). The author and ICF wish to thank the staff of OCFT and the World Food Programme (WFP) for their support and collaboration in the design and execution of the study. We also thank the study participants and their families, whose experiences will be invaluable in guiding future programming decisions.

The author of this report is Pablo Diego Rosell, who had the primary responsibility of cleaning the data, conducting all data analyses, and preparing the final report. This report benefited substantially from inputs provided by Professor Eric Edmonds, who provided methodological guidance and content expertise. The report also contains important input from Matt Holtman, Sarah Jones, Stacey Merola, and Joanne Liem, who were responsible for evaluation design and baseline report, and from Nahla Hassan (author of the midterm evaluation). The evaluation and final report received significant contributions from Nivine Osman and Nesreen Elmolla, of WFP; Holly Howell, of ICF; and Maureen Jaffe, Celeste Lemrow, Karina Jackson, and Sharon Heller, of USDOL.

Funding for this evaluation was provided by the United States Department of Labor under Task Order number DOLB129K33758. Points of view or opinions in this evaluation report do not necessarily reflect the views or policies of the United States Department of Labor, nor does the mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

CONTENTS

ACRONYMS

Acronyms	Definitions
2SLS	Two-stage least-squares
AR	Awareness Raising
CAPMAS	Central Agency for Public Mobilization and Statistics
CDA	Community Development Association
CDC	Cairo Demographic Center
CMEP	Comprehensive Monitoring and Evaluation Plan
CS	Community Schools
CWCLP	Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt
ICC	intra-cluster correlation
IGA	Income-Generating Activities
ILAB	Bureau of International Labor Affairs
ILO	International Labour Organization
IRB	Institutional Review Board
ITT	Intention to Treat
MDES	minimum detectable effect size
MOE	Ministry of Education
MOMM	Ministry of Manpower and Migration
NAP	National Action Plan
NCCM	National Council for Childhood and Motherhood
NGO	Nongovernmental Organization
OCFT	Office of Child Labor, Forced Labor, and Human Trafficking
OLS	Ordinary Least-Squares
RA	Rapid Assessment
THR	Take-Home Rations
TOT	Treatment on the Treated
UNICEF	United Nations Children's Fund
USDOL	U.S. Department of Labor

WFCL	Worst Forms of Child Labor
WFP	World Food Programme

EXECUTIVE SUMMARY

This is the final impact evaluation report of the Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt (CWCLP) project, implemented by the World Food Programme and funded by the U.S. Department of Labor.

The impact evaluation of the CWCLP project aimed to estimate the impact of the Community Schools (CS), Take-Home Rations (THR), Awareness Raising (AR), and Income-Generating Activities (IGA) components of the project on school and work-related outcomes among eligible children. Eligible children were those children in program communities that were aged 6 to 11 at baseline, not currently enrolled in school at baseline, and either working in or at risk of working in child labor.

A total of 116 villages, including 1,423 eligible children, were randomized at baseline to a control or a treatment condition. The 36 villages randomized to the control condition did not receive any program intervention. The 80 villages randomized to the treatment condition received a CS that offered transitional education based on the national primary school curriculum, using an active learning methodology. The CS operated 6 hours a day (from 8:00 a.m. to 1:00 p.m.), 6 days per week during the 9 months of the school year. During the 3 months of the summer vacation period, it operated 6 hours a day (from 8:00 a.m. to 1:00 p.m.), 3 days per week. Enrollment in the CS was offered to eligible children in the community. Families of eligible children were offered support to enroll their children. Education expenses, school supplies, and in-school snacks were also offered to all children enrolled in CS. Eligible children that enrolled in a CS also received a monthly THR conditional on 80 percent CS attendance. The THR, consisting of provisions of rice, wheat flour, and oil, was valued at around EGP 75.00 monthly (USD 11.00). Adding school expenses covered by the project, equivalent to around EGP 30.00 monthly, the total package would add up to approximately EGP 105.00 per month, or about 25 percent of an average child's monthly wage (typically around EGP 400.00).

Additionally, families of eligible children received support to gain access to small business loans and also received training on relevant IGAs. A total of 551 projects had started by September 2012 as a direct result of the project trainings, including a variety of micro-businesses in the agriculture, manufacturing, retail, and service sectors.

Finally, treatment communities received AR campaigns, including community activities targeting parents, children, formal and informal leaders, and employers through messaging on the value of education, children's rights, and strategies to offset any loss of income for vulnerable families and to promote CS activities.

Baseline data were collected from children and their households between October and November 2011 (the high agricultural season, when most harvesting occurs), whereas endline data were collected in March and April (a less active period for agricultural work). Data were collected from all eligible households, regardless of whether children actually enrolled in school. Data collected on both rounds included information on household composition, education and work status of household members, time allocation of children in the household, household attitudes,

and consumption. Child interviews collected information on school and work status, exposure to workplace hazards and injuries, and time allocation.

This evaluation found significant evidence that the CWCLP project had a positive impact on schooling outcomes of eligible children, including ever school attendance (attended school at least once, ever), school enrollment, school attendance, and time spent on school-related activities (including time spent attending school, commuting to and from school, and doing homework). Impacts on school-related outcomes were larger for girls, for younger children (6–8 years old) and for those children that were not involved in economic activities at baseline.

The evaluation also found a significant reduction on the allocation of time to economic activities for boys (defined as time spent on economic activities per week), and a significant reduction on the allocation of time to unpaid household services for girls (defined as time spent on unpaid household services per week). The fact that these impacts were significant for a specific sex only is likely the result of differential time allocation patterns, with boys spending roughly twice as much time in economic activities than girls, and girls spending roughly twice as much time in unpaid household services than boys. The evaluation, however, did not detect an impact on overall participation in economic activities (defined as having performed any economic activity in the last 12 months), exposure to workplace hazards, or occurrence of work-related injuries, either for the full sample or any of the subgroups.

These findings, which are consistent with previous randomized control trials on the effect of schooling incentives on child labor (e.g., Edmonds & Shrestha, 2012), suggest that the current treatment may need some reformulation in order to reach the stated objective of reducing participation in hazardous child labor. One such strategy would be to include a stronger component to increase knowledge and understanding (among children and parents) of the hazards of agricultural labor.

There are several limitations to the study. Seasonality in the data and the possibility of biased recall suggest that impacts identified by this evaluation may be more readily applicable to off-seasonal work-related outcomes. If recall bias was indeed an issue (a possibility that we cannot test), there is a possibility that the current results are less informative about program effects on child work during the high agricultural season. It is worth noting, in any case, that our estimates of program impacts are based on treatment/control comparisons, so impact estimates are not seasonally biased, as would be the case with a pre/post comparison. The lack of administrative data and the impossibility of evaluating individual program components separately are also limitations of the study.

Finally, we must note that the results from this impact evaluation are not generalizable beyond the study population. The sample selected for this evaluation is representative for those communities in which the CWCLP project is currently operating. These communities represent a very specific set of small rural hamlets with no primary school where children were engaged in or at risk of participating in hazardous child labor. The project may operate differently in other parts of the country, so the sample is not generalizable to the whole population of Egypt. However, we believe that the impact of this intervention might be replicable in populations similar to those in the sample, that is, children ages 6–11 living in small rural hamlets with no

primary school where children were engaged in or at risk of participating in hazardous child labor.

1. INTRODUCTION

ICF International (ICF) provides evaluation services to the U.S. Department of Labor's Office of Child Labor, Forced Labor, and Human Trafficking (USDOL/OCFT). OCFT is part of the USDOL's Bureau of International Labor Affairs. The office conducts research on international child labor, forced labor, and human trafficking; funds and oversees the efforts of organizations to eliminate exploitative child labor around the world; and assists in the development and implementation of U.S. government policy on international child labor, forced labor, and human trafficking issues.

Under task order DOLB129K34042, ICF is providing technical assistance and services to conduct an impact evaluation on the project titled Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt (CWCLP). The CWCLP project represents a partnership between the World Food Programme, the International Labour Organization (ILO), and the United Nations Children's Fund, along with a number of local organizations that implemented a set of interventions to reduce child labor and strengthen communities.

The impact evaluation examines the combined effects of Community Schools (CS), take-home rations (THR), Awareness Raising (AR), and Income-Generating Activities (IGA) on the participation in worst forms of child labor and school enrollment. Specifically, the report first provides an overview of child labor in Egypt and the overarching CWCLP interventions, as well as a description of the intervention components that are targeted for the impact evaluation. Next, the report describes the evaluation methodology, including its objectives, participants, and definition of treatments. The hypothesized outcomes are presented next, followed by a section on sample selection and randomization, the empirical methods used to screen data and test hypotheses, as well as the control variables used, participant flow across the different phases of the project, and participant recruitment. Finally, results from the endline comparison of the groups are presented, followed by a discussion of the findings and the limitations of the study.

1.1. BACKGROUND

As in most countries, child labor in Egypt results from a complex combination of factors. These include inadequate levels of household income and food security, limited quality and accessibility of education services, inadequate enforcement of child labor legislation, lack of awareness of the potential dangers of child labor, and cultural norms favoring children's early participation in work. The CWCLP project addresses each of these factors with a specific set of interventions.

1.1.1. CHILD LABOR IN EGYPT

Child labor is a common phenomenon in Egypt. While the estimated number of working children varies by source, it is clear that child labor is a significant issue in the country. According to the 2011 Report on the Worst Forms of Child Labor, Egypt has a population of approximately 993,417 working children (6.7 percent of children 5–14 years of age), many whom are working

in agriculture and domestic labor. Some other occupations in which children in Egypt are involved include the production of bricks, glass, and leather; fishing; blacksmithing; construction; carpentry; mechanical repair; mining; petty street trading; and possibly limestone quarrying.[1] A nationally representative survey conducted by the ILO and the Central Agency for Public Mobilization and Statistics (CAPMAS) in 2010 provides an estimate of 1.6 million children aged 5–17 engaged in child labor in Egypt, among which 64 percent worked in agriculture.[2] However, the National Council for Childhood and Motherhood and CAPMAS estimate that 3 million children are involved in labor in Egypt based on a 2009 study.[3] The study indicates that 64 percent of child workers are engaged in agriculture.[4] Children working in the agriculture sector work long hours, sometimes in extreme temperatures. Their work often leads them to use dangerous machinery, spray pesticides, inhale fumes or dust, bend for protracted periods, and carry heavy loads.[5]

1.1.2. POLICY CONTEXT

The Labor Law, No. 12 of 2003, stipulates that children cannot be employed until age 14, except that children may be trained, starting at age 12. Minors cannot work more than 4 consecutive hours, more than 6 hours per day, after 7:00 p.m., or overtime hours. However, the law explicitly excludes those working in agriculture, or those employed as domestic workers or working for family members.

The Child Law, Law No. 126, was enacted in 2008 and sets age limits for child employment. Children aged 15 and older are eligible for regular employment and children aged 12 and older are eligible for seasonal employment or apprenticeships. However, this law excludes domestic work, work in a family-run business, and agricultural work. This is significant, as the CWCLP focuses primarily on children who work in agriculture. Furthermore, Egypt identified 44 specific hazardous occupations under Decree 118 of the Ministry of Manpower and Migration. This did not include agriculture, however. This leaves a significant gap in protection for child laborers in this potentially dangerous field.

Local trade unions report that Egypt's labor laws are well enforced in the formal sector. By contrast, the Government does not seem to be enforcing the labor laws effectively in the informal

[1] U.S. Department of Labor's Bureau of International Labor Affairs. (2012). "Egypt," in Findings on the Worst Forms of Child Labor—2011, Washington, DC. Retrieved from http://www.dol.gov/ilab/programs/ocft/2011TDA/egypt.pdf.

[2] International Labour Organization and the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS). (2012). *Working children in Egypt: Results of the 2010 National Child Labour Survey.* Cairo: ILO. Retrieved from http://www.ilo.org/public/libdoc/ilo/2012/469638.pdf.

[3] Abu al Khair, Waleed. "Child Labour in Egypt a Growing Problem," Al-Shorfa.com, [online], October 14, 2010 [cited April 25, 2012]. Retrieved from http://al-shorfa.com/en_GB/articles/meii/features/main/2010/10/14/feature-02.

[4] International Labour Organization and the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS). (2012). *Working children in Egypt: Results of the 2010 National Child Labour Survey.* Cairo: ILO. Retrieved from http://www.ilo.org/public/libdoc/ilo/2012/469638.pdf.

[5] U.S. Department of Labor's Bureau of International Labor Affairs. (2012). "Egypt," in Findings on the Worst Forms of Child Labor—2011, Washington, DC. Retrieved from: http://www.dol.gov/ilab/programs/ocft/2011TDA/egypt.pdf.

sector, including small factories and workshops, where observers have reported that employers often violate child labor laws.[6]

1.2. OBJECTIVES

The objective of this impact evaluation was to examine the effect of the CWCLP program on children engaged in or at risk of participating in exploitative child labor. The components of the program were implemented in combination and include the provision of CS, THR, AR, and IGA. Specifically, this evaluation aimed to understand the impact of the CWCLP program on school-related and work-related outcomes among eligible children.

The evaluation was designed to provide information relevant to program managers, policy planners, researchers, and funding agencies, among others. The evaluation results were intended to help establish whether this particular intervention works as expected. The results are also expected to contribute to future program evaluation efforts by helping to build the body of knowledge concerning the impacts that can be expected of child labor interventions and by highlighting the methodological challenges associated with this kind of evaluation research. The specificity of the target population means that the results of this study may not be readily extrapolated to other contexts. That said, the issues raised by this evaluation are potentially relevant in many contexts.

[6] U.S. Department of State. (2013). "Egypt," in Country Reports on Human Rights Practices—2012. Washington, DC: U.S. Department of State. Retrieved from http://www.state.gov/j/drl/rls/hrrpt/humanrightsreport/index.htm#wrapper.

2. METHODS

2.1. EVALUATION PARTICIPANTS

The baseline survey was carried out in five governorates of Egypt (Assuit, Menya, Sohag, Fayoum, and Sharqiyah), in 116 small rural villages[7] that were identified as having no existing primary school serving a substantial segment of children who were school aged and where children were at risk or already participating in exploitive work. In order for a child to be considered eligible, they needed to meet the following criteria:

- Child between 6 and 11 years old;
- Child not currently enrolled in a national government school;
- Absent from school for at least the last 2 years; and
- Child engaged in or at risk of exploitative child work in agriculture. Children at risk of exploitative child work were defined by the project as the siblings of child laborers under age 15 and living in the same household.

Each household could have a maximum of two children who could participate in the study, while the other children could be enrolled in the project without participating in the study. If a single household had more than two eligible children, two children were selected at random for participation.

2.2. COMBATING WORST FORMS OF CHILD LABOR BY REINFORCING POLICY RESPONSE AND PROMOTING SUSTAINABLE LIVELIHOODS AND EDUCATIONAL OPPORTUNITIES IN EGYPT PROJECT INTERVENTIONS

Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt (CWCLP), funded by the U.S. Department of Labor (USDOL) in December 2010, is a 4-year, $9.5 million project implemented by the World Food Programme (WFP) and its sub-partners, the International Labour Organization (ILO) and the United Nations Children's Fund (UNICEF). The primary objective of CWCLP is to use a multi-pronged approach to reduce the worst forms of child labor in the targeted communities. Through CWCLP, the WFP, ILO, and UNICEF provide beneficiaries aid, using an approach toward the elimination of worst forms of child labor.

In quantitative terms, the program-wide goals of CWCLP state that partners will provide 5,000 children who are under the age of 15 and already engaged in exploitative child labor with educational support, including transitional education opportunities and Take-Home Rations (THR). Three thousand children ages 14–17 will receive a package of incentives, as well as on-

[7] Although 116 villages were randomized to a treatment or control condition, the final baseline and endline samples of children included 115 villages, as no eligible children could be interviewed in one treatment community, which had to be excluded.

the-job and off-the-job training opportunities, including apprenticeships. Eight thousand children who have been identified as "at high risk" of entering labor (primarily the siblings of children who are working) will receive a package of incentives, including THRs, enhanced educational opportunities, and facilities. Finally, 5,000 heads of household (with a focus on mothers who are vulnerable or at-risk of having a child who labors) will be provided livelihood development and financial empowerment activities.

In particular, WFP, ILO, and UNICEF, in collaboration with local partners, designed CWCLP with five central components to meet their stated goals:

1. **Reducing exploitive child labor and providing services to promote education:** A CWCLP education component includes support for those children who are already involved in formal education, such as state-supported schools. Additionally, CWCLP targets children engaged in child labor or who are at risk of joining the labor pool due to their level of poverty or because they have siblings who are working. For these target children, non-formal educational services are provided in the form of CS.

2. **Promoting sustainable livelihoods in targeted households:** A livelihood is defined in this project as a means of living, and the capabilities, assets (including both material and social resources, such as, food, potable water, health facilities, educational opportunities, housing, and time for participation in the community), and activities required for it. A livelihood encompasses income, as well as social institutions, gender relations, and property rights required to support and sustain a certain standard of living. It includes access to and benefits derived from social and public services provided by the state, such as education, health services, microfinance, and other infrastructure. Sustainable livelihood programs seek to create long-lasting solutions to poverty by empowering their target population and addressing their overall well-being. The sustainable livelihoods component aims to provide small business training and technical assistance to households, and support mothers of beneficiary children in implementing Income-Generating Activities (IGA).

3. **Raising awareness of exploitive child labor and the importance of education for all children:** This component seeks to raise awareness of the negative effects of child labor and the importance of education through the development of targeted communications messages and mass community meetings and awareness-raising events.

4. **Supporting the review and revision of legislation on child labor and improving the capacity of organizations to participate in this revision:** This component aims to strengthen policies addressing child labor by supporting the review and revision of legislation impacting child labor. The intent is to implement these changes by 2015. The project plans to cooperate with the Ministry of Manpower and Migration (MOMM), National Council for Childhood and Motherhood and all relevant stakeholders on the translation of the National Strategy for the Elimination of Child Labor into a National Action Plan. These activities will provide a framework for the Government to develop a national decree for the elimination of child labor.

5. **Improving national research capacity on child labor:** The project supports research, evaluation, and the collection of reliable data on child labor through population surveys. This activity is important to illuminate the root causes of child labor and to identify effective strategies, policies, and good practices to combat it.[8]

As the segment above shows, the CWCLP program consists of integrated interventions targeted at children, families, community members, and relevant policy-makers to help eliminate the worst forms of child labor. This impact evaluation, undertaken by ICF International (ICF) in partnership with USDOL and WFP, *does not* evaluate all of the components above. Rather, it examines the effect of a specific combination of interventions that were implemented in the first year of the program. The combination of interventions includes Awareness Raising (AR), IGA, Community Schools (CS), and THR. These components are described in more detail in the following section.

2.3. DEFINITION OF TREATMENT

Child eligibility for the intervention was determined by the local data collection firm, El Zanaty and Associates, at baseline. Eligible children were those children in the program communities that were aged 6 to 11 at baseline, not currently enrolled in school, and either working in or at risk of working in child labor. Eligible children were offered the opportunity to enroll in a CS and to receive a THR, conditional on CS attendance. Heads of households of eligible children (with a focus on mothers) were then selected to receive livelihood support through IGAs. All treatment communities received AR campaigns. It is expected that most beneficiaries were offered all four components of the treatment, although IGA assignment followed an additional screening process. Each component is described below.

Community Schools

A CS is a project-implemented primary education school serving small rural communities that do not have access to conventional primary schools. The CS component provides non-formal educational services to children engaged in or at risk of engaging in child labor. All CS offered the national, Ministry of Education (MOE)–sanctioned curriculum for primary schools. The only difference from government schools was in the pedagogical methodology. CS implement an active learning methodology, which is made possible by their reduced class sizes relative to government schools. Children are expected to attend a CS for 2 years. Students are required to enter and pass an exam to obtain their diploma. This exam is based on the national curriculum and put forward and supervised by the Directorate of Education of each governorate. The child who passes the exam receives a national primary certificate and is thus eligible to be streamlined

[8] While this information is not related to national research capacity on child labor, it was important to mention that CWCLP is developing a new child labor monitoring system in close cooperation with ILO. The new child labor monitoring system will capture all services provided to the child beneficiaries, as well as their families. Services include educational services and provision of daily meals and THRs. The project is working to have a web-based application so that nongovernmental organizations (NGOs), partners, as well as CDAs would be able to access the application online from their work places. The application will have various security levels and will facilitate information display. This will be especially useful for program staff as they monitor the progress of implementation and outcomes.

and complete his preparatory (equivalent to lower secondary) and secondary education (equivalent to upper secondary) in the formal education system.

Families of eligible children receive support to enroll their children (e.g., funds to cover the cost of birth certificates/national identification numbers, as needed) and assistance with other formal procedures, such as completing enrollment information. CWCLP covers school-related expenses such as registration fees, school supplies, and in-school snacks to all children enrolled in a CS, for an estimated value of EGP 415.00 per student/year if the child is in a formal school, and around EGP 300.00 if the child is in a community school (school fees not necessary).

Once an eligible community has been identified, CWCLP program staff negotiates with the community to find donated space for a classroom—often not in a school building per se, but rather in a commercial or residential space—and trains local residents, who are typically not education professionals, to serve as teachers. The recruitment of teachers is done by the directorates of education in targeted governorates. Hired teachers are university graduates whose degrees qualify them to work as teachers, according to the standards of the MOE. Hired teachers are provided with ongoing capacity building on how to utilize active learning techniques and methodologies.

In order to remain open, a CS needs to enroll a minimum of 25–30 children.[9] CS operate all year-round, with a 6-day week (Saturday through Thursday) during the regular school season (September to June), and a 3-day week during the school vacation season (July and August). CS are available for children between 6 and 13 years of age, although the sample of study participants only included children 6 to 11, as the program runs for 2 years and those children aged 12 and 13 would age out of the program.

Take-Home Rations

CWCLP also provides students and their families with incentives to send them to school through a THR program consisting of provisions of rice, wheat flour, and oil. The estimated value of the THR is around EGP 75.00 (USD 11.00). Adding school expenses covered by the project, equivalent to around EGP 30.00, the total package would add up to approximately EGP 105.00, or about 25 percent of an average child's monthly wage (typically around 400.00 EGP).

The THR is offered to all CS students, conditional on 80 percent CS attendance per month. Verification of attendance is done by teachers, and lists are kept in each CS and are verified by the Community Development Association (CDA) responsible for the operation of the school. Non-governmental Organizations (NGOs) and the WFP monitoring and evaluation team conduct monitoring visits to check children's attendance and compare it to distribution lists of THRs on a monthly basis. THRs are distributed directly to the child's family.

[9] One village that was originally assigned to the treatment group in El Menia governorate had to be moved to the control group because not enough eligible children could be identified. The effects of treatment non-compliance are addressed in the Treatment on the Treated results section (3.3.2.).

Income-Generating Activities

Project partners work with local communities and organizations to select highly vulnerable women from among the families of eligible children who met the criteria above. The main criteria for selection was that the woman was the mother of one of the project-targeted children. More specifically, the following profiles were prioritized:

- Low income mothers of children at risk;
- Mothers who have the will and desire to start an IGA;
- Mothers who have interest and are committed to attend training;
- Mothers who have a preliminary idea for an IGA that they wish to implement;
- Mothers who have skills (even primary skills) to run an economic activity; and
- Mothers who have the physical and mental capacity to lead a small project.

With technical support from CWCLP, partnering NGOs and local business associations are expanding existing microfinance schemes that can then provide these women with access to small business loans. These women are also provided with training on relevant IGAs implemented by partner NGOs, assisted by WFP and ILO, and in cooperation with MOMM. The credit and training combined provide an opportunity for IGAs that will reduce the need for school-aged children to work. According to the midterm implementation evaluation, 551 projects had started as a direct result of the project trainings, including a variety of microbusinesses in the agriculture, manufacturing, retail, and service sectors. It was, however, too early to know at that time the success rate of these projects and the income that they might be generating for the beneficiaries.

Awareness Raising

Community activities target parents, children, formal and informal leaders, and employers through messaging developed at appropriate levels. The messages are thematic and include the value of education, children's rights, and strategies to offset any loss of income for vulnerable families and to promote the CS activities. These messages were distributed during mass community meetings and AR events.

2.4. OUTCOMES

It is hypothesized that if children had the option to go to a CS (that is, a local school is put in place where one did not exist before) and if the family is provided with financial and social support to offset the opportunity cost of children attending school rather than working, then there would be an overall increase in school enrollment and school attendance. As school enrollment and school attendance increases, child participation in agricultural work, especially hazardous work, will decrease.

The CS component is being implemented along with THRs. This is because CWCLP partners are hoping to increase school enrollment and school attendance by providing children and families with a supplement (THR) to mitigate the risk that sending children to school rather than to work might represent to families' well-being and health. Clearly, these two components have a crucial

link to one another. Furthermore, because of the importance of THR on the enrollment/attendance of children into CS, these two interventions were not withheld from any child. As such, the impact evaluation does not examine the impact of each intervention individually, rather, their combined effects, along with those of the IGA and AR components.

Based on these assumptions, two central null hypotheses are tested, one related to the treatment effect on school-related outcomes, the other on child labor. Comprehensive data collection before and during the program period allows us to quantify the effect in a variety of areas related to these two central hypotheses. The specific study hypotheses assessed in the final analysis are presented next, grouped by subject.

2.4.1. SCHOOLING

This evaluation was designed to test the central null hypothesis that

1. Implementation of the treatment does not increase school enrollment in the 2011–2012 school year.[10] School enrollment is measured as an indicator that the child attended either a government, community, or other type of school at some point during the 2011–2012 school year.[11] Treatment is designed to increase school enrollment by providing access to schools where they did not previously exist and by incentivizing school attendance through THRs.

In addition to this central hypothesis, the evaluation collected information to address several key questions beyond the impact of the program on student enrollment. This information is used to test the null that the treatment had no effect on the following schooling-related outcomes:

1.1 Ever school attendance, measured as currently attending, attended school in the 2011–2012 school year, or attended ever;
1.2 Current school attendance, measured as self-reported current attendance at the time of the interview (2012–2013 school year);
1.3 Attendance to a community school, measured as self-reported current attendance to a community school at the time of the interview (2012–2013 school year); and
1.4 Time spent on school-related activities, measured as minutes spent going and returning from school, attending school, and study/homework in the last week (2012–2013 school year).

It is hypothesized that all these outcomes will increase as a consequence of increased school enrollment.

[10] It should be noted that the primary objective of the initiative is to decrease hazardous and/or exploitative child labor. School enrollment is an intermediate step to accomplishing this goal.

[11] This report assumes that ever school attendance in a year and school enrollment are equivalent. If a child has attended at least once, it is assumed that the child has enrolled. However ever-attendance in a year and enrollment can be different things in practice, especially when schooling is not costly to the subject. Children may enroll in school but never actually attend. This project was not able to match data on enrollment from administrative school records, so this assumption cannot be tested.

The operational definition of all school-related outcomes is presented in Appendix A. For all outcomes, data were collected from children. In the case of time allocation variables, which children, particularly younger ones, may find difficult to report, household reports are also included for validation purposes. In total, the null of no program impact is tested for six school-related outcomes.

2.4.2. CHILD LABOR

Additionally, this evaluation was designed to test the central null hypothesis that

2. Implementation of the treatment does not decrease child participation in hazardous child labor. Hazardous child labor is measured as self-reported exposure to a list of 10 workplace hazards. Treatment is designed to increase school enrollment and school attendance. It is expected that the impact of the treatment would largely come from shifting time children spend in other activities such as economic activities and unpaid household services to education. As time spent in economic activities decreases, this should in turn decrease exposure of children to workplace hazards.

In addition to this central hypothesis, the study also tests the null that the treatment had no effect on the following work-related outcomes:

2.1 Work-related injuries, measured as self-reported occurrence of any work-related injuries from a list of 10 types of injuries;
2.2 Participation in economic activities. This was measured as performing any economic activity in the last 12 months;
2.3 Time spent on economic activities, measured as minutes spent on work-related activities in the last week; and
2.4 Time spent on unpaid household services, measured as minutes spent on household chores (e.g., cooking, shopping, washing, dressmaking, etc.) in the last week.

It is hypothesized that all these outcomes will decrease as a consequence of increased school enrollment.

The operational definitions of all work-related outcomes are presented in Appendix A. For all outcomes, data were collected from children by default. In the case of time allocation variables, household reports are also included for validation purposes. Standard data collection methodologies on child labor issues (e.g., ILO Statistical Information and Monitoring Programme on Child Labour program) collect information by interviewing both the child and a responsible adult. There is no consensus regarding which of the sources is more reliable (adults or children), although our experience with qualitative and quantitative child labor data collection indicates that children are reliable informants on issues related with their direct, concrete experiences, such as their school and work-related activities, whereas adults may be more reliable informants on issues involving complex estimations, such as time allocation. This is why this study includes household and child reports for time allocation outcomes. In total, the null of no program impact is tested for nine work-related outcomes. Of these, six are specific to economic activities. Considering the limited number of outcomes, and the fact that most reflect different principles, no significance correction was made for the test of multiple hypotheses.

2.5. SAMPLE SIZE

The overall sample size for this evaluation was determined based on a power analysis conducted during the design phase of the study. Part of that analysis is presented here, fixing those parameters that have been determined during the course of study design and allowing others to vary.

During the evaluation design process, WFP provided a sampling frame of 180 hamlets or villages in the five governorates where the CWCLP project works that were eligible to receive a community school. In order to be eligible, a community had to have children that were working or at risk of working in exploitive child labor activities, and there must not have been an existing primary school accessible by some or all of the community's population. Based on an initial power analysis and consideration of the available resources, a target sample size of approximately 120 villages was initially proposed. Four villages were excluded from randomization because they were found, during data collection, to have an insufficient number of children eligible for a community school. Of the remaining 116 villages, 80 were randomly selected into the intervention group and 36 into the control group.

The following analysis presents estimated minimum detectable effect size (MDES) for the sample-based program impact estimate b_0 for a cluster randomization design using the formula proposed by Bloom (2005):

$$MDES(b_0) = \frac{M_{J-2}}{\sqrt{J}} \sqrt{\rho \frac{1-\rho}{n}} \sqrt{\frac{1}{P(1-P)}}$$

Here "J" represents the number of clusters and "n" the average number of individuals per cluster. "P" is the proportion of clusters assigned to the intervention condition. In this way, the formula allows for adjustment for unequal sample allocation ratios. Although balanced samples are optimal for maximizing power, the loss of power associated with moderately unbalanced allocation ratios is not severe, and the advantages of being able to include a larger number of villages in the intervention group are substantial.

M_{J-2} is calculated as a sum of t statistics, corresponding to the critical t value for the chosen level of significance α and the t value corresponding to the desired statistical power 1-β. For a two-tailed test this is:

$$M_{J-2} = t_{\alpha/2} + t_{-1\beta}$$

The intra-cluster correlation (ICC, or ρ) is a critical parameter that describes the relative variance in outcomes across hamlets versus within hamlets. This is important because, while randomization was done by cluster (that is to say, by village/hamlet), outcomes are reported at the individual level. The larger the ICC, the more similar children's responses are to the project within each hamlet, and consequently, the less new information is provided from sampling more children within each hamlet. ICC also encapsulates variability in responses to intervention across hamlets due to factors such as differences in crop cycles, local culture and gender roles, and

modes of production. When ICC is high, it is necessary to include more hamlets in the study in order to get enough information to tell whether the project had an impact. An ICC can range from zero to one. Higher values correspond to greater similarity within hamlets in response to an intervention. In educational and health impact evaluation studies conducted primarily in the United States and other developed countries, ICCs typically range between 0.10 and 0.25, indicating that there is a significant amount of variance in outcomes for students who receive the same intervention (Bloom, 2005).

A set of basic hierarchical models were developed using HLM version 7 software[12] to estimate ρ in order to lay the foundation for future analyses. All models are "empty," in the sense that only the outcome variables were specified, with no predictor variables. These empty models are estimated in order to generate variance component estimates for levels 1 and 2 in the case of the normal models, and for level 1 in the case of the Bernoulli models (in the Bernoulli models, error variance is treated as fixed at $\frac{\pi^2}{3}$).[13] These estimates are used to calculate the intra-cluster correlation for each outcome. The results are presented in Table 1. Estimated ρ ranges between .11 and .28.

Table 1. Estimated Intra-Cluster Correlation ρ

Models with all children	Model type	Estimated variance component		ρ
		Level 1 (child) n = 2,705	Level 2 (village) n = 116	
Ever school	Bernoulli	1.22	3.29*	0.27
Ever work	Bernoulli	0.69	3.29	0.17
Attended school last year	Bernoulli	1.13	3.29	0.26
Work time	Normal	50.17	381.18	0.12
Chore time	Normal	12.77	98.52	0.11
Models with working children only	Model type	Level 1 (child) n = 2,286	Level 2 (village) n = 116	ρ
Exposure	Bernoulli	1.25	3.29	0.28
Injury	Bernoulli	1.08	3.29	0.25
Log time	Normal	0.08	0.53	0.14
*For Bernoulli models, error variance is treated as fixed at $\pi^2/3$.				

Using the MDES formula above, Table 2 shows MDES under three scenarios (with the range of ρ values expected), fixing the following parameters to

$$J = 116,\ n = 23,\ P = 0.69,\ \alpha = 0.05,\ 1-\beta = 0.80$$

[12] Stephen Raudenbush, Tony Bryk, & Richard Congdon, 2010. HLM 7 Hierarchical Linear and Nonlinear Modeling. Scientific Software International, Inc.

[13] Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE publications Ltd.

Under these assumptions, for the lowest expected values of the ICC, MDES = 0.21, a small effect according to Cohen's conventions. With higher values of ICC, an MDES = 0.33 should be detectable.

Table 2. Intra-cluster Correlation (ρ) and MDES

ρ	MDES (Standard Deviations)	MDES (Percentage Points)
0.10	0.21	10.5
0.20	0.27	13.5
0.30	0.33	16.5

2.6. RANDOMIZATION

Randomization was conducted by ICF. The initial sampling frame of 180 communities was used to select randomly a subset of villages or hamlets to be included in the evaluation study. Based on an initial power analysis and consideration of the available resources, a target sample size of approximately 120 villages was initially proposed. In subsequent discussions, the WFP decided that it had sufficient resources to provide CS in 80 villages that were meeting the target to be included in the project document and in the signed cooperative agreement between WFP and USDOL. In order to serve as many children as possible and reach as many villages or hamlets as possible, rather than dividing the sample group in half (which would optimize power), ICF planned to assign 80 villages into the intervention group and 40 villages into the control group.

Subsequent investigation by the survey team showed that four villages in the sampling frame were, in fact, not eligible to receive a CS because there was an existing government school there, or because the number of eligible children was too low to meet the minimum enrollment required by WFP to start a new community school.

As a result, 116 villages were ultimately selected for inclusion in the study. After the baseline data collection was completed, ICF randomized the villages and hamlets into intervention and control groups, with 80 randomly selected into the intervention group and 36 into the control group.

2.6.1. SEQUENCE GENERATION

A balanced randomization process was used to generate the allocation sequence, with stratification by blocks. Blocks consisted of governorates. Within each governorate, villages were sorted by a random computer-generated number, and roughly the first 69 percent were selected into the intervention group, with the remainder being the control group. The randomization was done using standard commercial spreadsheet software.

2.6.2. ALLOCATION CONCEALMENT

Proper randomization depends on the generation of an unpredictable randomized allocation sequence and concealment of that sequence until assignment occurs. Allocation concealment is defined as the "strict implementation of a random allocation sequence without foreknowledge of treatment assignments. Allocation concealment refers to the technique used to implement the sequence, not to generate it" (Schulz and Grimes, 2002, p. 614). Allocation concealment should not be confused with blinding of study subjects. Blinding, which is discussed later, refers to the subject and/or the researcher being unaware of the subject's treatment assignment. Allocation concealment impedes those who admit participants to a randomized controlled trial from knowing their upcoming assignments.

In the current study, randomization was done after baseline data were collected and eligible children had been identified, so the allocation sequence was concealed from program implementers and eligible subjects before admission into the program. Further, randomization was done by ICF using a random number generator, which precluded study participants from deciphering their assignment before it occurred.

2.6.3. IMPLEMENTATION

Immediately after randomization, ICF provided WFP with a list of villages that would participate in CWCLP and those that would constitute the control group. WFP began the implementation process as soon as it was provided this list on November 11, 2011. Prior to the follow-up study, ICF worked with USDOL and WFP to understand the timeline and process of implementation of the various activities.

Eligible children in the villages assigned to the intervention group were offered to attend CS and receive THRs. Based on the fact that not all children initially identified for each class met the eligibility criteria when assessed for full eligibility, classrooms needed to be augmented with newly enrolled children partway through the intervention period in order to maintain full classes. This was done so that the CS met the necessary criteria of enrolling a minimum of 25–30 children in order to remain open. The additional children who are added will participate fully in the interventions, but were not included in the study.

While any subsequently enrolled children were not included in the current evaluation, follow-up data for those children who declined to participate or who dropped out was collected in order to support intention to treat analyses.

2.6.4. BLINDING

Since randomization of the villages occurred after baseline data collection, enumerators and study participants were blinded to intervention assignments while conducting the baseline survey. Blinding was not maintained during the endline survey, as enumerators needed to know the status of the village for appropriate questionnaire routing. Study participants were also obviously aware of their intervention assignment as they may either be enrolled in a CS or aware of the program in any other way.

2.6.5. VALIDATION OF RANDOMIZATION

Data from the baseline survey were used to analyze equivalence between intervention and control children in gender, age, school attendance, and time spent on study. Both groups were also compared on work-related outcomes including participation in economic activities, exposure to hazardous working conditions, time spent on economic activities, and time spent on unpaid household services. Finally, regression models were estimated to test for differences between the control and intervention groups at baseline. Since the intervention and control groups were randomly selected, all differences across the groups must be due to chance. Significance tests are designed to assess the likelihood that these differences are due to chance, so they are in theory redundant. Nonetheless, significance tests are included to validate the randomization process and to help the reader identify differences between the treatment and control groups that are unexpectedly large relative to sample variability.

Descriptive statistics for the participants at the beginning of the study are presented in Table 3 below. There was a slightly higher proportion of girls in the intervention group (55 percent) compared to the control group (50 percent). The children in the intervention and control groups were the same age, on average 8.9 years old. Less than one-third of the children in both the intervention and control groups had ever attended school (28 and 32 percent, respectively). Most children (78 percent in the treatment group and 76 percent in the control group) reported that they had worked. Only a small percentage of children in both groups attended school during the 2010–2011 school year (9 percent in the treatment group and 8 percent in the control group).

The study finds that children in the intervention group report spending less time in economic activities, on average, than the control group children. Children in the treatment group reported spending 1,545 minutes per week on average, or nearly 26 hours, on economic activities, compared to the 1,704 minutes (28 hours) reported by the control group. Though children in the intervention group worked fewer hours, they were significantly more likely to report working in hazardous conditions (70 percent vs. 60 percent in the control group). Children in both groups appear to spend a similar amount of time on unpaid household services (around 12 hours per week).

For all time allocation variables, household reports were included to validate data from child reports. Both reports seemed to match on average, although household informants reported a higher amount of time on school and study activities per week (4 hours for both groups) than child reports.

In conclusion, the intervention and control groups were roughly equivalent for all outcomes except exposure to workplace hazards. It appears that the randomization was not completely successful in balancing the intervention and control groups along these dimensions, but it is these baseline differences that are controlled by including baseline values as covariates when estimating project impacts with endline data. In any case, since this difference was due to chance, the randomization process remains valid. Any statistical controls implemented during endline analysis are only to reduce variance in the final impact estimates and improve statistical power.

Table 3. Baseline characteristics of groups after randomization

Characteristics by treatment group and significance of group differences [Egypt, 2013]

Baseline Characteristics	Treatment			Control			*p*
	Mean/ Count	S.E. or Percent	n	Mean/ Count	S.E. or Percent	n	
Girls	536	54.6%	982	220	50.0%	440	*0.21*
Age	8.9	0.1	982	8.9	0.1	440	0.89
Ever attended school	279	28.4%	982	141	32.0%	440	0.48
Attended school last year (2010–2011)	88	9.0%	982	36	8.2%	440	0.75
Currently attending school	0	0.0%	982	0	0.0%	440	-
Currently attending community school	0	0.0%	982	0	0.0%	440	-
Minutes spent on school and study per week	0	0.0%	982	0	0.0%	440	-
Minutes spent on school and study per week (household reports)	229.6	42.8	983	235.9	46.0	440	0.92
Working in hazardous conditions	690	70.3%	982	263	59.8%	440	0.03*
Participated in economic activities	764	77.8%	982	334	75.9%	440	0.62
Minutes spent on economic activities per week	1,544.6	63.1	982	1,703.7	1,314.2	440	0.19
Minutes spent on economic activities per week during the school season (household reports)	1,557.1	67.2	983	1,686.8	97.7	440	0.28
Minutes spent on economic activities per week during the non-school season (household reports)	1,643.4	61.5	983	1,782.1	100.8	440	0.24
Weekly minutes on unpaid household services	747.8	30.7	982	743.7	44.9	440	0.94
Weekly minutes on unpaid household services during the school season (household reports)	738.3	28.3	983	693.4	44.4	440	0.40
Weekly minutes on unpaid household services during the non-school season (household reports)	750.6	28.0	983	704.0	43.7	440	0.37

*Statistically significant at p<.05
**Statistically significant at p<.01

2.7. STATISTICAL METHODS

2.7.1. DATA SCREENING

Before proceeding with any hypothesis testing, study variables were screened using univariate and bivariate analyses. First the distributional properties of the variables were examined to verify if they meet the univariate requirements for bivariate and multivariate analysis (normality, outliers, missing cases). For example, consumption data is notoriously skewed, with the majority of households clustering at the bottom of the distribution. In such cases, skewed variables can be transformed (e.g., using logs) to improve the normality assumption.

The resulting variables were examined using bivariate analysis (scatterplots, bivariate correlations, chi-square tests) to assess analytical assumptions (linearity, homoscedasticity) and to identify significant associations with the outcome variables. Data analysis includes descriptive statistics (means and standard errors) for key variables and graphics to show variable distribution and covariance.

2.7.2. HYPOTHESIS TESTING

The study hypotheses are formulated as a comparison of means or proportions between treatment and control groups at the study end line. The null hypotheses shown in section 1.3 are tested by regressing each outcome for child *i* from village *v* observed at endline, y_{iv1} on a constant (a), an indicator Z_v that takes on a value of 1 if the community was assigned as a treatment community, a vector of controls based on baseline attributes X_{iv0}, a vector of controls for baseline time allocation E_{i0}, and a mean zero error term that is corrected for arbitrary heteroskedasticity (using a White correction), and clustered at the village-level, ε_{iv}:

$$y_{iv1} = a + \beta Z_v + CX_{iv0} + DE_{i0} + \varepsilon_{iv}$$

The use of Ordinary Least-Squares (OLS) to calculate impacts on binary outcomes violates the OLS assumption of normally distributed homoskedastic errors, particularly when outcome means are close to zero or to unity. However the value of a randomized controlled trial (RCT) lies in its simplicity and transparency. Compared to observational studies, the client is able to identify differences in the data separately from differences in modeling assumptions used by the researcher. Given that the objective of this evaluation is to compare differences in means and nothing more, our preference is to be transparent and use a simple linear regression to condition on observables. OLS regression coefficients are much more intuitive and readily interpretable than coefficients derived from logistic regression (e.g., odds ratios). This study did, in any case, conduct robustness checks on the calculations of impacts for binary outcomes using a logistic regression approach. None of the results presented in this report were sensitive to the choice of modeling approach.

The main analyses presented in this report are Intention to Treat (ITT) analyses, in which the primary predictor is the village's assignment to the intervention or control group. ITT analysis eliminates the need for imperfect estimations of impact based on pre/post comparisons. Random

assignment underlying the ITT analysis also eliminates the potential for selection bias (as long as there is assignment compliance), although take up of treatment may clearly be non-random. Analytical strategies to take compliance and take-up into account are presented next.

Compliance with the initial assignment was imperfect. Out of 115 villages, four villages crossed-over, with two moving from the control group to the treatment group, and another two moving from the treatment group to the control group. These shifts were initiated by WFP program staff in adjustment to the conditions on the ground for the establishment of program activities. These cases are instances of partial compliance, with some of the villages that were assigned to the treatment condition not getting the treatment, and vice versa.

Non-compliance may represent an analytical problem because the causal impact of the program on the outcomes of interest may be diluted. In the presence of partial compliance, both ITT and Treatment on the Treated (TOT) analyses are of interest. For program evaluation purposes, the ITT effect is the primary measure of interest, as it encapsulates the real-life impact of the intervention in the program areas, including any non-compliance problems. The TOT estimate, on the other hand, explains the direct impact of the treatment, and it is useful to evaluate the merits of the intervention, in the absence of implementation problems.

To account for partial compliance, both ITT and TOT analyses are conducted, using a modified dummy variable that encodes for the actual (versus assigned) placement of a community school in a village. This TOT estimate provides an upper bound on the impact estimate (see, for example, Avvisati et al., n.d.). The operational definition of the ITT group is straightforward: children in villages that were initially assigned to the treatment group. The operational definition of the TOT group is a little more nuanced, as it requires a definition for compliance. Compliance is naturally defined at the village level, since that is the level of randomization; thus, a village-level definition of the TOT group would include the children that were in villages where a community school was eventually established.

Alternatively, the TOT group can be defined as those children who actually took up the program. For this analysis, all students who reported participating in the program, regardless of their original assignment, were pooled together for the estimation of the treatment effect. The project was not able to obtain administrative records of program enrollment that could be linked to the impact evaluation data at the individual level, so program take-up is operationally defined using child self-reports of current attendance to a CS, irrespective of baseline assignment. This approach cannot determine whether children are also receiving a THR, which is conditional on regular attendance, or the IGA and AR components, although it is assumed that a majority are.

This report uses both definitions of TOT, including treatment on the treated villages and treatment on the treated children to show the full range of impact estimates. General TOT inflation factors in this report are used to provide an overall order of magnitude of the ITT versus TOT estimates. These overall TOT inflation factors are calculated using the Wald estimator:

$$\beta_W = \frac{E[\mathrm{Y_{iv1}} \mid Z_v = 1] - E[\mathrm{Y_{iv1}} \mid Z_v = 0]}{E[T_{\mathrm{v}} \mid Z_v = 1] - E[T_{\mathrm{v}} \mid Z_v = 0]}$$

Where the top part of the fraction, $E[\mathrm{Y_{iv1}} \mid Z_v = 1] - E[\mathrm{Y_{iv1}} \mid Z_v = 0]$ represents the ITT estimate, $\mathrm{E[T_v \mid Z_v = 1]}$ represents the percentage of units (villages or children) initially assigned to the treatment group that received the intervention, $\mathrm{E[T_v \mid Z_v = 0]}$ represents the percentage of units initially assigned to the control group that received the intervention. In order to mitigate omitted variable bias, the TOT estimates of program impacts are calculated using an instrumental variables approach called two-stage least-squares. In the first stage, the probability that the treatment is positively associated with random treatment assignment is estimated by using random assignment as an instrumental variable. In the second stage, this probability is used to estimate the treatment impact on the outcome of interest.

Finally, attrition bias analyses are conducted to elucidate any problems caused by attrition, including weighting adjustments to compensate for non-random attrition.

2.7.3. CONTROL VARIABLES

As indicated earlier, the sample size selected would allow the study to detect a medium effect size. To increase our chances to detect smaller effect sizes, additional controls are added to the regression equations to maximize power. Because the randomization was generally balanced on outcomes at baseline, these controls mostly serve to reduce variance and improve power. The effect of covariates on MDES in a cluster design is accounted for by Raudenbaush's (1997) equation:

$$MDES(b_0) = M_{J-g*-2}\sqrt{\frac{\rho(1-R_2^2)}{P(1-P)J} + \frac{(1-\rho)(1-R_1^2)}{P(1-P)Jn}}$$

Where R_1^2 and R_2^2 are the proportion of Level 1 (villages) and Level 2 (individuals) variance explained by covariates, g* is the number of group covariates used, and Jn the total number of individuals.

As Bloom (2006) notes, it is good practice to specify covariates in advance of analysis to avoid specification searching. Also, control variables are most effective when the lagged value of the outcome of interest at baseline can be controlled to minimize the effect of preexisting random variance. The best predictors of future outcomes are typically past outcomes, because past outcomes reflect most factors that determine future outcomes.

Heeding recommendations from Bloom (2006), control variables were pre-specified in the analysis plan prepared ahead of this report. Controls for each outcome of interest include all outcome variables at baseline. This is particularly relevant in the case of working in hazardous conditions, which was significantly unbalanced between treatment and control at baseline. The inclusion of working in hazardous conditions status at baseline as a control variable or covariate statistically removes the differences between the treatment and control group on this outcome. Additionally, the following vector of control variables was included:

- Age at baseline, expressed as a set of dummies;
- Sex;

- Sex by age interaction, expressed as a set of dummies;
- Age stopped school, as a vector of dummy variables indicating time since child last attended school interacted with the child's age, both at baseline;
- Household head's education at baseline;
- Household head's occupation at baseline;
- Subject's relationship to household head at baseline;
- Counts for number of elder dependents (50 and older) at baseline;
- Counts for number of young dependents (0–14) at baseline;
- Counts for number of prime-age adults (15–49) at baseline;
- Household size at baseline;
- Minutes spent with family or friends in rest or playing per week at baseline;
- Minutes spent caring for other family members like elder or children per week at baseline;
- Minutes spent on personal activities like reading, watching TV, etc., per week at baseline; and
- Minutes spent on sleep and relaxation per week at baseline.

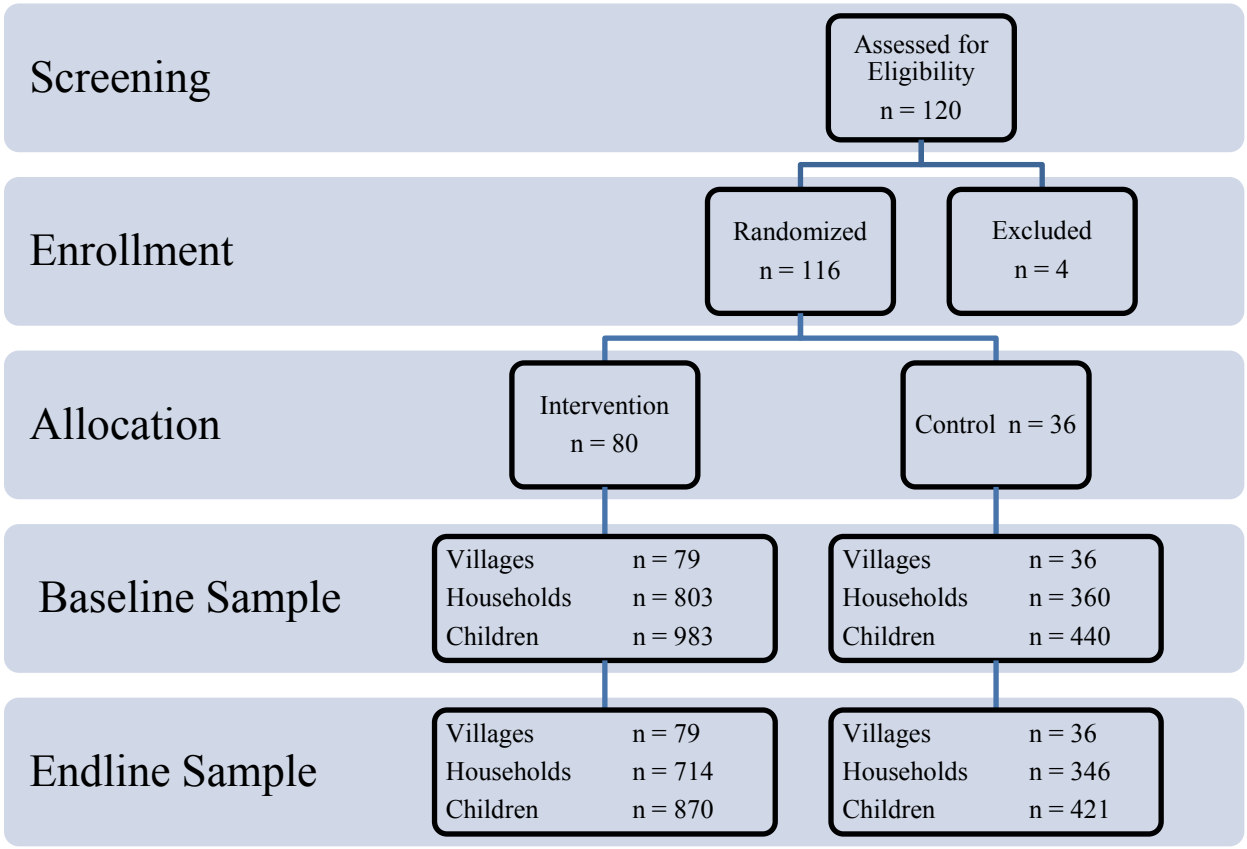
Finally, governorate was included as an additional control variable since it was used as a blocking variable during sample selection. Governorate is controlled for with a vector of dummy variables that take a value of 1 if a child was observed in a given governorate at baseline and 0 otherwise.

3. RESULTS

3.1. PARTICIPANT FLOW

The flow of participants through the various stages of the study and the allocation of villages are shown in Figure 1. An initial sampling frame of 180 villages was selected as eligible for randomization, of which 120 villages were randomly selected. Of those, 116 were subsequently randomized (Figure 1). Of the 116 villages, 80 were randomly assigned to the intervention group and 36 to the control. Four villages were excluded from randomization because they were found, during data collection, to have an insufficient number of children eligible for a community school. Table 12 (see Appendix B) shows the flow of individuals in the evaluation, and the corresponding response rates at each stage.

Figure 1: Randomization of Villages and Final Baseline and Endline Samples



3.2. RECRUITMENT

Prior to the fielding of the baseline survey in October 2011, Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt (CWCLP) partners developed the "Child labor in the Agriculture sector survey in Egypt." Using this information, CWCLP contracted the Cairo Demographic Center to

develop a Rapid Assessment (RA) to identify the areas (villages) where there were a high number of cases of children laboring in the agricultural sector. Based on the RA report's findings, CWCLP partners provided ICF International (ICF) with a list of villages and hamlets where the program components would be implemented and the study would take place. Once the list was translated, it was provided to ICF to use for randomization purposes.

In order for the baseline survey to be carried out and ensure that the program was implemented in a timely manner, households needed to be identified that might have eligible children to participate in the CWCLP program. As described above, in order for a child to be considered eligible, the child needed to meet the following criteria:

- Child between 6 and 11 years old;
- Child not currently enrolled in a national government school;
- Absent from school for at least the last 2 years; and
- Child engaged in or at risk of exploitative child work in agriculture. Children at risk of exploitative child work were defined as the siblings of child laborers under age 15 and living in the same household.

ICF and the U.S. Department of Labor (USDOL) proposed a strategy to have the local data collection firm El Zanaty and Associates, who was conducting the baseline survey, first conduct a screening to identify eligible households and children. The World Food Programme (WFP) accepted this recommendation and offered to have their local nongovernmental organizations (NGOs) work with the firm to help identify households and children, in a targeted way, by using their knowledge of local community members. As such, El Zanaty and Associates worked with the local NGOs to find potential households and children. If the household met the screening criteria, it would be included in the study roster. Data collection staff visited households in each village until they identified a maximum of 30 eligible children (the maximum number of children that could be accommodated by any individual community school), at which point they stopped and moved on to other villages.[14]

Once 30 eligible children in each village were located, the enumerators from the El Zanaty group conducted the baseline survey with the head of household and with the children who were eligible to participate. It was decided that each household would have a maximum of two children who could participate in the study, while the other children could be enrolled in the project without participating in the study. Therefore, if a single household had more than two eligible children, the enumerator drew names randomly for those who could participate in the study. Specifically, they listed the names of all the eligible children in a household on individual slips of paper and then randomly drew two names out of a hat to determine which children would participate in the study.

Unfortunately, there was a misunderstanding on the part of the data collection firm regarding the eligible age of the children. Their understanding was that children were eligible up to the age of 13, because the Community Schools (CS) are available to children up to the age of 13. However, the screening should only have included children up to age 11, because the program runs for 2

[14] One ICF staff member visited Egypt and went to the field to monitor training and data collection conducted by El Zanaty and Associates.

years and children above the age of 11 would age out of the program. Due to this misunderstanding, approximately 37 percent of the children ultimately were not eligible.[15] Therefore, once the baseline data were collected, WFP revisited the villages where the ineligible children were and selected new children to attend the schools and participate in the program. The new children, however, were not a part of the study, in order to ensure consistency across study participants.

3.2.1. BASELINE SAMPLE

Data collection for the baseline survey occurred between October and November 2011. From each village an average of 12 eligible children (between 6 and 11 years of age, not attending school at baseline) were identified and included in the study, representing about 1 percent of youth in the villages. Fourteen eligible families refused to participate, citing guidance from their religious leaders as reason not to participate in interviews. A total sample of 1,423 eligible children participated in the interview (including complete and partially complete surveys). A total sample of 1,163 surveys in households with eligible children was completed in 115 communities (There were no eligible children in one treatment community, which had to be excluded).

3.2.2. ENDLINE SAMPLE

Data collection for the endline survey occurred between March and April 2013.[16] Data collected during both rounds included information on household composition, education and work status of household members, time allocation of children in the household, household attitudes, and consumption. Child interviews collected information on school and work status, exposure to workplace hazards and injuries, and time allocation.

A total of 1,060 household surveys were completed, including 1,291 children that were also interviewed at baseline, out of an initial 1,423 correctly eligible children at baseline, representing a recapture rate of 91 percent, and an average of 11 respondents per village.

The recapture rate includes all individuals that could be interviewed at endline, and excludes all non-respondents. This inability to collect data for all individuals in the baseline sample can be defined as study attrition, and it represents two threats to our analysis, including diminished power and attrition bias.

[15] A detailed report was prepared by the project and sent to USDOL and ICF in May 2012, explaining the different reasons for exclusion of children. While the reduction in the number of children presents a challenge, as the remainder of the report will demonstrate, the impact evaluation design chosen was a cluster randomized control trial. Therefore, the number of children within the cluster is less relevant than the number of clusters in determining effect sizes. When ICF learned of the change in cluster size, it re-ran the power analyses using the new cluster size and determined the sample size was still sufficient to show moderate effects.

[16] It would have been optimal to conduct follow-up data collection during the same time of the year as baseline (due to issues of seasonality discussed later in the report). However, for administrative reasons, this was not possible.

Diminished Power

Attrition and non-response means that the final effective sample is smaller than initially anticipated. The number of individuals per cluster has, however, a smaller impact on power and minimum detectable effect size (MDES) than the number of clusters (Bloom, 2006). Since all but one of the clusters initially included in the study remained at follow-up, attrition should not have a large impact on the final power achieved. For example, using the examples from Table 1, Table 4 shows the MDES for different ρ values, with an average of n = 23 individuals per cluster (baseline scenario) versus the actual endline scenario of n = 11 individuals per cluster. As the table shows, the MDES are only slightly higher in spite of the substantial decrease in the number of individuals per cluster.

Table 4. Intra-cluster Correlation (ρ) and MDES

ρ	MDES (n = 23)		MDES (n = 11)	
	Standard Deviations	Percentage Points	Standard Deviations	Percentage Points
0.10	0.21	10.5	0.24	12.0
0.20	0.27	13.5	0.30	15.0
0.30	0.33	16.5	0.40	20.0

Attrition Bias

When study attrition is correlated with the outcomes of interest, program impact may be under or overestimated. For example, if children interviewed at follow-up are more likely to be enrolled in school than those that could not be interviewed, then the study would overestimate the impact of the program on enrollment rates. Since there is no follow-up data on attritors, it is difficult to estimate attrition bias. However, it is possible to test for any significant effects of treatment as a predictor of attrition, controlling for observable differences used in the regressions. This test should provide an indication of the potential for attrition bias, by identifying non-random differences in attrition by treatment status. This would indicate that the effects of attrition are different for the treatment and control groups. A model with experimental assignment as predictor, controlling for observables, did find a significant effect for experimental assignment (β = 0.08, Robust S.E. = 0.18, p = 0.00), with age, age stopped school, head of household's occupation, education, and relationship as significant covariates.

An examination of baseline mean-differences between non-attritors and attritors (see Table 5) shows that attritors in the treatment group are significantly younger, less likely to have attended school last year and less likely to have participated in economic activities. It is thus likely that attrition may have an impact on average treatment effects.

Table 5. Baseline characteristics of groups by attrition status

Characteristics by treatment group and significance of group differences [Egypt, 2013]

	Treatment				Control			
Baseline Characteristics	Total	Non-attritor	Attritor	*p*	Total	Non-attritor	Attritor	*p*
Age	8.9	9.0	8.0	.00**	8.9	8.9	8.9	0.9
Girls	55%	56%	47%	0.14	50%	50%	58%	0.4
Ever attended school	28%	29%	23%	0.19	32%	32%	32%	1
Attended school last year (2010–2011)	9%	10%	4%	.01*	8%	8%	11%	0.8
Working in hazardous conditions	70%	71%	62%	0.08	60%	59%	68%	0.5
Injured at work	50%	51%	47%	0.47	45%	45%	53%	0.6
Participated in economic activities	78%	79%	70%	.04*	76%	76%	79%	0.8
Number of subjects	983	870	113		440	421	19	

*Statistically significant at p<.05

**Statistically significant at p<.01

To compensate for this bias, non-attritors can be reweighted so that on average they have the same baseline distribution as the original sample. This can be done using a stepwise weight adjustment known as iterative proportional fitting or raking to achieve known population margins. Raking weights were computed using Bergmann's Stata IPFWEIGHT module. Weights were computed separately for treatment and control groups to adjust for any difference on the baseline characteristics shown in Table 5 between the non-attritor group and the total group. The resulting weight adjustments are accurate (with a deviation tolerance of one centesimal point) and relatively modest, with a minimum weight of 0.89 and a maximum of 1.28. All endline analyses are conducted using these raking weights.

3.3. ENDLINE RESULTS

The study results shown in the following sections present the endline means and standard errors (clustered at the village level) for the six school-related outcomes and the nine work-related outcomes by treatment status. All outcome variables are by default collected from child reports, unless "household reports" are specified.

Additionally, each table shows the Ordinary Least-Squares regression adjusted mean difference between the program and control groups (β), the standard error (S.E.) of this coefficient, and the statistical significance (*p*) of the null that a treatment group did not differ from the control group. Based on the preliminary analyses presented in Section 2.5 and Section 3.3, this report assumes that a valid randomization of study subjects was conducted. If the difference between control subjects and treatment subjects is causal and the control mean is a counterfactual for treatment subjects (two conditions that follow from a valid randomization), all group differences can be considered causal. The β coefficient can thus be interpreted as the causal program impact on each outcome of interest, and is discussed in the remaining of the report as such.

Finally, the results are presented first for the Intention to Treat (ITT) estimates for both school and work-related outcomes, followed by the village-level Treatment on the Treated (TOT) estimates and the child-level TOT estimates.

3.3.1. INTENTION TO TREAT RESULTS

Table 6 shows the ITT results for the six school-related outcomes at endline. The null of no program impact is widely rejected for all school-related outcomes. Beginning with school enrollment, 14.5 percent of eligible children in control villages had enrolled in school during the 2011–2012 year, compared to 62.3 percent among eligible children in treatment villages. Conditioning on observables, this represents a significant causal impact equivalent to an increase in school enrollment of 46.4 percentage points (pp).

Results for the three school attendance outcomes are also consistent with a large treatment effect. School attendance was significantly higher for the treatment group (86.4 percent) than the control group (49.9 percent), with an estimated program impact of 38.8 pp. Regarding current school attendance, 23.5 percent of control subjects were attending school at the time of the interview (March–April 2013), compared to 71.2 percent among treatment subjects,[17] representing an impact equivalent to an increase in current school attendance of 48.8 pp.

The largest school-related impact was, more specifically, on current attendance to a CS, which is consistent with the program primary target of implementing CS in communities with no existing primary school. Very few (3.8 percent) control subjects reported attending a CS at the time of the interview,[18] compared to more than half (58.0 percent) of treatment subjects, representing an impact equivalent to an increase in current CS attendance of 53.9 pp.

In addition to school enrollment and school attendance, the program also shows a significant impact on the time children spend in school-related activities, such as going and returning from school, studying, and doing homework. Subjects in the treatment group report spending a net average of 1,443 minutes per week (about 24 hours) on school-related activities, compared to the 511 minutes (less than 9 hours) reported by children in the control group. The lower average minutes by the control group is driven by the fact that many fewer children in the control group are attending school. These children are included in the computation of averages as spending 0 minutes in school-related activities. Conditioning on observables, and irrespective of current school attendance status, village assignment to the program increases the time children spend on school-related activities by 951 minutes (16 hours) per week. This finding is consistent with household reports of the time children spent on school-related activities, although when household reports are considered, the impact of the program is even greater, at 1,092 minutes per week (18 hours).

[17] It would seem like school enrollment should be mechanically greater than current school attendance. However, that is not necessarily true based on the definitions used in this evaluation, where school enrollment is measured for the school year 2011–2012, whereas school attendance refers to the school year 2012–2013.

[18] No control subjects were expected to be attending a CS at endline. It is thus possible that the 3.8 percent of control children that reported attending a CS represent either respondent error or crossover of study subjects. If either is true, the impact of the program is potentially larger than reported, although the difference is arguably small enough that it can be safely ignored. According to CWCLP program staff, it is most likely that this represents mis-reporting, as they exert strict controls over the roster of students eligible to attend a CS.

An interesting finding that can be gleaned from Table 6 is that subjects enrolled in school in the treatment population appear to spend less time in school (1,443.2 minutes / 62.3 percent enrolled = 2,317 minutes among the enrolled) than subjects enrolled in school in the control population (510.5 minutes / 14.5 percent = 3,520.7 minutes). The difference between the two groups would amount to more than 3 additional hours spent on school-related activities per day (including school attendance, commuting to school, and time spent studying or doing homework at home). This implies that subjects induced into school by the treatment are on average spending less time in school than those that would seek out schooling absent the treatment. A possible explanation for this finding is that children in control villages (which had no school to begin with and did not receive a CS as part of their experimental assignment) are spending more time going and returning from school, as they must necessarily be attending a school outside their village. Time spent on school-related activities was not disaggregated to this level, so the extra commuting time hypothesis is not testable with the current data.

Subgroup analysis for all school and work-related outcomes by sex, age, and work status at baseline is presented in Appendix C. For simplicity, these tables omit the outcome means by assignment group to focus on the size of the program impact and its significance. Subgroup analysis indicates that the school-related impacts discussed above were significant for all subgroups. However, impacts on school-related outcomes were larger for girls, for younger children (6–8 years old), and for children that were not involved in economic activities at baseline.

Table 6. School-related outcomes at endline by treatment status (ITT)

Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E.	β	S.E.	*p*
School enrollment (2011–2012, %)	14.5	2.7	62.3	3.9	46.4	4.8	.00**
Current school attendance (2012–2013, %)	23.5	2.7	71.2	3.0	48.8	3.3	.00**
Current community school attendance (2012–2013, %)	3.8	1.7	58.0	3.6	53.9	3.7	.00**
Ever school attendance (%)	49.9	4.2	86.4	2.1	38.8	3.3	.00**
Time on school-related activities	510.5	63.6	1443.2	67.0	950.9	79.6	.00**
Time on school-related activities (household reports)	390.9	73.4	1518.5	99.1	1092.5	119.2	.00**
Sample Size	418		866		1284		

*Statistically significant at p<.05
**Statistically significant at p<.01

It is hypothesized that the additional time spent on school-related activities will result in a reduction of the time spent in economic activities, and that child participation in economic

activities in general, and in hazardous child labor in particular, will decrease as a result. Table 7 shows the ITT results for the nine work-related outcomes at endline. In contrast to school-related outcomes, results for work-related outcomes are mixed.

In the case of involvement in economic activities in general, as well as hazardous child labor and prevalence of work-related injuries, both control and treatment subjects show similar levels, with a regression adjusted mean difference not significantly different from zero. The study therefore fails to reject the null that village assignment has no impact on overall involvement in economic activities, hazardous child labor, and work-related injuries. Subgroup analysis did not identify any significant impact for the sex, age, and work status groups analyzed.

To put these results in context, it is worth mentioning that overall levels of involvement in economic activities, hazardous child labor, and work-related injuries are sharply down from their baseline values for both groups. In the absence of group differences at endline, it is not possible to assume program causality, so any analysis of this sharp reduction in the participation in economic activities can only be speculative. This is discussed in greater depth in the context of seasonality effects in section 4.1.

Analysis of time allocation variables shows a more nuanced picture of program impacts on work-related outcomes. Children in control villages, irrespective of activity status, report spending an average of 676 minutes (11 hours) per week in economic activities, compared to 474 minutes (8 hours) per week among children in treatment villages. Conditioning on observables, this represents a significant impact equivalent to a reduction in time spent on economic activities of 195 minutes (3 hours) per week.

This reduction is corroborated by household reports, which also identify a significant impact of village assignment on time spent on economic activities of 211 minutes (3½ hours) per week during the school season. Subgroup analysis (see Appendix C) shows that the reduction in time spent on economic activities was larger for boys, with an estimated reduction of 223 minutes per week (nearly 4 hours), and a reduction of 275 minutes during the school season (4½ hours). The different results by sex are probably explained by the fact that unpaid economic activities are much more prevalent among boys, whom spend roughly twice as much time on economic activities than boys (740 minutes per week for boys in either the treatment or control group, versus 363 minutes for girls). The reduction in time spent on economic activities was also larger for children who were involved in economic activities at baseline and for older children, although in the latter case none of the group impacts were significant.

Household informants were further asked about time allocation during the school off-season.[19] In this case, the impact of program assignment on time spent on work activities was not significantly different from zero ($\beta = -93$, $p = .39$). This was also the case with all subgroups analyzed, reinforcing the hypothesis that time spent in school-related activities is shifted from time spent in economic activities.

[19] The full-time school season goes from September to June. During this season, students are expected to attend school 6 days per week. In the case of CS, students were expected to attend school part-time (3 days per week) during the off-season (July and August). High demand for labor is most likely during the agricultural harvest season, which is most intensive between October and November. See section 4.1. for further discussion on seasonality.

Unpaid household services are not considered an economic activity, and yet they may represent a significant burden for the child and add on to the negative impact of economic activities on children's welfare opportunities. Children in control villages, irrespective of activity status, report spending an average of 930 minutes (more than 15 hours) per week in unpaid household services, compared to 811 minutes (about 14 hours) per week among children in treatment villages. Conditioning on observables, this represents a significant impact equivalent to a reduction in time spent on unpaid household services of 135 minutes (2 hours) per week. The impact estimates on unpaid household services are validated by household reports, which identify a significant impact of 137 minutes per week during the school season. As in the case of economic activities, the impact of program assignment on time spent on unpaid household services during the school off-season was not significantly different from zero.

Subgroup analysis (see Appendix C) shows that the reduction in time spent on unpaid household services activities was particularly larger for girls, with an estimated reduction of 197 minutes per week (more than 3 hours), and a reduction of 240 minutes (4 hours) during the school season.

It is worth noting that the total decline in economic activities (195 minutes) and unpaid household services (135 minutes) is on aggregate only 35 percent of the additional time spent on school-related activities (951 minutes). Both school and work-related activities include commuting time, so the increase in time spent on school-related activities among treatment children may be associated with a shift in time spent in other activities, such as time with family or friends, caring for other family members, leisure, or rest.

Table 7. Work-related outcomes at endline by treatment status (ITT)
Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E.	β	S.E.	*p*
Economic Activities							
Involved in hazardous child labor (%)	27.3	3.2	27.4	2.7	0.6	3.4	.87
Suffered work-related injuries (%)	22.3	2.6	19.6	2.2	-2.0	2.9	.49
Involved in economic activities (%)	30.7	3.8	32.4	2.9	1.2	4.0	.77
Minutes spent on economic activities per week	676.3	99.1	474.2	52.7	-194.5	95.5	.04*
Minutes spent on economic activities per week during the school season (household reports)	707.6	101.8	500.5	56.6	-211.1	104.2	.04*
Minutes spent on economic activities per week during the non-school season (household reports)	739.9	104.9	632.4	63.7	-93.0	108.1	.39
Non-economic Activities							
Minutes spent on unpaid household services per week	929.5	52.5	811.4	40.4	-135.0	40.0	.01**
Minutes spent on unpaid household services per week during the school season (household reports)	904.3	59.9	771.6	51.2	-137.0	51.3	.01**
Minutes spent on unpaid household services per week during the non-	936.0	61.8	892.4	54.0	-57.3	55.9	.31

school season (household reports)			
Sample Size	418	866	1284
*Statistically significant at p<.05 **Statistically significant at p<.01			

3.3.2. TREATMENT ON THE TREATED RESULTS

The ITT results presented above represent the impact of initial assignment, which is a valid estimate of the effect of the treatment on the project areas. This impact is of interest for program evaluation purposes, and also whenever the program is expected to be scaled-up.

However, in cases where the actual treatment is different from assignment due to subject non-compliance or treatment refusal, the impact of the treatment itself may differ from the initial ITT estimate. This TOT impact may be of interest when the treatment can be delivered in different ways from those implemented in the current project, rendering project-specific non-compliance effects irrelevant, or when the project is not expected to be scaled up.

In order to obtain a TOT estimate, it is first necessary to determine the degree of non-compliance and take-up among subjects initially assigned to the control and treatment groups. This can be done at the village level, which is the level of randomization, and the individual level, which is the level where the treatment is actually received. These two levels are used to create two different definitions, including a village-level TOT and child-level TOT.

Treatment on the Treated Villages

Treated individuals include the children who were in villages where a community school was eventually established. The Wald estimator for this TOT estimate is derived from the percentage of the treatment and control subjects receiving the treatment. The four non-complying villages only covered 27 children, so the compliance rates were still fairly high: 98.4 percent of the children initially assigned to the treatment group were eventually in a treatment village, and only 3.1 percent of the children in the control group were eventually in a treatment village.[20] The resulting Wald inflation factor is 1.05, indicating that the true TOT estimate would be five percent larger than the ITT estimate.

The interpretation of this Wald estimator is valid based on a set of assumptions (Angrist, Imbens, and Rubin, 1996). First, that the eventual treatment offer is random (independence assumption). Second, that the offer increases the probability of taking up the program equally for all subjects (monotonicity assumption). Third, that the offer of the program does not have an independent effect on the outcome variables except through actually joining the program (exclusion restriction). We have no prior reason to believe that the monotonicity assumption does not hold. We expect that offering Awareness Raising (AR), Income-Generating Activities (IGA), CS, and

[20] In the presence of non-perfect compliance in the control group, that is, when some individuals in the control group end up being treated, the Wald estimator gives us the effect of the treatment on the compliers, which is known as the local average treatment effect (LATE). However, since the number of non-compliers in the control group is small, this report refers to the TOT estimate for simplicity.

Take-Home Rations (THR) would make all subjects more likely to attend school and less likely to work. Also, we expect the exclusion restriction assumption to hold. Other than the effect from AR, IGA, enrolling in a CS, and receiving a THR, we do not foresee any indirect effects on outcomes of experimental assignment. Of these assumptions, the first one is most likely violated in the case of village-level TOT, although the number of individuals involved means that the effect of this violation is probably small.

TOT program impacts were estimated an instrumental variables approach (2SLS). The TOT estimates resulting from the 2SLS estimator are shown in Tables 8 and 9. The tables also show the mean and S.E. for each group, with control and treatment defined in terms of eventual establishment of a CS in the village. As already discussed, any differences between the ITT and village-level TOT estimates of impact are small given the reduced number of subjects involved.

Table 8. School-related outcomes at endline by treatment status (TOT Villages)

Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E	β	S.E.	*p*
School enrollment (2011–2012, %)	14.3	2.3	62.5	3.9	48.7	4.7	.00**
Current school attendance (2012–2013, %)	23.5	2.5	71.3	2.9	51.2	3.2	.00**
Current community school attendance (2012–2013, %)	3.1	1.2	58.4	3.6	56.5	3.5	.00**
Ever school attendance (%)	49.6	4.1	86.6	2.1	40.7	3.3	.00**
Time on school-related activities	498.0	48.1	1,450.0	68.5	997.8	74.8	.00**
Time on school-related activities (household reports)	383.5	63.0	1,520.4	99.5	1,153.8	115.1	.00**
Sample Size	419		865		1284		

Note: 'Treated Villages' defined as those villages where the program was eventually established, irrespective of initial experimental assignment.

*Statistically significant at p<.05
**Statistically significant at p<.01

Table 9. Work-related outcomes at endline by treatment status (TOT Villages)
Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E.	β	S.E.	*p*
Economic Activities							
Involved in hazardous child labor (%)	26.5	3.2	27.8	2.7	0.6	3.5	.86
Suffered work-related injuries (%)	21.5	2.6	19.9	2.3	-2.0	2.9	.47
Involved in economic activities (%)	29.9	3.7	32.8	2.9	1.2	4.1	.76
Minutes spent on economic activities per week	650.7	97.3	486.4	53.8	-204.0	96.7	.04*
Minutes spent on economic activities per week during the school season (household reports)	690.3	100.9	509.2	57.2	-222.9	105.7	.04*
Minutes spent on economic activities per week during the non-school season (household reports)	717.3	103.1	643.5	64.4	-98.0	109.9	.37
Non-economic Activities							
Minutes spent on unpaid household services per week	946.3	51.6	803.1	40.4	-141.7	39.5	.01**
Minutes spent on unpaid household services per week during the school season (household reports)	919.5	58.9	764.5	51.2	-144.7	51.4	.01**
Minutes spent on unpaid household services per week during the non-school season (household reports)	950.2	60.0	885.6	54.5	-60.5	56.6	.29
Sample Size	419		865		1284		

Note: 'Treated Villages' defined as those villages where the program was eventually established, irrespective of initial experimental assignment.

*Statistically significant at p<.05
**Statistically significant at p<.01

Treatment on the Treated Children

An alternative and perhaps more conventional TOT estimate results from scaling up the ITT estimates by the proportion of subjects that eventually took up the program. In the current evaluation program take-up is operationally defined as current attendance to a CS. The Wald estimator for this TOT estimate is thus derived from the percentage of the treatment and control subjects that are currently attending a CS. The take-up rates are much lower at the subject level than compliance at the village level: 61.0 percent of the children initially assigned to the treatment group were currently attending a CS, and only 3.8 percent of the children in the control

group were currently attending a CS.[21] The resulting Wald inflation factor is 1.85, indicating that the true TOT estimate would be 85 percent larger than the ITT estimate.

As in the case of the village-level TOT estimate, the interpretation of this Wald estimator is valid based on the independence, monotonicity, and exclusion restriction assumptions. As in the case of the village-level TOT estimate, the independence assumption is likely violated for a small number of children. The exclusion restriction would entail assuming that the experimental assignment affects school and work-related outcomes exclusively via the administration of CSs and THRs. Since experimental assignment was random, this assumption is fairly safe.

TOT program impacts were estimated using an instrumental variables approach (2SLS). The TOT estimates resulting from the 2SLS estimator are shown in Tables 9 and 10. The tables also show the mean and S.E. for each group, with control and treatment defined in terms of current attendance to a CS. These TOT estimates give an approximation to the effect of the treatment under-perfect take-up.

TOT estimates for school-related outcomes are shown in Table 10. Since current school attendance is by definition 100 percent for the treatment group, the impact of the treatment is most informative on time spent on school-related activities. The impact on the treated subjects would be an increase of 1,765 minutes (29 hours) on school-related activities per week, with household informants reporting an even greater impact of 2,046 minutes (34 hours).

Table 10. School-related outcomes at endline by treatment status (TOT Children)

Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E.	β	S.E.	*p*
School enrollment (2011–2012, %)	23.5	2.7	81.1	4.2	86.1	7.3	.00**
Current school attendance (2012–2013, %)	25.7	2.0	100.0	0.0	90.6	4.3	.00**
Current community school attendance (2012–2013, %)	0.0	0.0	100.0	0.0	-	-	-
Ever school attendance (%)	57.3	3.2	100.0	0.0	71.9	4.7	.00**
Time on school-related activities	528.8	44.2	2,042.1	36.3	1,765.1	100.6	.00**
Time on school-related activities (household reports)	594.8	65.7	1,975.9	109.9	2,046.0	171.6	.00**
Sample Size	767		517		1284		

Note: 'Treated Children' defined as those children that were currently attending a CS, irrespective of initial experimental assignment.
*Statistically significant at p<.05
**Statistically significant at p<.01

[21] As noted previously, this is most likely the result of respondent error. For estimation purposes we must, however, base our calculations on data as reported.

The TOT estimates for work-related outcomes are shown in Table 11. The impact of the TOT subjects would be a decrease of 361 minutes (6 hours) on economic activities per week. This would go up to 412 minutes (7 hours) during the school season. The impact on unpaid household services would be a reduction of 251 minutes per week (4 hours), with this effect more likely during the school season than the school off-season.

Table 11. Work-related outcomes at endline by treatment status (TOT Children)
Outcome means and standard errors by treatment groups and significance of group differences [Egypt, 2013]

	Control		Treatment		Program Impact		
	Mean	S.E.	Mean	S.E.	β	S.E.	*p*
Economic Activities							
Involved in hazardous child labor (%)	32.9	2.5	19.2	3.0	1.1	6.1	.86
Suffered work-related injuries (%)	24.8	2.1	14.0	2.5	-3.7	5.1	.47
Involved in economic activities (%)	37.1	2.8	24.2	3.4	2.2	7.2	.76
Minutes spent on economic activities per week	761.5	70.5	212.7	35.7	-361.0	168.7	.03*
Minutes spent on economic activities per week during the school season (household reports)	770.1	71.7	269.8	43.4	-412.6	184.0	.03*
Minutes spent on economic activities per week during the non-school season (household reports)	848.6	74.8	400.7	57.2	-189.8	192.8	.33
Non-economic Activities							
Minutes spent on unpaid household services per week	914.9	41.1	753.7	33.1	-250.6	71.1	.01**
Minutes spent on unpaid household services per week during the school season (household reports)	883.0	47.0	721.1	43.2	-267.1	89.6	.01**
Minutes spent on unpaid household services per week during the non-school season (household reports)	927.1	47.9	886.5	48.2	-114.9	98.9	.25
Sample Size	767		517		1284		

Note: 'Treated Children' defined as those children that were currently attending a CS, irrespective of initial experimental assignment.

*Statistically significant at p<.05
**Statistically significant at p<.01

4. DISCUSSION

4.1. INTERPRETATION

The impact evaluation of the Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt (CWCLP) project aimed to estimate the impact of the project on school and work-related outcomes among children engaged in or at risk of exploitative child labor. These program impacts were evaluated 1 year after CWCLP program implementation began.

This evaluation found significant evidence that the CWCLP project had a positive impact on the schooling situation of eligible children. Assuming that children would enroll in school if a Community School (CS) is made available would be a reasonable assumption, but in the context of the impact evaluation the effect must still be quantified. This is not only to verify that the program had the expected impact, but also to quantify this impact. The evaluation also found a significant impact on the allocation of time to economic activities and unpaid households services. The evaluation failed, however, to detect any impact on overall participation in economic activities, exposure to workplace hazards, or occurrence of work-related injuries. Each outcome is discussed in more detail below.

Program impact on school-related outcomes

The program increased school enrollment, school attendance, and the time children spent on school-related activities relative to the control group. As the majority of children in the treatment villages were attending CS, most of the impact on school-related outcomes can be directly linked to the implementation of CS. These findings are informative about the importance of availability of an adequate education infrastructure, but they cannot be dissociated from the impact of Take-Home Rations (THR).

In spite of the significant impact of the program, 29.5 percent of eligible children in the treatment villages were still not attending any type of school at the time of the endline survey, and 42.2 percent were not attending a CS. There are two alternative explanations for this. First, it is possible that some children were not offered the opportunity to attend a CS. Indeed, among children who were currently attending some other type of school, 42.5 percent had not been offered a place in the CS run by the CWCLP program. Among children who had never attended school, 45.8 percent had not been offered a place in the CS. The two groups of children that report not being offered a place in a CS combined represent nearly 8 percent of all the children in the treatment group.

A second possibility is that, in spite of being offered a place in the school, children or their parents were not interested in taking-up the offer. This second group of children is more interesting from a theoretical perspective, as it may help explain the child and household-level decisions driving take-up. Among children currently attending school, the main reason they cite for not taking-up the offer is that they went to a government school instead. This group of children would have probably gone to school even in the absence of a CS. In fact, based on the

estimated counterfactual from Table 6, it can be expected that about 20 percent of children (those currently attending school minus those who report attending a CS in the control group) would have gone to a government school absent treatment. Among those children who never attended school, the main reason cited for not taking-up the offer is lack of interest in school (53 percent), followed by considering school "not important" (20 percent). Only a few cases cited having to work to generate income for the household (7 percent) or not having money to pay for schooling (11 percent).

These combined findings would suggest that failure to take-up is driven either by failure to deliver treatment to eligible children or by lack of interest in schooling, rather than the cost of attending school or the opportunity cost of foregone income. This may provide some rationale to believe that the THR and school-costs support component of the program is sufficient to offset the impact of fixed schooling costs and the opportunity cost of working in the study population.

Finally, it is possible that the impact of the program on school-related outcomes is somewhat subdued due to implementation challenges. According to the midterm evaluation (Hassan, 2012), the political environment following the 2011 revolution, changes in government, and the security situation made it hard for CWCLP to develop the required partnerships in a timely and efficient manner, which caused delays in project implementation. Specifically, security problems delayed the accreditation of some CS by the Ministry of Education. Partner nongovernmental organizations (NGOs), Community Development Associations, and other stakeholders noted that the difficulties encountered in securing community donations for the establishment of CS and finalizing contracting of teachers caused delays in opening CS in some areas. Most notably, due to political challenges and difficulties identifying an NGO with appropriate capacity for implementation in the Sharquiya Governorate, implementation did not begin there until the end of 2012, meaning that endline data were collected after only 3–4 months of program implementation. In order to maintain the integrity of the sample and the Intention to Treat (ITT) estimates, the sample from Sharquiya Governorate was kept in the final analysis. However, the impact of this delay on the final estimates is probably small, with only 38 children in the endline sample affected (about 3 percent of the sample).

Program impact on work-related outcomes

The program decreased the time children spend in economic activities and unpaid household services relative to the control group. Since household informants report that this decrease is only significant during the school season, and children in the treatment group had much higher levels of school enrollment and attendance than children in the control group, program impacts on work-related outcomes can be logically attributed to the implementation of the project.

These findings confirm the causal link hypothesizing that children, as a result of the intervention, would shift time from economic activities and unpaid household services to school-related activities. It seems, however, that children do not only make room for schooling by shifting time from economic activities, but they also shift time from other non-economic activities.[22] This evaluation focused on unpaid household services, as they share some of the same deleterious

[22] This study cannot determine exact agency in child time-allocation decisions. It is likely that most decisions about child time allocation are made at the household level. However, for simplicity, this report refers to children as agents.

effects of economic activities on child welfare opportunities. The project significantly reduced the amount of time that children spend in such unpaid household services.

This impact evaluation did not, however, find evidence of reduced involvement in economic activities in general, exposure to workplace hazards, or occurrence of work-related injuries. It is possible that as a result of the intervention, children shifted into types of work that are easier to combine with transition education rather than completely withdrawing from work. It is also possible that children simply reduce the amount of time they work in their preexisting occupations. Qualitative reports from the midterm evaluation (Hassan, 2012) are consistent with this hypothesis. Some children reported working fewer hours, while others reported quitting the more labor intensive agricultural activities altogether to help with the sale of farm produce in the market after finishing homework. In this evaluation, this effect would show up in the number of hours worked, but not on children's overall involvement in economic activities, which would be consistent with our findings.

An associated finding that is worth examining is the sharp decrease in economic activity across groups from baseline to endline. More than three-fourths of children in both groups were involved in economic activities at baseline. This proportion was down to about a third for both groups by endline. This large reduction was not expected, particularly for the control group. Rather, the opposite effect was likely: as children in the control villages grow older, their participation in economic activities would tend to increase.

One possibility to explain such an effect is that treatment spillovers might have contaminated the control group. However, this explanation is unlikely. The cluster randomization design limited the likelihood of spillovers, and strict program monitoring of eligible children meant that there were no opportunities for children in the control group to enroll in a program CS or receive a THR. Anecdotal evidence from the World Food Programme (WFP) suggests the possibility of spillovers from the other project components. It is possible that the improved livelihoods components may have resulted in an improved overall economic situation and that these benefits are shared, particularly among extended family members living in nearby control villages. A spillover of the Awareness Raising (AR) component is also possible, as both information and the staff involved in the program travel across treatment groups. In some cases, CS teachers lived in control villages while they worked in treatment villages.

Besides spillovers, which appear to be a minor threat considering the cluster design and strict program monitoring, anecdotal evidence from El Zanaty and Associates and WFP indicates that economic activity throughout Egypt is indeed reduced since the 2011 Egyptian revolution. This trend is supported by unemployment data. In June 2011, official unemployment was 11.8 percent (versus 9 percent in June 2010), the highest rate in 10 years. It is possible therefore that external events may have obscured any impact that the project could have on participation in economic activities and the associated measures of exposure to workplace hazards and workplace injuries.

However, leaving these explanations aside, the significant decrease in economic activities from baseline is most likely explained by seasonality in the data. Baseline data were collected between October and November 2011, which the WFP program identifies as the high agricultural season in the targeted governorates. Most harvesting occurs in October and November, especially the cotton and rice harvesting, which are considered the most important harvest events in these areas.

Planting of some crops such as wheat, tomatoes, and onions also occurs during this season. Endline data, on the other hand, was collected in March and April 2013, a period that is only associated with less important agricultural events such as the onion, tomato, and wheat harvests. Although a 12-month reference period was used for work-related outcomes, it is possible that respondents discount seasonal work from their reports of economic activities. Seasonality in the data and the possibility of biased recall would mean that impacts identified by this evaluation may more readily be applicable to off-seasonal work-related outcomes. There is a possibility that this evaluation is less informative about treatment effects on child work during the high agricultural season. In any case, it must be clarified that the effects found by this evaluation are not in any way biased by seasonality effects. Program impacts are derived from a control/treatment comparison, not from a pre/post comparison. Control/treatment comparisons are independent of seasonality effects, as the season is the same for both groups. Finally, indicators such as involvement in economic activities, school enrollment, and school attendance are based on a 12-month reference period, so even if program impacts are seasonally elastic, the indicators should capture impacts over a full 12-month period. Only the possibility of biased recall could explain failure to identify an effect on work-related outcomes. However, we can only speculate about the actual existence of such bias.

The comparison of ITT versus Treatment of the Treated (TOT) estimates of causal impacts on work-related outcomes is in any case enlightening about the potency of the intervention as a tool to reduce child labor. This evaluation shows that the components of the project are an effective way to increase school enrollment, school attendance, and time spent on school-related activities. The intervention also achieves a reduction in time spent in economic activities and unpaid household services. However, the evaluation does not find that the intervention has a measurable impact on other work-related outcomes, including overall participation in economic activities, exposure to workplace hazards, and work-related injuries. As the TOT estimate shows, this is probably true even if it were possible to improve program take-up among eligible children.

These findings are consistent with previous experimental evidence indicating a moderate or non-existent effect of schooling support and THRs on child labor (see for example Edmonds & Shrestha, 2012). However, the program intervention decreased time spent on economic activities by about 200 minutes, while it increased time spent on school activities by about 1,000 minutes per week. Every 5 hours of school resulted in about 1 hour less of work. This result shows that substantial schooling (an extra 16–19 hours a week) can be provided to children during times they would have otherwise worked on farms. Current findings and prior evidence suggest that the current treatment may need some reformulation in order to reach the stated objective of reducing participation in hazardous child labor. Possible alternatives include a stronger component to increase knowledge and understanding (among children and parents) of the hazards of agricultural labor.

4.2. LIMITATIONS

In addition to the limitations discussed above regarding the possibility of biased recall, there are several limitations of the study, including lack of administrative data, lack of evaluation of individual program components, and the lack of generalizability of the findings.

A limitation of the current evaluation was the inability to use administrative data on program take-up as well as key outcomes such as school attendance. Although such data were collected by WFP as part of its monitoring system, these data could not be linked to the impact evaluation data at the individual level. Using these administrative data could have helped with the validation of household or child reports, and could have also reduced the impact of non-response and attrition. Validating child and household reports by comparing them to administrative records could be the subject of further research.

In the absence of administrative data, take-up was defined using self-reports of CS attendance, which do not identify the other components of the program and may be liable to respondent error. Key outcomes such as economic status or school attendance were also measured using self-reports. Although household and child-level reports were generally consistent and self-reports appeared to be internally valid, these reports may be externally liable to social-desirability bias and/or manipulation motivated by expectations of some future benefit.

Administrative data might not have been a source of perfect information, though. For example, it is possible that data on key outcomes might have been available only in program areas with a more active monitoring system. Even if such data could have been complemented with survey data from control areas, this could introduce method bias in the control-treatment comparisons. For reasons such as these, it is usually recommendable to integrate monitoring and evaluation components so that both program and survey data can be used to evaluate key outcomes.

A limitation of the evaluation design was the lack of evaluation of individual program components and/or dose-response assessment. All project components were evaluated in combination, so the impact estimates cannot disentangle the individual effects from each component. We do, however, anticipate that the THR and CS components might be responsible for the majority of the effects identified by the evaluation. This is because the evaluation was done after a year of program implementation, so it is more likely to capture short-term effects, such as those expected from the CS and THR components. It is expected, on the other hand, that the effect of the AR and Income-Generating Activities might be more noticeable in the midterm.

Additionally, there are priors, both in the literature and based on anecdotal evidence from the CWCLP program, that the CS and THR components might have an asymmetrical impact. For example, qualitative reports from the midterm evaluation report (Hassan, 2012) indicate that while the THR component was highly valued by mothers, they would still send their children to school if the THR component were to stop. It is also possible that the impact of the THR component is heterogeneous at different levels. Estimating the relative impact of each project component would be useful to fine-tune the program before scaling-up. This would require the use of additional multiple treatment groups, for example a CS-only and CS+THR treatment group. A more ambitious design would assess dose-response by setting the THR at different values for additional treatment groups.

Finally, the results from this impact evaluation are not generalizable beyond the study population. The sample selected for this evaluation is representative for those communities in which the CWCLP project is currently operating. The project may operate differently in other parts of the country, so the sample is not generalizable to the whole population of Egypt. Furthermore, because only the first 30 eligible children who were identified in each village were

included in the study, results may not generalize well to those children who may enter the CS component later in the year. Lack of direct generalizability is a limitation of most RCTs that do not rely on nationally or regionally representative samples. However, we believe that the impact of this intervention might be replicable in populations similar to those in the sample, that is, children ages 6–11 living in small rural hamlets with no primary school where children were engaged in or at risk of participating in hazardous child labor. It is hoped in any case that the results of this impact evaluation will add to the knowledge base regarding the impact of this type of educational intervention on child labor.

BIBLIOGRAPHY

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). *Identification of Causal Effects Using Instrumental Variables*. Journal of the American Statistical Association, 91 (434) 444–455.

Avvisati, F., Gurgand, M., Guyon, N., & Maurin, E. (n.d.). Getting Parents Involved: A Field Experiment in Deprived Schools. Available at: http://www.restud.com/wp-content/uploads/2013/07/MS16235manuscript.pdf

Bloom, H. S. (2006). The Core Analytics of Randomized Experiments for Social Research. *MDRC Working Papers on Research Methodology.* Available at: http://www.mdrc.org/sites/default/files/full_533.pdf

Bloom, H. S. (2005). Randomizing groups to evaluate place-based programs (pp. 115–172), in Howard S. Bloom (ed.) *Learning more from social experiments: Evolving analytic approaches*. New York: Russell Sage Foundation.

Edmonds, E. and Shrestha, M. (2012). *The Schooling Incentives Project Evaluation. Research Project on Children Working in the Carpet Industry in India, Nepal and Pakistan.* ICF International. Available at: http://www.dol.gov/ilab/programs/ocft/pdf/FinalReportSIPEStudy.pdf

Hassan, N. (2012). *Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt*. Midterm Evaluation report.

Holtman, M. C., Jones, S. E., Merola, S. S., & Liem, J. (2012). *Evaluation of USDOL services to prevent child labor in Egypt: Randomized trial to evaluate "Combating Worst Forms of Child Labor by Reinforcing Policy Response and Promoting Sustainable Livelihoods and Educational Opportunities in Egypt" (CWCLP) project, 2012 baseline report*. ICF International. Available at: https://ocftdata.dol.gov/Downloads/Evaluations/Reports/Egypt_Baseline_Report.pdf

Raudenbush, Stephen W. (1997). *Statistical Analysis and Optimal Design for Group Randomized Trials*. Psychological Methods 2(2): 173–85.

Schulz, K. F. & Grimes, D. A. (2002). *Allocation concealment in randomised trials: defending against deciphering.* The Lancet 359, 614–618. Available at: http://apps.who.int/rhl/LANCET_614-618.pdf

APPENDIX A: DEFINITION OF OUTCOME VARIABLES

School-Related Outcomes

Outcome	Definition	Questionnaire	Denominator	Variable definition
1. School enrollment	Child attended school during the 2011–2012 school year	Child	All children	Enrolled if C104 = 1
2. Current school attendance	Child is attending school at the time of the interview (2012–2013 school year)	Child	All children	Currently attends if C103 = 1
3. Current attendance to a community school	Child is attending a community school at the time of the interview (2012–2013 school year)	Child	All children	Currently attends if C103 = 1 & C106a = 2
4. Ever school attendance	Child is currently attending (2012–2013 school year), attended school in the 2011–2012 school year, or attended ever	Child	All children	Ever attended if C103 = 1 or C104 = 1 or C201 = 1
5. Time on school-related activities (child report)	Minutes spent on school-related activities per week	Child	All children	C134AH*60+C134AM ('0' if not attending school)
6. Time on school-related activities (adult report)	Minutes spent on school-related activities per week	Household	All children	(H102_1A*5+H102_1B+H102_1C + H102_2A*5 + H102_2B+H102_2C)*60 ('0' if not attending school)

Work-Related Outcomes

Outcome	Definition	Questionnaire	Denominator	Variable definition
1. Hazardous child labor	Child is exposed to workplace hazards	Child	All children	Hazardous labor if "yes" to any C222 or C132
2. Work-related injuries	Child has suffered a work-related injury in the last 12 months	Child	All children	Injured if "yes" to any C223 or C133
3. Economic activities	Child performed any economic activity in the last 12 months	Child	All children	Child worked if C121A = 1 or C215A = 1
4. Time spent on economic activities	Minutes spent on economic activities activities per week	Child	All children	Time = C134BH*60+C134BM or C224BH*60+C224BM ('0' if not working)
5. Time spent on economic activities (school season)	Minutes spent on economic activities per week	Household	All children	If (H101A = 1) time = (H102_3A*5) + H102_3B + H102_3C)*60 If (H101A = 2) time = (H103_1A*5) + H103_1B + H103_1C)*60
6. Time spent on economic activities (non-school season)	Minutes spent on economic activities per week	Household	All children	Time = (H103_1A*5) + H103_1B + H103_1C)*60
7. Time spent on unpaid household services	Minutes spent on unpaid household services per week	Child	All children	Time = C134BH*60+C134BM or C224BH*60+C224BM
8. Time spent on unpaid household services (school season)	Minutes spent on unpaid household services per week	Household	All children	If (H101A = 1) time = ((H102_4A*5) + H102_4B + H102_4C)*60 If (H101A = 2) time = ((H103_2A*5) + H103_2B + H103_2C)*60
9. Time spent on unpaid household services (non-school season)	Minutes spent on unpaid household services per week	Household	All children	Time = ((H103_2A*5) + H103_2B + H103_2C)*60

APPENDIX B: PARTICIPANT FLOW

Table 12. Participant flow

CWCLP Program Evaluation [Egypt, 2013]

Enrollment	n	%		
Assessed for eligibility	2,707			
Ineligible	1,284	47.4		
Randomized	1,423	52.6		
	Treatment		**Control**	
Randomization	n	%	n	%
Allocation	983	69.1	440	30.9
Baseline Survey	n	%	n	%
Surveyed	983	100.0	440	100.0
Not Surveyed	0	0.0	0	0.0
Endline Survey	n	%	n	%
Surveyed	870	88.5	421	95.7
Not Surveyed	113	11.5	19	4.3

APPENDIX C: SUBGROUP ANALYSIS

School-Related Outcomes

Table 13. School-related outcomes at endline by treatment status and sex (Intention to Treat)

Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (Boys)			Program Impact (Girls)		
	β	S.E.	*p*	β	S.E.	*p*
School enrollment (2011–2012, %)	42.2	5.0	.00**	48.6	6.2	.00**
Current school attendance (2012–2013, %)	41.5	3.8	.00**	53.8	5.0	.00**
Current community school attendance (2012–2013, %)	44.8	3.5	.00**	62.2	4.9	.00**
Ever school attendance (%)	27.4	3.7	.00**	49.9	4.8	.00**
Time on school-related activities	757.5	98.0	.00**	1094.3	108.4	.00**
Time on school-related activities (household reports)	931.5	131.9	.00**	1209.3	147.9	.00**
Sample Size	593			691		

*Statistically significant at p<.05

**Statistically significant at p<.01

Table 14. School-related outcomes at endline by treatment status and age at baseline (Intention to Treat)

Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (6–8 years)			Program Impact (9–11 years)		
	β	S.E.	*p*	β	S.E.	*p*
School enrollment (2011–2012, %)	55.0	5.4	.00**	41.0	5.2	.00**
Current school attendance (2012–2013, %)	52.0	3.6	.00**	46.0	4.4	.00**
Current community school attendance (2012–2013, %)	61.6	4.6	.00**	47.2	3.7	.00**
Ever school attendance (%)	46.0	3.8	.00**	34.0	4.1	.00**
Time on school-related activities	1,011.0	93.9	.00**	900.1	96.6	.00**
Time on school-related activities (household reports)	1,345.2	133.0	.00**	935.0	134.5	.00**
Sample Size	498			786		

*Statistically significant at p<.05

**Statistically significant at p<.01

Table 15. School-related outcomes at endline by treatment status and work status at baseline (Intention to Treat)

Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (Not Working)			Program Impact (Working)		
	β	S.E.	*p*	β	S.E.	*p*
School enrollment (2011–2012, %)	65.0	6.2	.00**	42.9	5.2	.00**
Current school attendance (2012–2013, %)	61.4	5.4	.00**	44.9	3.9	.00**
Current community school attendance (2012–2013, %)	72.0	4.8	.00**	49.4	4.1	.00**
Ever school attendance (%)	58.8	5.0	.00**	34.2	3.6	.00**
Time on school-related activities	1,264.5	141.8	.00**	853.9	88.4	.00**
Time on school-related activities (household reports)	1,550.9	150.6	.00**	1009.0	126.7	.00**
Sample Size	285			999		

*Statistically significant at p<.05

**Statistically significant at p<.01

Work-Related Outcomes

Table 16. Work-related outcomes at endline by treatment status and sex (Intention to Treat)
Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (Boys)			Program Impact (Girls)		
	β	S.E.	*p*	β	S.E.	*p*
Economic Activities						
Involved in hazardous child labor (%)	-2.4	4.1	.55	1.7	4.7	.71
Suffered work-related injuries (%)	-5.5	3.4	.11	3.9	4.2	.93
Involved in economic activities (%)	-2.5	4.1	.54	3.9	5.7	.50
Minutes spent on economic activities per week	-223.7	102.6	.03*	-156.9	138.8	.26
Minutes spent on economic activities per week during the school season (household reports)	-275.4	104.9	.01*	-159.1	153.4	.30
Minutes spent on economic activities per week during the non-school season (household reports)	-125.7	111.9	.26	-70.8	158.2	.66
Non-economic Activities						
Minutes spent on unpaid household services per week	-69.2	46.7	.14	-197.1	64.2	.00**
Minutes spent on unpaid household services per week during the school season (household reports)	-13.7	51.6	.79	-240.2	75.8	.00**
Minutes spent on unpaid household services per week during the non-school season (household reports)	27.2	53.9	.62	-121.4	83.5	.15
Sample Size		592			691	

*Statistically significant at p<.05
**Statistically significant at p<.01

Table 17. Work-related outcomes at endline by treatment status and age at baseline (Intention to Treat)

Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (6–8 years)			Program Impact (9–11 years)		
	β	S.E.	*p*	β	S.E.	*p*
Economic Activities						
Involved in hazardous child labor (%)	2.3	3.9	.55	-1.0	4.3	.82
Suffered work-related injuries (%)	1.0	3.0	.75	-4.3	4.2	.30
Involved in economic activities (%)	1.8	4.6	.69	0.5	4.8	.93
Minutes spent on economic activities per week	-165.3	95.2	.09	-203.4	132.5	.13
Minutes spent on economic activities per week during the school season (household reports)	-148.0	93.1	.12	-256.6	144.6	.08
Minutes spent on economic activities per week during the non-school season (household reports)	-85.5	101.2	.40	-101.5	147.6	.49
Non-economic Activities						
Minutes spent on unpaid household services per week	-111.5	53.8	.04*	-145.0	59.6	.02*
Minutes spent on unpaid household services per week during the school season (household reports)	-156.0	71.8	.03*	-120.3	60.5	.05*
Minutes spent on unpaid household services per week during the non-school season (household reports)	42.4	78.7	.59	-55.1	63.7	.39
Sample Size		501			786	

*Statistically significant at p<.05

**Statistically significant at p<.01

Table 18. Work-related outcomes at endline by treatment status and work status at baseline (Intention to Treat)

Estimated impacts and significance of differences by experimental assignment [Egypt, 2013]

	Program Impact (Not Working)			Program Impact (Working)		
	β	S.E.	*p*	β	S.E.	*p*
Economic Activities						
Involved in hazardous child labor (%)	2.6	5.5	.63	0.1	4.0	.98
Suffered work-related injuries (%)	1.4	3.6	.68	-3.2	3.6	.38
Involved in economic activities (%)	0.5	6.5	.93	1.1	4.6	.81
Minutes spent on economic activities per week	-98.1	129.9	.45	-226.2	119.9	.06
Minutes spent on economic activities per week during the school season (household reports)	-102.2	137.6	.46	-268.4	130.7	.04*
Minutes spent on economic activities per week during the non-school season (household reports)	-13.3	140.4	.92	-138.7	132.6	.30
Non-economic Activities						
Minutes spent on unpaid household services per week	-70.3	71.1	.33	-163.7	48.2	.00**
Minutes spent on unpaid household services per week during the school season (household reports)	-190.5	79.7	.02*	-135.2	57.9	.02*
Minutes spent on unpaid household services per week during the non-school season (household reports)	-31.5	89.9	.73	-70.2	62.0	.26
Sample Size		285			998	

*Statistically significant at p<.05

**Statistically significant at p<.01