January 2017

Tayari Programme

Longitudinal Midterm Report for the Tayari Early Childhood Development and Education Programme

Prepared for

Children's Investment Fund Foundation
7 Clifford Street
London
W1S 2FT
UNITED KINGDOM

Prepared by

Dunston Kwayumba, Dr Benjamin Piper, Arbogast Oyanga and Moses Oyagi
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709
USA

RTI Project Number 0214471

RTI INTERNATIONAL

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

APBET	Alternative Provision of Basic Education and Training
APHRC	African Population and Health Research Center
CHA	Community Health Advisor
CHV	Community Health Volunteer
DAT	Direct Assessment Tool [for pre-primary; adapted for Tayari from UNICEF MELQO instrument]
DICECE	District Centre for Early Childhood Education
DiD	differences-in-differences
ECD	early childhood development
ECDE	early childhood development education
GER	gross enrolment ratio
KICD	Kenya Institute of Curriculum Development
MDG	Millennium Development Goal
MECP-Kenya	Madrassa Early Childhood Programme Kenya (MECP-Kenya)
MELQO	Measuring Early Learning and Quality Outcomes
MoE	Ministry of Education
NACECE	National Centre for Early Childhood Education
NER	net enrolment ratio
NESP	National Education Sector Plan
PP1, PP2	pre-primary 1 and 2
PRIMR	Primary Math and Reading Initiative
RCT	randomised controlled trial
RTI	RTI International (registered trademark and trade name of Research Triangle Institute)
SCECDC	Sub-County ECD Coordinator
SD	standard deviation
SES	socioeconomic status
SMS	(text messaging service)
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
WHO	World Health Organization

EXECUTIVE SUMMARY

The Tayari Early Childhood Development and Education Programme is an intervention implemented by the Kenyan Ministry of Education and four counties, with technical support from RTI International and funding from the Children's Investment Fund Foundation (CIFF). Tayari is focused on designing, piloting and testing the cost effectiveness of investments to increase school readiness for children in Early Childhood Development and Education (ECDE) centres in the four counties.

The Ministry of Education selected Uasin Gishu, Siaya, Laikipia and Nairobi counties for implementation in public ECDE centres and Nairobi Alternative Provision for Basic Education and Training (APBET) centres in nonformal settlements. To better understand the combination of interventions that have the largest impact on school readiness, Tayari is testing three different interventions. Treatment 1 focuses on providing instructional support and training to teachers teaching in pre-primary 1 (4-year-olds) and pre-primary 2 (5-year-olds), using the ECDE approach implemented successfully elsewhere in Kenya, but without a formal impact evaluation. Treatment 2 provides the same amount of training and support to teachers, but also supplies teachers' guides to teachers in four subjects and workbooks to PP1 and PP2 children in the subjects of language and mathematics. Treatment 3 offers the same investment as Treatment 2, but also supports the county health system to utilise the ECDE centres as a locus for their support for referrals and health record-keeping, with particular emphasis on low-cost, locally managed hand-washing support.

The Tayari intervention is designed as a randomised controlled trial, with the African Population and Health Research Center (APHRC) as the external evaluator estimating the effect of the three treatment groups on school readiness. To complement the external evaluation of APHRC, the Tayari team is undertaking a longitudinal tracer study of children in control and Treatment 2 ECDE centres, to better understand how school readiness skills transition over time. The longitudinal study assessed randomly selected children during the January 2016 baseline, and is proceeding to follow them by reassessing the same children at the October 2016 midterm and the October 2017 endline assessment.

For the midterm evaluation, if any selected child was not in attendance at school during the day of the midterm data collection, the Tayari team used the contact information collected at baseline, plus information available at the ECDE centre, to locate and assess that child. Attrition at the October 2016 midterm assessment was 11.2%.

Using weighted data and multiple regression analysis to compare the gains in school readiness between Treatment 2 and control schools, we found that Tayari had a larger impact on school readiness at the October 2016 midterm assessment round than it was expected to have at the October 2017 endline. ***Figure ES1*** below shows the change in school readiness results for the treatment and control groups.

Figure ES1: School Readiness Index results

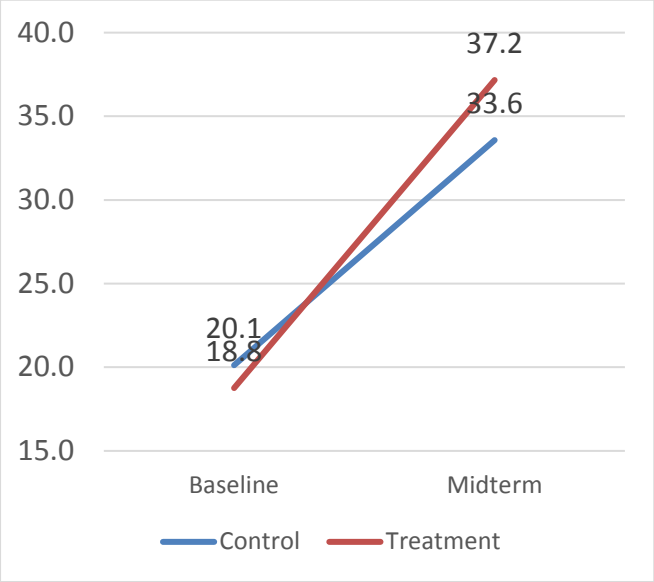
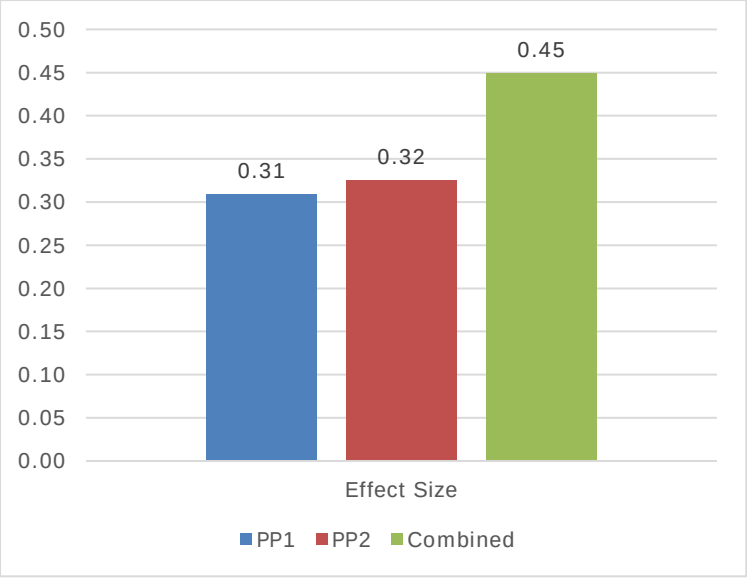


Figure ES1 shows that the learning increases in the Tayari group were much larger than in the control group (18.4 compared with 13.5 points on a School Readiness Index). These results are encouraging after the first several months of Tayari implementation. The results are for the combined sample, not differentiating by class level.

In order to reveal whether these results are significant and meaningful, ***Figure ES2*** presents the effects of Tayari by class level. These include PP1 and PP2, but also the somewhat large percentage of classrooms that did not differentiate by PP1 and PP2, and instead had a combined classroom.

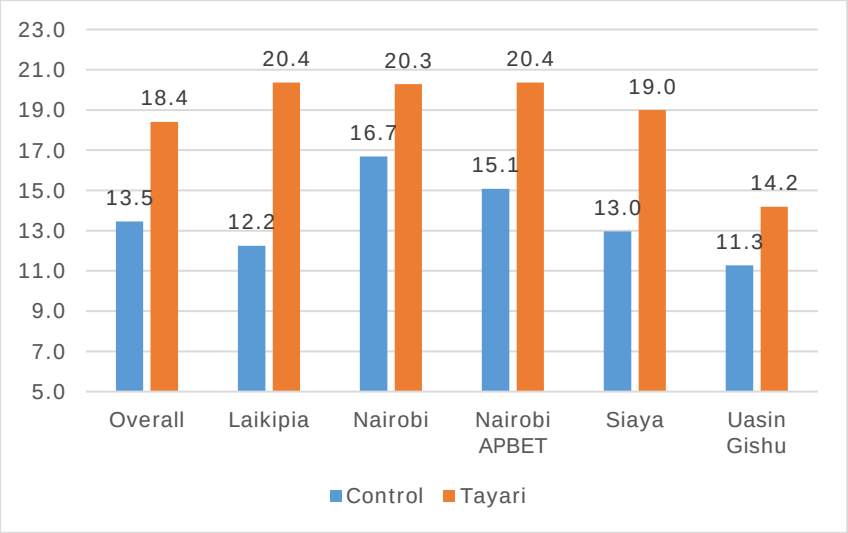
Figure ES2. Tayari effect size by class level



The overall effect of Tayari was 0.34 standard deviations (SD). This average breaks out as 0.31 SD for PP1, 0.32 SD for PP2 and 0.45 SD for Combined classrooms. Encouragingly, as noted earlier, these results show that Tayari had a larger effect at the midterm assessment then it was expected to have at the endline assessment, after more than two years of classroom implementation.

While the design of the Tayari intervention is not structured to be able to make county-level comparisons, the rich data and relatively large samples at each county allow for a general comparison between the counties at the midterm assessment. ***Figure ES3*** shows the gains in school readiness for the Tayari treatment and control by county and for the Nairobi APBET portion of the study.

Figure ES3. Impact of Tayari on school readiness by county



The results showed higher school readiness outcomes for each of the four counties, Nairobi APBET and the entire sample. Not all of these differences are statistically significant, given the relatively small county-level samples and the associated weighting.

The purpose of the Tayari midterm longitudinal impact evaluation was to determine whether and how the Tayari programme was having an impact on school readiness. The initial midterm results compiled by APHRC showed consistently significant impacts on school readiness, for each level, and for both boys and girls. The large longitudinal sample and the additional measures assessed in the Tayari longitudinal assessment mean that by the endline assessment, Tayari will be able to produce nuanced information about how school readiness in Kenya transitions over time and what instructional factors had the largest impact on learning.

Recommendations

Based on these midterm longitudinal findings, we present the following recommendations:

- **Consider the magnitude of the first-year results from Tayari.** The results presented in this report show effects that are commensurate with those of the Primary Math and Reading (PRIMR) Initiative (2011–2014) after its first year, and larger than many or most other large-scale ECDE improvement programmes. This finding suggests that Tayari has the potential to be a successful intervention at scale.
- **Continue to implement pilots at medium scale, through government mechanisms.** The Tayari design ensures that implementation of the programme occurs through the government systems at both the national and county levels. The training, the coaching and the support to Tayari are being implemented by Kenyan officials and field-based officers. Thus, the Tayari results are likely indicative of what would happen if the programme were scaled up to additional sites.
- **Consider scaling up the Tayari intervention.** While we recommend waiting until the endline assessment in October 2017 to determine whether these encouraging results continue, we suggest that Kenya begin now to consider

scaling up the Tayari programme more broadly, based on the encouraging findings.

- **Improve school readiness measures.** The reliability statistics presented in this report show that the Tayari assessment tools meet the accepted research standards for reliability and discrimination. Given that few school readiness tools have been utilised in developing countries, it is particularly important for the Tayari team to work on refining and improving the new set of tools that Tayari is experimenting with from the Measuring Early Learning and Quality Outcomes (MELQO) activities funded by CIFF. In particular, some of the executive function and socio-emotional items did not perform as well as hoped, and additional work should be undertaken to better assess these portions of school readiness.
- **Review the selection of and results from APHRC public control schools.** While these longitudinal results were largely similar to APHRC's results, children in the APHRC public control schools improved their school readiness results by significantly more than the control children in the longitudinal samples presented here. The results for the APHRC external evaluation control schools also increased by more than those for the APBET control schools, which – based on APHRC's work in the sector and RTI's work in APBET schools – is not consistent with results in any other study over the past several years. In order to obtain a fair analysis of Tayari's impact in public schools, these odd results should be carefully considered, and potentially additional schools included in the control sample to determine whether these results are consistent at the endline assessment. In addition, given the differences in scoring on the school readiness index for the APHRC external evaluation and the longitudinal study, we suggest joint training of assessors going forward for the January 2017 and October 2017 assessments.

1. KENYA'S ECDE CONTEXT

Early childhood development (ECD)[1] is a field that serves a crucial developmental period for children and lays the foundation for their holistic development (Pence, 2011; World Health Organization [WHO] & UNICEF, 2012). ECD has been an increasingly important area for investment in Kenya, with the government having signed and ratified several international conventions that relate to the provision of ECD. These conventions include the African Charter on the Rights and Welfare of the Child, the United Nations Convention on the Rights of the Child (1990) and its domestication through the enactment of the Children Act of 2001; the Education for All goals (1990); the Millennium Development Goals (MDGs, 2000–2015) and the recently adopted Sustainable Development Goals (2015). More importantly, the Kenyan Ministry of Education (MoE) has invested in the design and development of a revised Integrated ECD Policy Framework to replace the 2006 National ECD Policy Framework and to align the ECD services provision with the Kenya Constitution of 2010. The MoE recently announced that the fundamental design of the new curriculum structure will start with two years of ECD.

Beyond the new ECD policy and the revised curriculum plans, several other key documents focus on the importance of ECD access for Kenyan children. These include the Sessional Paper No. 1 of 2015 (Republic of Kenya, 2015), the Sessional Paper No. 14 of 2012 (Republic of Kenya, 2012b) the Basic Education Act of 2013 (Republic of Kenya, 2013) and the National Education Sector Plan (NESP) of 2013–2018 (Republic of Kenya, 2014). Each of these documents states Kenya's desire to integrate ECD services for 4- and 5-year-old children into the formal (and eventually, free) education system. The NESP goes beyond simple advocacy to describe how ECD can be implemented in the decentralised system, and describes opportunities for counties to improve access and raise the quality of ECD, with particular attention to improved pedagogical methods.

Kenya's Vision 2030 underscores the need for increased access to high-quality education from the formative years of development, and the integration of ECD into basic education. Empirical results show that inequalities in access to ECD services persist in Kenya, and children from low-income households are less likely to access ECD compared to their counterparts from well-to-do households. For those who do attend ECD, they are more likely to enrol in school at the right age, transition to primary schooling, and complete primary schooling without dropping out or repeating grades (Ngaruiya, 2006; UNICEF, 2012; UNESCO, 2015).

[1] ECD refers to policy and programmes for children from birth to eight years, and their parents and caregivers who protect children's right to develop their full cognitive, emotional, social and physical potential (UNICEF, 2001). Early childhood development education (ECDE) is offered in institutions bearing various names, including ECDE centre, pre-primary, preschool, pre-units, nursery, and day care. The 2006 policy framework adopted *ECD* as an encompassing term which also includes the ECD curriculum-defined baby care, pre-primary 1 and pre-primary 2.

1.1 Kenya ECD policy

The 2006 National ECD Policy Framework (Republic of Kenya, 2012a) defined ECD as services for children from conception to eight years, and stated that every child shall have a right to life, survival and development and parental care. This language is repeated in the draft 2016 Integrated ECD Policy Framework, which is anchored in the principles of human rights, calling for more mobilisation of resources by ensuring integration of ECD programmes into the national and county governments' development plans. Key among these is the call for providing capitation grants for learners in the pre-primary sector and also ensuring a seamless transition from ECDE centres to primary schools.

The Basic Education Act of 2013 defines pre-primary as the education offered to children of four to five years before they join Class 1. The policy environment in Kenya suggests that ECD services are a partnership between the MoE (for policy guidelines, standards and curriculum setting) and counties (for funding and implementation). At the local level, counties continue to work with parents and local communities to manage ECD centres, including hiring and paying teachers and other caregivers. The draft 2016 Policy Framework proposes that county governments solicit resources through partnerships with local or national organisations.

1.2 Access to ECD in Kenya

ECD serves a large proportion of the Kenyan population, as approximately 20% of the Kenya population is below five years old (Republic of Kenya, 2012a). The 4- to 5-year-old cohort (official pre-primary age) is projected to increase from 3 million in 2014 to 3.3 million in 2018 (Republic of Kenya, 2014). Kenya has made remarkable progress in ECD enrolment, with official MoE statistics suggesting a net enrolment ratio (NER) of 32.9% in 2005 and, more recently 71.8% in 2014. Encouragingly, ECD enjoys nearly net gender parity (0.99; Republic of Kenya, 2014). Projections suggest up to an 81.4% net enrolment rate by 2018 (Republic of Kenya, 2014). However, there remains the problem of over-age children in ECD as the gross enrolment ratio (GER) was 74% in 2014, and the difference between the GER and the NER suggests significant enrolment from over-age children. To serve this growing number of children, the number of ECDE centres increased from 38,247 in 2009 to 40,211 in 2014, and is projected to increase by 10% more to 44,232 in 2018 (Republic of Kenya, 2014). During the same period, overall ECD enrolment is expected to increase from 2.2 million (2014) to 2.7 million (2018) children.

In spite of the impressive national increases in enrolment, significant regional inequalities remain in access and participation at the ECDE level. Many counties have low levels of enrolment compared to the target of universal access and the MDG target of 80% access by 2015. Low NER of less than 10% is typical of counties in northeastern Kenya, particularly Wajir, Mandera and Garissa counties. High NER levels above the MDG target are experienced in Nyeri, Kiambu, Kisumu and Nairobi counties, with the rest of the country having a wide range of enrolment ratios. Tayari is working in Siaya, Uasin

Gishu, Nairobi and Laikipia counties, which represent a reasonable cross-section of the ECDE enrolment population. ***Table 1*** presents the number of ECDE centres and enrolment in the Tayari pilot counties.

Table 1. Number of ECDE centres and enrolment ratios in pilot counties

	No. of ECDE centres		GER (%)			NER (%)		
County	Public	Private	Girls	Boys	Total	Girls	Boys	Total
Laikipia	317	192	77.7	82.2	79.9	80.8	76.3	78.6
Nairobi	213	1,841	70.5	82.3	76.3	68.8	80.8	74.6
Siaya	744	140	74.6	72.3	73.5	72.5	70.3	71.4
Uasin Gishu	498	313	58.8	62.8	60.8	57.6	61.85	59.7

Source: Ministry of Education, Science and Technology, Republic of Kenya (2014).

1.3 Successes and Challenges in the Decentralisation of ECD

As mentioned above, the Kenyan Constitution makes ECD (child care services and pre-primary) a county function and allocates the responsibility for planning, developing and implementing ECDE to the county governments (UNESCO, 2015). This transition towards the county level of ECD services has been complex. A few successes should be noted, including increased enrolment in ECDE centres; construction of new ECDE centres and model ECDE centres; recruitment of teachers in some counties; and the provision of some centre-based feeding programmes.

However, significant challenges remain for implementing ECD through the counties:

- Lack of clear management and administration of ECD due to inadequate collaboration between the county and national government.
- Lack of effectively established and utilised ECD management committees at the school level.
- The disintegration of the NACECE-DICECE[2] system following reforms at the Kenya Institute of Curriculum Development (KICD) and the transfer of personnel after the devolvement of pre-primary to counties.
- The decision to devolve some functions of ECD (construction, management) while retaining others at the national level (teacher training and deployment).
- Limited clarity regarding the funding mechanism. Clear capitation grants are not yet in place for ECD, and at the same time, counties are not provided with

[2] The National Centre for Early Childhood Education (NACECE) was based at the KICD and served as a centre for ECD curriculum and curriculum support materials development, resource mobilisation and capacity building of district officers and other personnel in ECDE. District Centres for Early Childhood Education (DICECEs) were decentralised units of the NACECE at the district level. These structures are currently moribund.

unit-based funding for ECD to allow them to manage the financial planning of the sub-sector.

- Poor instructional quality at ECD centres.
- Low-quality teacher management systems, which are affected by the lack of a scheme of service for ECD personnel and confusion regarding whether the Teachers' Service Commission or the counties should hire them.
- Shortage of qualified ECD personnel at the counties, as many experienced DICECE officers were transitioned back to the national level, and counties have not successfully replaced them.
- Existence of different curricula being implemented and an overemphasis on academic examinations for very young children.
- Inappropriate and inadequate play and learning materials and inadequate nutrition and health support services (Republic of Kenya, 2006).
- The ECDE policy required that in order to improve access, ECDE centres should be located not more than 1 kilometre away from households. This has resulted in the mushrooming of private ECD centres. However, counties are unable to efficiently manage the quality of ECD delivery in these centres.
- Unclear roles in special-needs education. Given that special-needs education remains the responsibility of the national government, it is unclear how counties and the national level will work together to support special-needs children at the ECD level (Murunga, 2015).

2. BACKGROUND TO THE TAYARI PROGRAMME

Results from the Primary Math and Reading (PRIMR) Initiative implemented by RTI International (2011–2014) indicated that well-designed, syllabus-based teachers' guides; intensive teacher training; and targeted ongoing teacher support through coaches improved instruction led to significant gains in lower primary pupils' literacy and numeracy performance (Piper & Mugenda, 2016; Piper, Ralaingita, Akach, & King, 2016). It also showed that an investment in improved training and instructional support had modest impacts on learning outcomes, even without access to learning materials. Based on these results, the literacy portion of the PRIMR Initiative was subsequently scaled up to all public and 1,000 private primary schools in Kenya, while the numeracy component is being scaled up to all public primary schools. Building on the strength and success of PRIMR, the Tayari programme was designed to test whether a programme dependent on the existing MoE and county structures can also improve ECDE outcomes, as it did in primary.

2.1 Goals and Objectives of Tayari

The overarching goal of the Tayari ECD programme is to support the Kenyan government to develop a cost-effective and scalable ECD model to enhance the proportion of pre-primary children transitioning to Class 1 with the requisite school

readiness skills and competencies in literacy, numeracy, executive functioning and social-emotional skills. The specific objectives of the Tayari intervention are to:

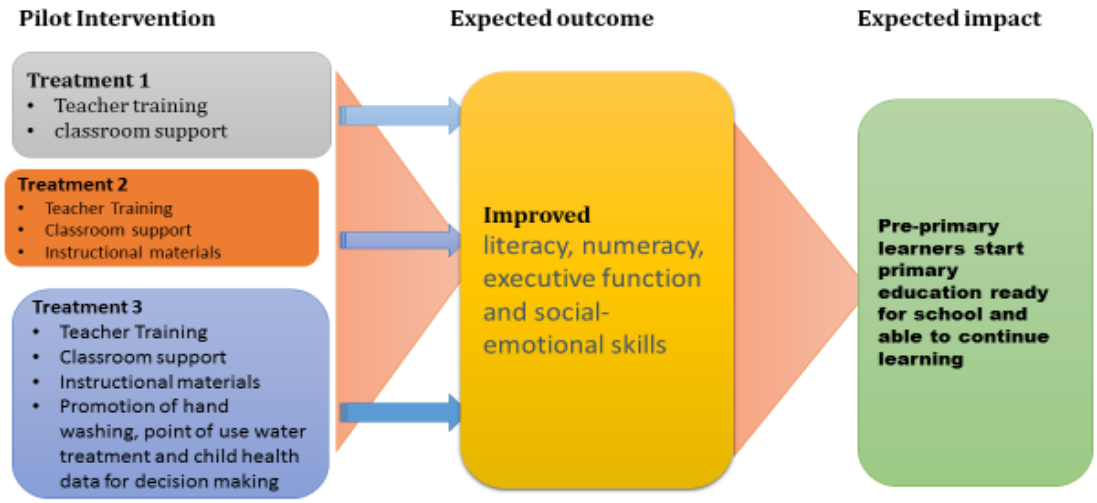
- Develop a tested, cost-effective, affordable and scalable ECDE model.
- Pilot test and expand provision of quality ECDE through a tested and cost-effective ECDE model.
- Promote global uptake of a tested ECDE model through the provision of open-source ECDE materials and leverage communication of activities at the global level.

2.2 Overall Tayari Implementation

The MoE determined that the Tayari intervention would be piloted in the four counties of Laikipia, Nairobi, Siaya and Uasin Gishu. The programme is being implemented in public ECDE centres across the four counties. In Nairobi, in addition to the public centres, the programme is also being piloted in Alternative Provision of Basic Education and Training (APBET) ECDE centres.

In order to contribute to our knowledge of how Kenya can improve the quality of ECD delivery at low-cost, the Tayari programme tests the impact of three different treatment groups (see ***Figure 1***) against a control group.

Figure 1. Tayari theory of change



2.2.1 Treatment 1: Training and Support

The Tayari Training and Support intervention, or Treatment 1, builds on the best practice of other ECDE programmes implemented previously in Kenya. In particular, the Madrassa Early Childhood Programme–Kenya (MECP-Kenya) was selected to support the implementation of this treatment arm. This treatment group focuses on training and supporting Sub-County ECD Coordinators (SCECDCs) under the county governments to more effectively train and support teachers in ECDE centres to improve their pedagogical

skills. ECDE officers in this treatment group are given electronic tablets to assist in classroom observation. The tablets are programmed to guide the officers in observing practices that are included in the teacher training, and to suggest feedback on these practices. The observational tool is therefore equipped for giving feedback to each observed teacher on how they are providing instruction in classrooms.

In addition, the teachers in target schools are trained for between 3 and 5 days at the beginning of each term (three terms per academic year) on improved pedagogical practices. Some of these aids are given directly to teachers, while the teachers are also taught and supported to make other materials during and after training. The training intervention uses teaching and learning materials that schools already have and are available in the market. The following key elements are included in this Training and Support treatment group:

- **SCECDC training with tablets:** SCECDCs are trained to supervise teachers in the ECDE centres for which they are responsible. The training is tailor-made and aims at helping the SCECDCs to familiarise themselves with and use tablet-based technology to determine whether their teachers are teaching in a manner consistent with the enhanced pedagogical training.
- **Enhanced teacher training:** SCECDCs, in turn, train the teachers in their zones, clusters, or wards on how to improve the quality of their instruction across subjects. This training covers how to increase active learning, to increase instructional time, and to develop more teaching aids. The teacher training focuses on helping teachers to improve their pedagogical practices using instructional resources already available in the centres.
- **Instructional aids:** SCECDCs work with teachers during the training on developing some key instructional resources. Whereas the Tayari programme provides some teaching aids, schools are encouraged to contribute towards this initiative through head teacher sensitisation and Centre Management Committee review meetings.

2.2.2 Treatment 2: Training, Support and Instructional materials

The second treatment—Training, Support and Instructional Materials—is similar to the first, but it includes provision of learners' books and teachers' guides as well. The Treatment 2 cohort receives the same number of days of training and tablets for instructional support for teachers in their zones, albeit in a separate training from the Treatment 1 group. However, unlike Treatment 1, which provides feedback on using existing instructional materials in the schools, the SCECDC observational tool in Treatment 2 is equipped to help coaches give feedback to teachers on how they are using the specific instructional materials developed by Tayari.

The Tayari books and teachers' guides were designed by a joint MoE, KICD and Tayari technical team that followed a structured scope and sequence based on the KICD syllabus. Learner books were matched with teachers' guides to help teachers implement high-quality instruction. Books for this treatment group were produced and printed at a

low cost and provided to schools at a ratio of one pupil per textbook (1:1). The following key elements are included in the Treatment 2 group:

- **SCECDC training with tablets:** The SCECDC officers are trained to supervise teachers in the ECDE centres for which they are responsible. The training helps the officers use tablet-based technology to determine whether the teachers are teaching in a manner consistent with the enhanced pedagogical training.
- **Teacher training:** SCECDCs train the teachers in their zones, clusters or wards on how to improve the quality of their instruction across subjects. This training covers how to increase active learning, to increase instructional time, and to develop more instructional materials. In this treatment group, the enhanced teacher training prepares teachers to use the books and teachers' guides developed by Tayari.
- **Books and teachers' guides:** Each learner in this treatment group receives low-cost instructional materials, including learner activity workbooks for language and mathematics and charts for life-skills development activities. The learner workbooks are distributed at a 1:1 ratio. Apart from receiving guides for language and mathematics, teachers also receive guides for social activities and life skills. It should be noted that the learner workbooks contain activities that are matched to the lesson and the teachers' guides. In addition, the number of lessons differs by activity area, with language and numeracy being taught daily, social skills taught twice per week, and life skills activities once per week.

2.2.3 Treatment 3: Training, Support, Instructional Materials and Health

Treatment 3 builds on Treatment 2. The SCECDCs are trained and provided with tablets that contain the same classroom observational tools linked to the Tayari books and teachers' guides as in Treatment 2. In addition, the Treatment 3 group supports the Kenyan government's health system by increasing connections between health interventions and the ECDE subsector, to use government-managed ways of integrating ECDE into the health system. The following key elements are included in the Treatment 3 group:

- **SCECDC training with tablets:** SCECDCs are trained to supervise teachers in the ECDE centres for which they are responsible. The training utilises tablet-based technology to determine whether the teachers are teaching in a manner consistent with the enhanced pedagogical training.
- **Teacher training:** SCECDCs train teachers in their zones, clusters or wards on how to improve the quality of instruction across the main activity areas. The training covers how to increase active learning, to increase instructional time, and to develop more instructional materials. In this treatment group, the enhanced teacher training also helps teachers use the books and teachers' guides effectively. Further, the teachers are trained on a child health promotion component.

- **Books and teachers' guides:** Each pupil in these treatment groups receives low-cost instructional materials at a 1:1 ratio. These materials contain a page of activities per lesson. The learner materials match and interact with teachers' guides developed by the programme that support teachers to teach the KICD curriculum.
- **Support for child health:** This component entails providing tablet-based observational tools to help Community Health Advisors (CHAs) adequately support the Community Health Volunteers (CHVs) to work with teachers to promote child health in the centres. The CHAs are given professional development and they in turn train the CHVs on the implementation of Tayari health interventions. Specifically, CHVs support the setting up of low-cost hand-washing facilities in ECDE centres; undertake point-of-use water treatment; and improve the capacity of ECDE centres to maintain and use child health records.

2.2.4 Control Group

In this randomised controlled trial (RCT), the three treatment groups are compared with a control group. ECDE centres in control wards, zones or clusters will not receive any intervention until after the end of the pilot. Thereafter, for ethical reasons, they will receive whichever of the three treatments is judged most cost-effective.

2.2.5 Tayari Implementation

Table 2 presents a summary of the overall implementation of Tayari project; details by county appear in ***Appendix A.*** The project is being implemented in two phases. The first phase of implementation began in January 2016; the second phase begins in January 2017, while the control group will begin implementation in January 2018.

During the first year, the programme was implemented in 635 schools reaching 35,450 children. The primary project activities included focused, practice-based teacher training. At the pinnacle of the training, 22 master trainers from the national and county governments, 26 SCECDCs and 8 instructional coaches were trained. The 26 SCECDCs and 8 instructional coaches trained and supported 1,286 teachers in the three treatment groups in phase 1 across the counties of implementation.

During the same period of implementation, 92,837 teaching and learning materials were procured, printed and distributed to the ECDE centres receiving the intervention. These included 50,011 pupil workbooks for numeracy, language, social and life skill activities; 5,481 teacher guides; 10,184 big books; 1,019 charts; and 25,804 manipulatives. Manipulatives provided to these centres comprised letter and number flashcards, object cards, counters and crayons. Similarly, 338 integrated registers for documenting health services for each child were distributed to the treatment group receiving the health intervention. Below, we cover specific interventions by treatment group.

Table 2. Tayari overall implementation design

Location	Treatment group	Schools starting 2016	Schools starting 2017	2018 zones (control)	2018 schools (control)	Total zones	Total schools
APBET	Training and Support	48	43			6	91
	Training and Support + Books / Teachers' Guides	44	52			6	96
	Training and Support + Books / Teachers' Guides + Health	35	31			4	66
	Control zones			6	92	6	92
Nairobi	Training and Support	26	17			4	43
	Training and Support + Books / Teachers' Guides	31	40			6	71
	Training and Support + Books / Teachers' Guides + Health	20	10			2	30
	Control zones			4	48	4	48
Uasin Gishu	Training and Support	45	38			4	83
	Training and Support + Books / Teachers' Guides	59	55			6	114
	Training and Support + Books / Teachers' Guides + Health	17	15			2	32
	Control zones			6	114	6	114
Laikipia	Training and Support	49	60			4	109
	Training and Support + Books / Teachers' Guides	98	52			6	150

Location	Treatment group	Schools starting 2016	Schools starting 2017	2018 zones (control)	2018 schools (control)	Total zones	Total schools
	Training and Support + Books / Teachers' Guides + Health	34	20			2	54
	Control zones			3	95	3	95
Siaya	Training and Support	39	47			4	86
	Training and Support + Books / Teachers' Guides	72	76			6	148
	Training and Support + Books / Teachers' Guides + Health	18	27			2	45
	Control zones			6	126	6	126
Total	Training and Support	207	205			22	412
	Training and Support + Books / Teachers' Guides	304	275			30	579
	Training and Support + Books / Teachers' Guides + Health	124	103			12	227
	Control zones			25	475	25	475
	Total treatment zones	**635**	**583**			64	1,218
	Total with control zones					**89**	**1,693**

Training and Support (Treatment 1)

As explained in Section 2.2.1, under the Training and Support treatment group, six SCECDCs and three instructional coaches were trained to train and supervise teachers in the 207 ECDE centres which they oversee. The training helped the SCECDCs and coaches use active-learning methodologies to train teachers on this approach. They used tablet-based technology to give feedback during and after lesson observations. The six SCECDCs and instructional coaches trained 439 teachers in 11 zones in the four counties. In addition, they reached and sensitised 207 head teachers of the schools they supported. Teachers were trained on how to improve the quality of their instruction across subjects. This training sought to help teachers to develop low-cost teaching and learning resources such as charts and counters. It also helped the teachers improve their making and use of professional records such as lesson plans, schemes of work, progress records and record of work covered. In this Treatment group, 12,006 learners were reached.

Training and Support + Books/Teachers' Guides (Treatment 2)

As indicated in Section 2.2.2, in the Treatment 2 group, 14 SCECDCs and 3 coaches were trained to train and supervise teachers in their zones in the four counties in the 304 ECDE centres for which they were responsible. The training enabled the SCECDCs and coaches to train, support and help teachers to use the instructional materials provided, such as teachers' guides and pupil workbooks. They also use tablet-based technology to give feedback during and after lesson observations.

Using the cascade mode of training, the trained SCECDCs and instructional coaches in turn trained 587 teachers in 15 zones. The training focused on improving the quality of instruction, giving more time for learners' participation, making and using low-cost teaching and learning resources, and using instructional materials provided by Tayari.

In this treatment group, 587 teachers and 17,532 learners were provided with teaching and learning materials respectively at the ratio of 1:1.

Training and Support + Books/Teachers' Guides + Health (Treatment 3)

As noted (see Section 2.2.3), in the Treatment 3 group, six SCECDCs and three coaches were trained to train and supervise teachers in the ECDE centres for which they were responsible. Similar to the training in Treatments 1 and 2, the SCECDCs and Coaches later trained 124 head teachers and 260 teachers in their six zones on how to improve the quality of their instruction across subjects.

The community health system at ECDE centres was also supported under this treatment group. In this regard, Tayari provided training for 14 CHAs, who in turn trained 30 CHVs across the four counties. In addition, CHAs were given tablets for data collection. These tablets supplied the system with data on health outcomes and activities at the 124 ECDE centres. They also facilitated linkages between the ECDE centres and the corresponding health facilities. The CHVs supported 124 ECDE centres to improve hygiene practices (in

addition to their other roles), with particular emphasis on hand-washing using soap or ash. CHVs fed information to the CHAs through an SMS-based questionnaire developed by the programme.

Teachers and learners under this treatment group also received teachers' guides and pupil workbooks, respectively. Across the four counties, the teaching and learning materials reached 260 teachers and 5,908 learners.

2.3 Challenges

2.3.1 Parallel Activities by SCECDEs

One of the main challenges during phase 1 included multiple parallel duties expected of the SCECDEs. These included leading educational activities such as meetings, conducting trainings, and setting and marking exams. This additional layer of responsibilities affected the number of times they supported teachers via classroom observations. On average, only 39% of ECDE centres were observed during the first year of implementation.

2.3.2 Low Staffing of SCECDEs

In some counties of intervention, such as Laikipia, during the first term of implementation, only two SCECDCs were deployed. However, the situation improved in the third term when an additional four were hired.

Terrain challenges with some schools located very far from each other reduced the number of support visits a SCECDC could manage. This problem was especially acute in Laikipia and thus fewer ECD centres were covered there (39%) compared to other counties.

Some of the ECD centres do not have physical structures; hence, children learn under trees. This situation hampers the effective use of instructional materials provided by Tayari. Lessons take longer, as learners spend more time completing individual and group activities.

Nairobi County experienced some challenges engaging the county stakeholders for Treatment 3, particularly the county heath teams. Some trainings for the health intervention occurred later than planned, which reduced the amount of time for CHAs and CHVs to visit ECD centres. This issue reduced the proportion of APBET ECD centres covered to 8%. Laikipia County is vast; hence, CHVs and CHAs had a challenge accessing some of the ECD centres in remote areas due to lack of regular means of public transport. As a result, on average, they covered 18% of the ECD centres.

Refer to Table 2 above for the overall implementation design for the Tayari programme. The numbers of ECDE centres by treatment group, treatment phase and county are indicated. Overall, in the first year, the programme was implemented in 635 ECDE centres across the four counties. During the second phase of implementation, which begins in January 2017, 583 additional ECDE centres will join the study. As noted earlier,

it is also planned that 89 ECDE centres in the control group will receive the most effective intervention beginning 2018.

3. RESEARCH DESIGN AND METHODOLOGY

This section covers the Tayari longitudinal midterm research design and the second (midterm) data collection point. In addition, the section describes the methodology used for the longitudinal data collection and analysis, including the sampling procedure, the tools used for data collection, the training of field teams, the actual ECD students who responded, and the data cleaning and data analysis strategies.

3.1 Study Design

The Tayari midterm data collection was the second evaluation round within a longer longitudinal study that utilises randomised controlled trial methods. The Tayari longitudinal study follows a set of children from PP1 and PP2 into primary school, and collects information from these children at different time periods. As described above, the Tayari RCT has three treatment groups. Tayari used random selection to select the zones, clusters and wards that were included in Tayari, and then randomly assigned all ECDE centres within a zone, cluster or ward to one of the three treatment groups or the control group.

Longitudinal studies have an advantage over cross-sectional studies in terms of their ability to describe how change happens over time within children. In order for a cross-sectional study design to precisely measure intra-individual change, the following assumptions must be met: (1) Different subjects come from the same population, (2) subjects match across age groups, and (3) different-aged subjects have similar socio-economic background. It is clear that these assumptions cannot be met in social studies like the Tayari RCT. Cross-sectional studies also require further statistical adjustments during analysis for accurate findings (Schaie, 1983).

3.2 Objective of the Tayari Longitudinal Midterm Study

The main objective of the midterm study was to analyse the effect of the programme on school readiness over the implementation period. A second objective was to investigate the factors that influenced the relationships between background variables and school readiness. Given the cost and complexity of longitudinal data collection, Tayari is only comparing Treatment 2 with the control group. An impact evaluation of Treatments 1 and 3 is being done by Tayari's external evaluator, African Population and Health Research Center (APHRC).

3.3 Sampling and Sample Size

The midterm study utilised the same procedures for sampling as in the baseline and followed up at the same schools for the second round of data collection (see Kwayumba & Piper, 2016). The minimum number of ECDE centres and APBET institutions required to detect a mean effect size of 0.2 standard deviations (SDs) was determined using a power calculation with the following parameters:

- Power of 80%[3];
- Centre-level intra-class correlation of 0.249[4];
- Significance level of 0.05[5];
- Proportion of variance explained by centre-level covariance of 0.50[6]; and
- 10 pupils for PP1 and 10 pupils for PP2.

For both groups, the sample size was proportionately distributed to the counties and hence to the wards, zones and clusters. Using the sampling procedure within the Stata survey analysis software, schools were then randomly selected to participate in the Tayari longitudinal tracer study in each county and the respective zones, wards or clusters based on the proportionate allocation. Note that both treatment and control arms required 80 schools. However, 86 schools were selected for the Treatment 2 group ***(Table 3)*** and 90 schools for the control group ***(Table 4).***

Table 3. Number of schools sampled for the Treatment 2 group

County	Name of zone, ward or cluster	Total number of schools in the zone, ward or cluster	Number of schools sampled
Nairobi APBET	Congo	13	3
	Kianda	14	4
	Mwiki	13	5
Nairobi	Eastleigh	7	2
	Juja Road	10	5
	Kilimani	14	6
Laikipia	Mugogodo East Ward	38	11
	Githiga Ward	26	6
	Marmanet Ward	44	12

[3] This value is used in most impact evaluation studies in the social sciences.
[4] Calculated from PRIMR data.
[5] This value is the standard in the social sciences.
[6] Calculated from previous studies in Kenya covering the early primary school Classes 1 to 3.

County	Name of zone, ward or cluster	Total number of schools in the zone, ward or cluster	Number of schools sampled
Uasin Gishu	Soy	32	9
	Kapyemit	8	2
	Tulwet	19	5
Siaya	West Uyoma	27	7
	West Gem	24	5
	Central Sakwa	21	4
Totals	**15**	**310**	**86**

Table 4. Number of schools sampled for the control group

County	Name of cluster, ward, or zone	Total number of schools in the zone, ward, or cluster	Number of schools sampled
Nairobi APBET	Gatina	12	2
	Makina	15	5
	Riruta	18	4
	Tassia	17	3
	Kariobangi North	14	2
	Kiamaiko	16	3
Nairobi	Riruta Zone	12	3
	Bahati	10	2
	Dandora	19	3
	Karen	7	1
Laikipia	Rumuruti Township Ward	27	4
	Mukogodo West Ward	35	8
	Tigithi Ward	33	7
Uasin Gishu	Karona	17	3
	Sugoi	25	4
	Chepkoilel	8	1
	Koisagat	16	3
	Cheptiret	19	3

County	Name of cluster, ward, or zone	Total number of schools in the zone, ward, or cluster	Number of schools sampled
	Kapseret	29	5
Siaya	Ugunja	24	4
	South Gem	29	5
	East Ugenya	23	5
	South Sakwa	21	4
	Usonga Ward	14	3
	Ukwala	12	2
Totals	**15**	**472**	**90**

3.4 Sampling of Pupils in the Sampled ECDE Centres

In each of the sampled centres, 10 PP1 and 10 PP2 pupils were targeted. In centres with more than the targeted number of pupils, systematic random sampling was used, whereby all the children in each class (PP1 and PP2) were requested to line up, and every n^{th} pupil was picked to achieve the required number. Systematic random sampling was done separately for girls and boys to achieve gender parity. However, in centres where there was no clear demarcation of PP1 and PP2 learners, 20 learners (10 boys and 10 girls) were selected. These learners constituted a third subgroup of learners, designated as a 'Combined' class.

3.5 Study Instruments

The Tayari longitudinal study employed several instruments that were previously used at baseline: Direct Assessment Tool (DAT), pupil questionnaire (completed by the teacher), head teacher questionnaire, teacher questionnaire and classroom observation schedules.

3.5.1 The Direct Assessment Tool

The DAT is the Monitoring Early Learning and Quality Outcomes (MELQO) tool adapted for Kenya. The DAT midterm tool for Tayari included items on social-emotional skills; pre-academic skills, such as language, literacy and numeracy; and areas that support learning across multiple domains, such as executive functioning, persistence, self-regulation, and approaches to learning. These subtasks were grouped into four main domains: literacy, social-emotional, executive functioning and numeracy. Further, for ease of administration and to avoid learner fatigue, a pair of the domains was merged to form two sub-questionnaires: one comprising literacy and social-emotional components, and the other one, executive functioning and numeracy. ***Table 5*** and ***Table 6*** describe these subtasks in detail.

Table 5. Content of the language and socio-emotional DAT

Section	Description 1	Description 2
Section A – Socio-emotional	Perspective: Showing empathy	The child is shown a picture of a crying child and is required to show empathy
	Understanding feelings	The child is required to identify positive and negative emotional stimuli
Section B – Language	Rhyme [a]	The child is required to listen to words and identify rhyming words
	Expressive listening	The child is required to listen to a list of objects and provide a group name
	Letter-name knowledge [a]	The child is required to identify letters of the alphabet
	Letter-sound knowledge [a]	The child's knowledge of letter sounds is tested
	Initial sound identification [a]	This section tests the child's ability to identify the beginning sound of a word
	Sequencing	This section tests the child's ability to arrange images to recount a story
	Listening comprehension [a]	The child is required to listen to a story and thereafter answer questions related to the story

[a] School Readiness Index subtask.

Table 6. Content of the executive functioning and numeracy DAT

Section		Description
Section A – Executive Functioning	Dimensional change: colour sort	This section tests the child's ability to correctly identify colours visually. The child is presented with pictures of lorries and cars of two colours and is asked to separate them by colour
	Dimensional change: picture sort	This section requires the child to sort images correctly. This child is presented with pictures of lorries and cars and is supposed to place them, regardless of colour, in specific locations as instructed by the assessor
	Forward digit span	The child is required to repeat a set of digits read out in the same order
	Backward digit span [a]	The child is required to repeat the digits read out in reverse order (that is, backwards)
Section B – Numeracy	Shape naming [a]	The child is required to identify shapes
	Number identification	The child is required to identify numbers

Section		Description
	Producing a set	The child is required to produce a set of bottle tops as instructed
	Quantity discrimination [a]	The child's ability to discriminate quantity is tested
	Mental addition and subtraction	The child is required to perform addition and subtraction problems
	Measurement vocabulary [a]	The child is required to identify objects according to size
	Oral addition [a]	The child is required to perform addition problems

[a] School Readiness Index subtask.

3.5.2 Pupil Questionnaire – Administered to Teachers

To collect meaningful data on pupil context, we decided to obtain this information from specific ECDE-level teachers (PP1 and PP2) using a pupil questionnaire. Administered by the supervisor to each selected child's teacher, this questionnaire aided in the collection of data on the child's educational history, social-emotional and life skills, any special needs, and health records.

3.5.3 Teacher Questionnaire

With the understanding that factors influencing children's school readiness include school and classroom factors, the longitudinal study collected data on classroom teachers, and whether they taught PP1 or PP2 or were in charge of the entire class for cases of combined classrooms. The teacher questionnaire contained sections on teacher qualifications (including whether they had undertaken specialised courses in ECDE); availability and use of teaching and learning materials; pedagogical practices; learner evaluation; and instructional support received from the ECDE head teacher.

3.5.4 Head Teacher Questionnaire

The head teacher questionnaire was included in this study to generate information about the ECDE centre necessary to understand ECDE delivery. It collected data on school background, head teacher profile, subjects taught, and school facilities.

3.5.5 Classroom Observation Tools

The study also included classroom observation tools for both numeracy and literacy. The tools included a snapshot section on the teaching and learning process, with observations captured at three-minute intervals. Components of the snapshot section centred on teacher focus (whether at the instant of the snapshot, the teacher was concentrating on the entire class or an individual pupil; whether a group of pupils was on or off task); instructional content (the specific aspect of teaching content the teacher was covering); teacher action (what the teacher was doing); student actions (what

pupils were doing at that particular instance); and languages used alongside teaching and learning materials (English, Kiswahili or mother tongue). Afterwards, the assessor took an inventory of classroom facilities, including instructional material and the classroom's physical infrastructure.

3.6 School Readiness Index

In a workshop attended by MoE representatives, RTI and APHRC, stakeholders agreed that a School Readiness Index should be developed comprising the subtasks indicated in ***Table 7.***[7] Agreement on these subtasks was reached after confirmation that the Tayari instructional improvement programme directly covers these areas, and hence, they are expected to reflect varying levels of impact (measurability) during the course of programme implementation. Each portion of the School Readiness Index was weighted 10% and the overall index score was determined from a total of 100%.

Table 7. Subtasks selected for the School Readiness Index

Subtask no.	Subtask	Domain
1.	Backward digit span	Executive functioning
2.	Rhyme	Literacy
3.	Letter-name knowledge	Literacy
4.	Letter-sound knowledge	Literacy
5.	Initial sound identification	Literacy
6.	Listening comprehension	Literacy
7.	Shape naming	Numeracy
8.	Quantity discrimination	Numeracy
9.	Measurement vocabulary	Numeracy
10.	Oral addition	Numeracy

3.7 Enumerator Recruitment and Training

RTI maintains a database of experienced enumerators, some having worked with RTI in Kenya since 2011. A total of 98 enumerators were recruited and trained. Of these, 94 were selected, based on performance, to participate in the data collection process.

Enumerators were first trained for a period of five days. The team was taken through objectives of training, a detailed walk-through on the administration of the DAT, data collection research ethics and data security.

[7] Ten other subtasks used in the study were not included in the computation of the School Readiness Index but are maintained for research purposes, to understand how the programme works.

Fifteen supervisors and deputy supervisors were selected from this group based on their qualifications and experience. This group received additional training that entailed administering the head teacher and teacher questionnaires as well as undertaking classroom observations for numeracy and literacy.

Direct assessment being undertaken in one of the sampled ECDE centres in Laikipia County

Pilot testing of questionnaires was also undertaken on the fourth day of enumerator training. This was done concurrently with the supervisors' and deputy supervisors' training. The field practice was undertaken in four ECDE centres in Nairobi. Feedback during debriefing sessions provided rich practical experiences that were used to further reinforce the training.

As part of the training, enumerators were taken through two sets of performance assessments to measure the degree of agreement among the raters. This exercise gauged the level of homogeneity in the ratings given. In the context of fieldwork, the degree of inter-rater reliability is an important aspect of the research (Armstrong, Gosling, Weinman, & Marteau, 1997). The average score for these tests was 96.0%, which is comparable to the baseline score of 96.5%.

After training, enumerators were teamed into groups of six, comprising the supervisor, deputy supervisor and four assessors. In total, 15 teams were deployed to collect data in the randomly selected 176 ECDE centres across the four counties. Data collection was undertaken between from 3rd to 21st October 2016.

3.8 Data Preparation and Analysis

Data were captured using the RTI-developed Tangerine® open-source application loaded on tablets. At the end of data collection in schools and on a daily basis, data were uploaded to the central server and made available for cleaning. This method meant that data cleaning could commence while data collection was still ongoing, allowing for follow-up of issues revealed during cleaning. Technical staff at Tayari's home office downloaded all the various assessments from the Tangerine database and exported them into Stata, which they then used to undertake the data cleaning and preparation procedures. They checked the data for inconsistencies resulting from faulty skip logic and data-entry errors.

The process of data cleaning entailed confirming the required number of respondents (children, teachers and the ECDE centres), as well as the number of classroom observations, for each ECDE centre included in the study. Further checks were undertaken to ensure that each tracked child completed the two portions of the DAT

tool. Erroneous pupil identification codes were recoded, duplicates were dropped and incomplete cases were scrutinised to ensure data accuracy. A single data set was then created from the data collected at various levels at the school. The data set comprised the literacy and social-emotional DAT, numeracy and executive functioning DAT, pupil questionnaire, head teacher questionnaire and teacher questionnaire. In addition, a similar data set collected at baseline was merged with the midterm data set to allow for comparative analysis. The entire data set was then weighted in response to the sampling design in readiness for analysis. All analyses were undertaken using the svy set of commands in Stata so that the results could be weighted. Primarily, analysis was done longitudinally based on the changes in learning outcomes from baseline to midterm.

After data cleaning, basic frequencies were generated to confirm the number of children assessed. In total, 3,448 children were assessed (1,023 in Nairobi, 703 in Uasin Gishu, 927 in Laikipia and 795 in Siaya).

Table 8 shows a comparison of children who were assessed at baseline and midterm. At midterm, if a child was not there during the lesson, the assessor used the contact information collected at baseline, plus information available at the ECDE centre, to locate and assess that child. If the child still could not be identified, another child was sampled as described below. The midterm figures then included the replacement children. As indicated, the total number of children assessed at midterm was more than those assessed at baseline (3,448 vs. 3,257), also explained below.

Table 8. Children assessed at baseline and midterm (cross-sectional sample)

	Baseline			Midterm		
County	Control	Treat-ment	Total	Control	Treat-ment	Total
Laikipia	339	522	861	359	568	927
Nairobi	169	211	380	180	240	420
Nairobi APBET	334	232	566	360	243	603
Siaya	436	319	755	469	326	795
Uasin Gishu	371	324	695	376	327	703
Totals	**1,649**	**1,608**	**3,257**	**1,744**	**1,704**	**3,448**

Despite the number of children assessed at midterm being more than those assessed at baseline, not all of these children were assessed at baseline. This number comprised children who joined the study at midterm to replace those who had dropped out, as well the additional number of children who were assessed during the midterm data collection but were not present at baseline even though they had been randomly selected for the first assessment. Given the longitudinal nature of this midterm study, we paid careful attention to the number of children who were no longer accessible by the midterm. The

attrition rate is indicated in ***Table 9.*** At the midterm, Tayari experienced 11.2% attrition, with higher rates experienced in the APBET centres. Reasons for attrition varied, but included transfers out of the county of programme implementation; promotion to higher grades (for example PP1 to PP2 or PP2 to Class 1); demotion to lower grades (for example PP1 to baby class, or PP2 to PP1); or unknown reasons.

As discussed above, some children assessed at baseline were not present at midterm. This accounts for the 11.2% attrition. Note that earlier, the monitoring team had established, by ECDE centre, the number of children assessed at baseline who were still present at these centres. During the data collection day, enumerators had a list of children assessed at baseline whose consent to participate in the longitudinal study had been obtained. Upon confirmation that a child had been transferred, replacement was done by selecting a child of the same gender, who sat where the transferred child used to sit or nearest to where the transferred child used to sit. In case the child transferred to an ECDE centre within the county of intervention, an effort was made to find and assess the child. This process resulted in the slightly higher number of children assessed at midterm (3,448) as compared to those assessed at baseline (3,257).

3.9 Analytic Methodology

This midterm report shares the results from school readiness changes of children who were assessed at both baseline and midterm. To analyse for impact, we used the svy command in Stata to generate gains between baseline and midterm. A differences-in-differences (DiD) model was also used to measure the effect sizes at midterm.

Table 9. Number of children assessed at baseline and midterm (longitudinal sample)

County	Baseline			Midterm			Percentage attrition		
	Control	Treat-ment	Total	Control	Treat-ment	Total	Control	Treat-ment	Total
Laikipia	339	522	861	283	462	745	16.5	11.5	13.5
Nairobi	169	211	380	146	179	325	13.6	15.2	14.5
Nairobi APBET	334	232	566	262	202	464	21.6	12.9	18.0
Siaya	436	319	755	410	297	707	6.0	6.9	6.4
Uasin Gishu	371	324	695	346	304	650	6.7	6.2	6.5
Total	**1,649**	**1,608**	**3,257**	**1,447**	**1,444**	**2,891**	**12.2**	**10.2**	**11.2**

4. RELIABILITY ESTIMATES

In this midterm evaluation study, we carried out reliability analyses of the DAT by computing and analysing Pearson correlations and Cronbach's alpha statistics. Strong correlations denote consistency in learners' performance across subtasks. In addition, Cronbach's alpha was used to measure the internal consistency of the tool. When internal consistency is combined with item-level evaluative psychometric methods, it offers insights on subtask functioning. In this analysis, we used Cronbach's alpha to estimate items' consistency in measuring the concepts of interest within each subtask (Aron, Coups, & Aron, 2013; Cronbach, 1951). We first present the Pearson correlations[8] for each of the four domains of the DAT, then the Cronbach's alpha scores for the entire tool, and lastly, the Pearson correlations and Cronbach's alpha scores for the School Readiness Index subtasks. It should be noted that the School Readiness Index subtasks, listed in Section 3.6, are among the subtasks used for the actual impact evaluation—i.e. determining the impact of Tayari on learners' school readiness.

4.1 Analysis of DAT by Domain

Keep in mind that the DAT covered four main domains: social-emotional, executive functioning, literacy, and numeracy. This section presents the findings of the reliability analyses for each of the domains, as well as a discussion of the first two domains, which have been combined as explained earlier in Section 3.5.1.

4.1.1 Analysis of Literacy Subtasks

In the literacy domain, there were seven subtasks. Specifically, we assessed rhyme, expressive listening, letter-name knowledge, letter-sound knowledge, initial sound identification, sequencing and listening comprehension. The results shown in ***Table 10*** indicated that the correlations between all these subtasks were statistically significant ($p < .001$). Reasonably strong correlations of between 0.30 and 0.44 were observed between subtasks: For instance, 0.30 between rhyme and initial sound identification; 0.32 between rhyme and listening comprehension; 0.35 between expressive listening and initial sound identification; 0.41 between expressive listening and listening comprehension; 0.30 between letter-sound knowledge and listening comprehension; 0.32 between initial sound identification and sequencing; 0.39 between initial sound identification and listening comprehension; 0.44 between sequencing and listening comprehension.

As expected, there was a moderately strong correlation (0.52) between letter-sound knowledge and initial sound identification, because these two subtasks measure the same sound concept. Likewise, the moderately strong correlations (0.44) observed

[8] We excluded the number identification subtask from the correlations because whereas other subtasks contained five items, the number identification subtask contained 20, thus making comparisons erratic. As a result, its measure departed significantly from those of the other subtasks within the numeracy domain.

between listening comprehension and sequencing were expected because, in an ideal world, learners needed to comprehend instructions to arrange items in the required sequence. The results indicated weak relationships of less than 0.30 between the rest of the subtasks.

Table 10. Pearson correlations for literacy subtasks

	Rhyme	Expressive listening	Letter-name knowledge	Letter-sound knowledge	Initial sound identification	Sequencing	Listening comprehension
Rhyme	1.00						
Expressive listening [a]	0.25***	1.00					
Letter-name knowledge	0.18***	0.21***	1.00				
Letter-sound knowledge	0.22***	0.23***	0.14***	1.00			
Initial sound identification	0.30***	0.35***	0.25***	0.52***	1.00		
Sequencing [a]	0.28***	0.28***	0.13***	0.24***	0.32***	1.00	
Listening comprehension	0.32***	0.41***	0.26***	0.30***	0.39***	0.44***	1.00

*p < .05, **p < .01, ***p < .001.
[a] Denotes non-school-readiness indicators.

Table 11 shows the Cronbach's alpha coefficient for the literacy subtasks. This table, and others like it in this report, provide two important pieces of information: the correlations between each subtask and the entire scale (i.e. item-test correlations); the correlations between each subtask and the scale including all other subtasks (i.e. item-rest correlations); and the overall Cronbach's alpha. We followed the accepted rules (Bland & Altman, 1997; Tavakol & Dennick, 2011), which we list below, to make decisions on the alpha statistics:

- $\alpha \geq 0.9$: Excellent
- $0.9 > \alpha \geq 0.8$: Good
- $0.8 > \alpha \geq 0.7$: Acceptable
- $0.7 > \alpha \geq 0.6$: Questionable
- $0.6 > \alpha \geq 0.5$: Poor
- $\alpha \geq 0.5$: Unacceptable.

The highest coefficient (0.72) was observed in the rhyme subtask. The lowest coefficients (0.68) were observed in initial sound identification and listening comprehension. The literacy subtasks had an overall reliability of 0.74, which falls in the acceptable range.

Table 11. Cronbach's alpha for literacy subtasks

Subtask	Item-test correlation	Item-rest correlation	Alpha
Rhyme	0.58	0.40	0.72
Expressive listening [a]	0.63	0.45	0.71
Letter-name knowledge	0.50	0.30	0.74
Letter-sound knowledge	0.60	0.43	0.71
Initial sound identification	0.72	0.57	0.68
Sequencing [a]	0.62	0.44	0.71
Listening comprehension	0.72	0.57	0.68
Total			**0.74**

[a] Denotes non-school-readiness indicators.

4.1.2 Analysis of Numeracy Subtasks

The numeracy subtasks were as follows: shape naming, number identification, producing a set, quantity discrimination, mental addition and subtraction, measurement vocabulary and oral addition. Except for number identification, with 20 items, the rest of the subtasks contained five items each. Based on the fact that the number identification subtask had a larger scale of measure, it was excluded from the calculation of reliability statistics.

Table 12 shows the Pearson correlations of the numeracy subtasks. Moderately high correlations (0.37) were observed between shape naming and producing a set, shape naming and quantity discrimination (0.38), shape naming and measurement vocabulary (0.36) and shape naming and oral addition (0.33). Pairings of producing a set and other subtasks also recorded moderately high correlations: between producing a set and quantity discrimination (0.51); producing a set and measurement vocabulary (0.36), and producing a set and oral addition (0.35). In the pairing of quantity discriminations and other subtasks, moderately high correlations were observed between quantity discrimination and measurement vocabulary (0.38); and between quantity discrimination and oral addition (0.39). The other moderately high correlations were observed between mental addition and subtraction and oral addition (0.34) and between measurement vocabulary and oral addition (0.33). The pairings between the remainder of the subtasks were observed to have weak correlations. Additionally, it worth noting that all the correlation coefficients for the numeracy subtasks were statistically significant ($p < .001$).

Table 12. Pearson correlations for numeracy subtasks

	Shape naming	Producing a set	Quantity discrim-ination	Mental addition and subtrac-tion	Measure-ment vocabu-lary	Oral addition
Shape naming	1.00					
Producing a set [a]	0.37***	1.00				
Quantity discrimi-nation	0.38***	0.51***	1.00			
Mental addition and sub-traction	0.16***	0.23***	0.25***	1.00		
Measure-ment vocabulary	0.36***	0.36***	0.38***	0.19***	1.00	
Oral addition [a]	0.33***	0.35***	0.39***	0.34***	0.33***	1.00

*p < .05, **p < .01, ***p < .001.
[a] Denotes non-school-readiness indicators.

Table 13 shows the Cronbach's alpha coefficients for the numeracy subtasks. The overall reliability was 0.75, which surpassed the acceptable threshold of 0.70. The highest alpha level was observed for the mental addition and subtraction subtask (0.75). Nevertheless, two subtasks had alpha levels below the 0.70 thresholds: producing a set (0.69) and quantity discrimination (0.68).

Table 13. Cronbach's alpha for numeracy subtasks

Subtask	Item-test correlation	Item-rest correlation	Alpha
Shape naming	0.65	0.47	0.71
Producing a set [a]	0.71	0.54	0.69
Quantity discrimination	0.73	0.58	0.68
Mental addition and subtraction	0.54	0.33	0.75
Measurement vocabulary	0.66	0.47	0.71
Oral addition [a]	0.69	0.51	0.70
Total			**0.75**

[a] Denotes non-school-readiness indicators.

4.1.3 Social-Emotional and Executive Functioning Analysis

Table 14 and ***Table 15*** show the Pearson correlations and Cronbach's alpha statistics for the socio-emotional and executive functioning subtasks of the DAT. As explained previously, the decision to combine these two subtasks for reporting the midterm results emanated from the fact that there are just two items in the socio-emotional subtask and four in the executive functioning subtask. Furthermore, the reported correlations were bivariate, thus not affected by the grouping.

Statistically significant ($p < .001$) bivariate correlations were observed for all measures, but with varying magnitudes. There was a moderately strong correlation of 0.56 between showing empathy and understanding feelings. Weak correlations of less than 0.30 were observed between pairs of all other subtasks except forward digit span and backward digit span (0.30). Table 14 shows the details of these findings.

Table 14. Pearson correlations for the six social-emotional and executive functioning subtasks

	Showing empathy	Under-standing feelings	Colour sort	Picture sort	Forward digit span	Backward digit span
Showing empathy [a]	1.00					
Understanding feelings [a]	0.56***	1.00				
Colour sort [a]	0.16***	0.23***	1.00			
Picture sort [a]	0.21***	0.23***	0.28***	1.00		
Forward digit span [a]	0.25***	0.24***	0.21***	0.21***	1.00	
Backward digit span	0.21***	0.21***	0.09***	0.16***	0.30***	1.00

*$p < .05$, **$p < .01$, ***$p < .001$.
[a] Denotes non-school-readiness indicators.

Table 15 shows the Cronbach's alpha coefficients for the combined socio-emotional and executive functioning subtasks. The overall alpha was observed to be 0.65; the highest was 0.64 for both the colour sort and backward digit span; and the lowest alpha level of 0.57 was observed in the understanding feelings subtask.

Table 15. Cronbach's alpha for the six social-emotional and executive functioning subtasks

Subtask	Item-test correlation	Item-rest correlation	Alpha
Showing empathy [a]	0.66	0.46	0.58
Understanding feelings [a]	0.68	0.48	0.57

Subtask	Item-test correlation	Item-rest correlation	Alpha
Colour sort [a]	0.54	0.30	0.64
Picture sort [a]	0.58	0.34	0.62
Forward digit span [a]	0.61	0.39	0.61
Backward digit span	0.54	0.30	0.64
Total			**0.65**

[a] Denotes non-school-readiness indicators.

4.1.4 School Readiness Index Items Analysis

As explained in Section 3.6, the School Readiness Index consists of 10 subtasks: backward digit span, rhyme, letter-name knowledge, letter-sound knowledge, initial sound identification, listening comprehension, shape naming, quantity discrimination, mental addition and subtraction, and measurement vocabulary. Recall that each of these subtasks assesses various learners' abilities, as outlined in ***Table 16:***

Table 16. School Readiness Index subtasks

Subtask	Description
Rhyme	Assesses learners' ability to listen to words and identify rhyming words
Letter-name knowledge	Assesses learners' ability to identify letters of the alphabet
Letter-sound knowledge	Assesses learners' knowledge of letter sounds
Initial sound identification	Assesses learners' ability to identify the beginning sound of a word
Listening comprehension	Assesses learners' ability to listen to a story and thereafter answer questions related to the story
Backward digit span	Assesses learners' ability to repeat the digits read out in reverse order – that is, backwards
Shape naming	Assesses learners' ability to identify shapes
Quantity discrimination	Assesses learners' ability to discriminate quantity
Mental addition and subtraction	Assesses learners' ability to perform addition and subtraction problems
Measurement vocabulary	Assesses learners' ability to identify size

The results in ***Table 17*** show statistically significant correlations for all the School Readiness Index subtasks. Moderately strong correlations of more than 0.30 were

observed between the pairing of quantity discrimination and other subtasks. Specifically, the following correlations with quantity discrimination were observed:

- 0.32 correlation with backward digit span
- 0.36 correlation with initial sound identification
- 0.35 correlation with listening comprehension
- 0.38 correlation with measurement vocabulary.

Other moderately strong correlations were observed between the following pairings: initial sound identification and backward digit span (0.39); initial sound identification and rhyme (0.30); listening comprehension and initial sound identification (0.39); shape naming and initial sound identification (0.34); shape naming and listening comprehension (0.36); measurement vocabulary and shape naming (0.36); and measurement vocabulary and quantity discrimination. Weak correlations of less than 0.30 were observed between the remaining pairs of School Readiness Index subtasks.

Table 17. Pearson correlations for the School Readiness Index subtasks

	Backward digit span	Rhyme	Letter-name know-ledge	Letter-sound know-ledge	Initial sound identifi-cation	Listening compre-hension	Shape naming	Quantity discrimi-nation	Mental addition and subtrac-tion	Measure-ment vocabu-lary
Backward digit span	1.00									
Rhyme	0.19***	1.00								
Letter-name knowledge	0.23***	0.18***	1.00							
Letter-sound knowledge	0.25***	0.22***	0.14***	1.00						
Initial sound identification	0.39***	0.30***	0.25***	0.52***	1.00					
Listening compre-hension	0.29***	0.32***	0.26***	0.30***	0.39***	1.00				
Shape naming	0.25***	0.24***	0.25***	0.26***	0.34***	0.36***	1.00			
Quantity discrimi-nation	0.32***	0.21***	0.27***	0.26***	0.36***	0.35***	0.38***	1.00		
Mental addition and subtraction	0.21***	0.06***	0.15***	0.19***	0.21***	0.21***	0.16***	0.25***	1.00	
Measurement vocabulary	0.27***	0.17***	0.20***	0.19***	0.26***	0.28***	0.36***	0.38***	0.19***	1.00

*p < .05, **p < .01, ***p < .001.

Table 18 shows the Cronbach's alpha coefficients for the School Readiness Index subtasks. The highest alpha level of 0.78 was observed on the mental addition and subtraction item. However, all the School Readiness Index subtasks had an alpha level of at least 0.75. The overall alpha level of 0.78 is above the acceptable threshold and denotes that these 10 items do indeed discern the school readiness of learners.

Table 18. Cronbach's alpha for the School Readiness Index subtasks

Subtask	Item-test correlation	Item-rest correlation	Alpha
Backward digit span	0.59	0.45	0.76
Rhyme	0.50	0.35	0.77
Letter-name knowledge	0.50	0.36	0.77
Letter-sound knowledge	0.57	0.44	0.76
Initial sound identification	0.69	0.58	0.75
Listening comprehension	0.65	0.53	0.75
Shape naming	0.62	0.50	0.76
Quantity discrimination	0.65	0.54	0.75
Mental addition and subtraction	0.45	0.30	0.78
Measurement vocabulary	0.57	0.43	0.76
Total			**0.78**

Table 19 shows the Cronbach's alpha for all[9] the subtasks of the DAT, which includes both the School Readiness Index and the indicators for research purposes. The item-test correlations show the correlation between a subtask and the composite score, and the item-rest correlations do the same between a subtask and the composite scores of all other subtasks. The results indicate that all subtasks had an alpha level of at least 0.85, which is above the acceptable level. The overall alpha level of 0.86 indicates that the DAT could discriminate between learners at different levels of school readiness and was thus valid.

Table 19. Cronbach's alpha for all subtasks

Subtask	Item-test correlation	Item-rest correlation	Alpha
Showing empathy^{r}	0.57	0.50	0.858
Understanding feelings^{r}	0.60	0.53	0.857
Colour sort^{r}	0.40	0.31	0.865

[9] We have purposefully omitted from this report the Pearson correlation output for the entire tool because it is too massive and detailed. Correlations are inferred based on the analysis at the subtask level.

Subtask	Item-test correlation	Item-rest correlation	Alpha
Picture sort[r]	0.43	0.35	0.864
Forward digit span[r]	0.53	0.45	0.860
Backward digit span[i]	0.54	0.46	0.859
Rhyme[i]	0.46	0.38	0.862
Expressive listening[r]	0.54	0.47	0.859
Letter-name knowledge[i]	0.43	0.35	0.864
Letter-sound knowledge[i]	0.52	0.45	0.860
Initial sound identification[i]	0.65	0.59	0.854
Sequencing[r]	0.53	0.45	0.860
Listening comprehension[i]	0.68	0.62	0.853
Shape naming[i]	0.59	0.52	0.857
Producing a set[r]	0.64	0.58	0.855
Quantity discrimination[i]	0.65	0.58	0.854
Mental addition and subtraction[i]	0.40	0.32	0.865
Measurement vocabulary[i]	0.55	0.47	0.859
Oral addition[r]	0.57	0.49	0.858
Total			**0.866**

i = School Readiness Index indicator (used for actual impact evaluation).
r = research indicator – i.e. item included in the evaluation for research purposes only.

5. FINDINGS

5.1 Impact on School Readiness

The Tayari ECD programme focuses on improving school readiness as measured by the School Readiness Index designed for Tayari. The research question that this report focuses on is whether Tayari can make a statistically significant impact on school readiness at low cost. ***Figure 2*** shows the first-year impact of Tayari from the baseline (January 2016) to the midterm (October 2016). At the baseline, school readiness in the control group was 20.1%, while it was 18.8% for the Tayari treatment group (Treatment 2). The control group increased its index outcomes from 20.1 to 33.6, while the treatment group increased from 18.8 to 37.2. The slopes of the lines show a much faster increase for the treatment group than for the control group during this period. While additional results will be needed at the endline to definitely say whether the Tayari

treatment had a statistically significant impact on school readiness, these results are encouraging for the first academic year of Tayari's implementation.

The causal effect of Tayari is estimated as an increase of 5.0% on school readiness over what the control increase is (p <.001), a 0.34 SD effect. This effect of 0.34 SD is larger than Tayari was expected to achieve after two full years of implementation. The increase of 5.0% on school readiness is larger than the negative relationship of having learning difficulties, which suggests that Tayari, as implemented even in its first year, was able to dramatically reduce the gaps for children with learning challenges and to increase their readiness for school in a meaningful way.

Figure 2. School Readiness Index results

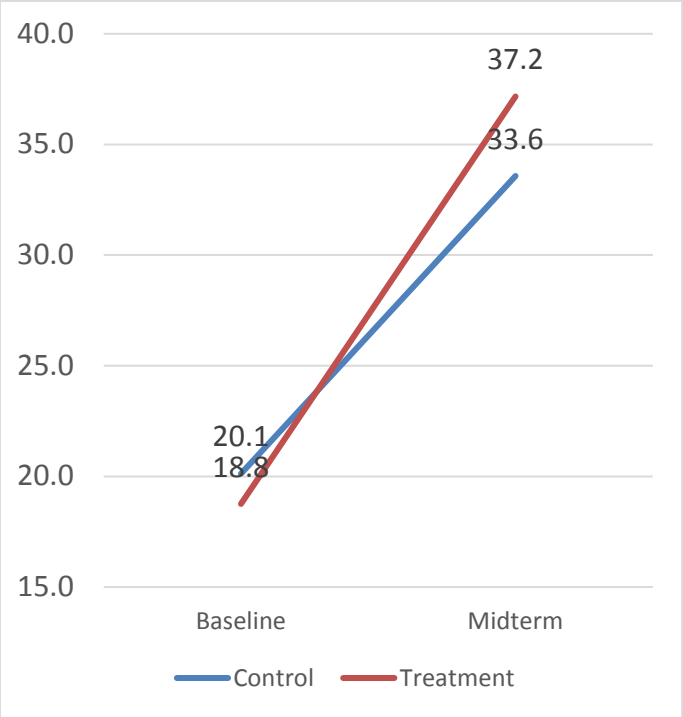
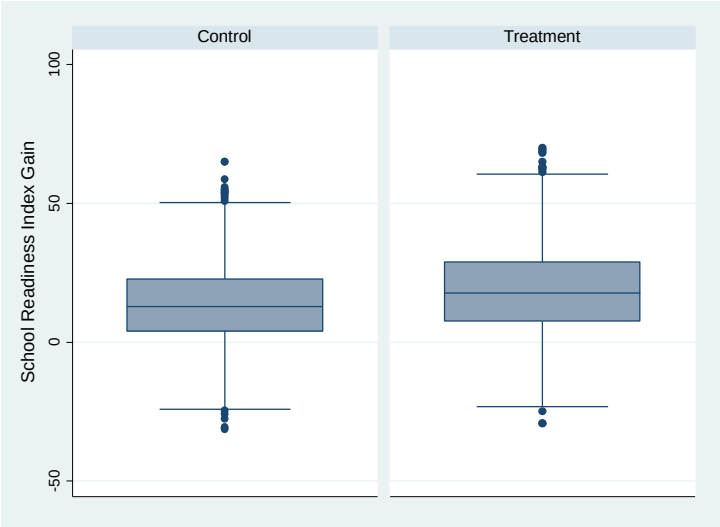


Figure 2 above shows the average change in school readiness. This average, however, hides heavy disparities in school readiness changes over time, within both treatment and control, as ***Figure 3*** shows. The range of gains in school readiness is wide, with a mean of 15.2% and a standard deviation of 14.8%. Given that wide range, it is worth noting that the effect of Tayari was felt throughout the distribution, with the 75th percentile, mean and 25th percentile of gains being quite different between treatment and control.

Figure 3. Boxplots of gains in School Readiness Index scores, control and treatment groups

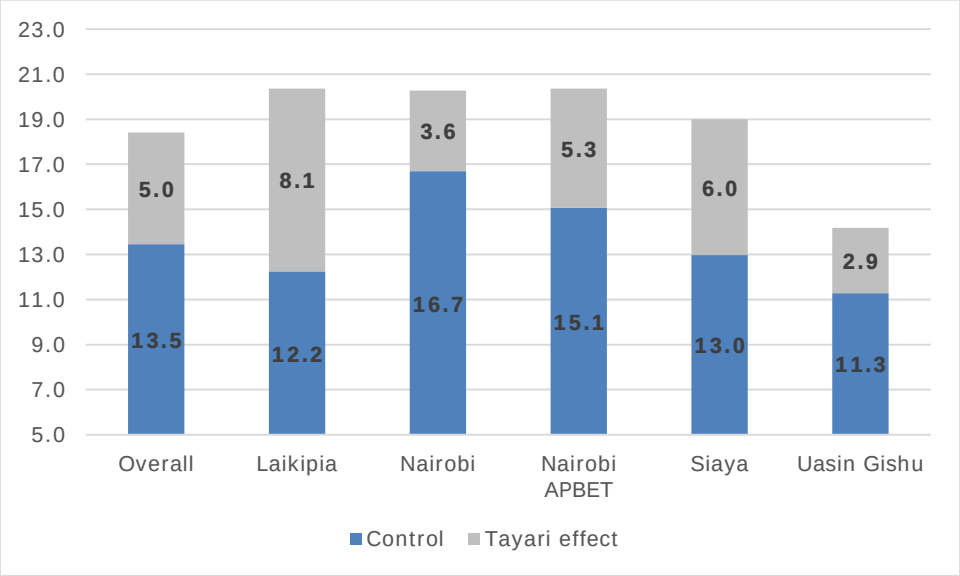


5.2 Impact on Readiness by County and Class

The Tayari longitudinal impact evaluation was not organised to be able to make statistically significant comparisons at the county level. However, given the larger-than-expected impact of Tayari at the midterm, we are able to share county-level comparisons on the impact of the programme for discussion. ***Figure 4*** presents the impact, by county, of the Tayari intervention. The blue bars indicate the gains in school readiness from the baseline to the midterm, while the grey bars show the additional causal impact of Tayari over and above the gains by the control group.

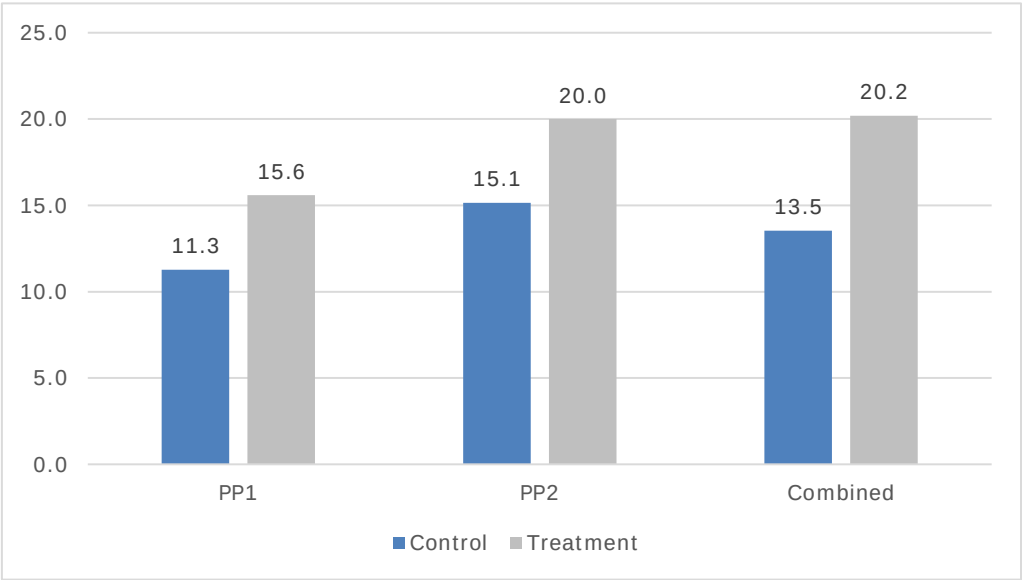
All four public school counties and the APBET schools in Nairobi appear to have had positive gains from Tayari, although given the modest sample sizes, not all of these comparisons are statistically significant, as expected from the design. The largest effects are 8.1% identified in Laikipia (*p*-value .01), 6.0% experienced in Siaya (*p*-value .03) and 5.3% in Nairobi APBET (*p*-value .05), with smaller, statistically insignificant gains of 3.6% for Nairobi public (*p*-value .47) and 2.9% for Uasin Gishu (*p*-value .14). These results are useful for each of the county teams to consider what strategies could be employed to increase the impact of Tayari by the end of the second year. Recall, however, that the research design was not structured to ensure county-level comparisons, so noting insignificant effects for some counties is not surprising.

Figure 4. Tayari impact on school readiness, by county (%)



While the analysis above focused on the impact of Tayari across class levels, the results below differentiate Tayari's impact by PP1, PP2, and Combined classrooms. ***Figure 5*** shows the longitudinal gains for each of these three groups. We found statistically significant impacts of Tayari for all three groups, with the largest gains felt in the Combined classrooms, where gains were 20.2% for treatment and 13.5% for control, a causal Tayari gain of 6.7 percentage points (*p*-value .01). The 4.9 percentage point gains in PP2 (20.0% for treatment, 15.1% for control) were statistically significant (*p*-value <.01) and slightly larger than the 4.3 percentage point gains for PP1 (15.6% for treatment, 11.3% for control) and were also statistically significant at the .03 level.

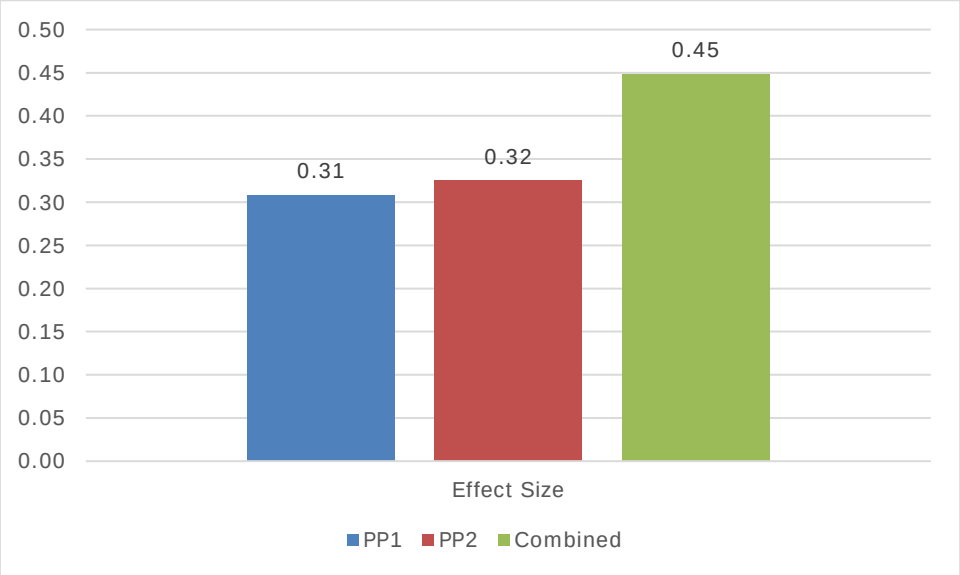
Figure 5. Gains in school readiness (%)



In order to assess whether the Tayari intervention was having a meaningful impact on learning outcomes, we were able to convert the longitudinal gains in Tayari by effect

sizes by dividing the effect by the pooled standard deviation. It is worth noting that Tayari expected to identify a 0.2 SD effect *after two years* of intervention. In fact, after the first year, Tayari's intervention effect magnitude was 0.31 SD for PP1, 0.32 SD for PP2 and 0.45 SD for Combined classrooms. These effect sizes are comparable to the effect of the PRIMR Initiative after the first year of the PRIMR intervention in literacy (Piper, Zuilkowski, & Mugenda, 2014) and larger than the effects identified in mathematics in the first year of the intervention (Piper et al., 2016). The overall effect size across the entire sample was 0.34 SD.

Figure 6. Tayari impact: Effect sizes (standard deviations)



5.3 Impact on the Percentage of Children at the School Readiness Cut-Off

At baseline, a School Readiness Index measure was developed from subtasks within the Kenya-specific DAT, as explained in Sections 3.5 and 3.6 above. In this section, we analyse data to examine if the initial eight months of the Tayari intervention had an impact on this School Readiness Index. We do so by computing the proportion of children reaching the following bands:

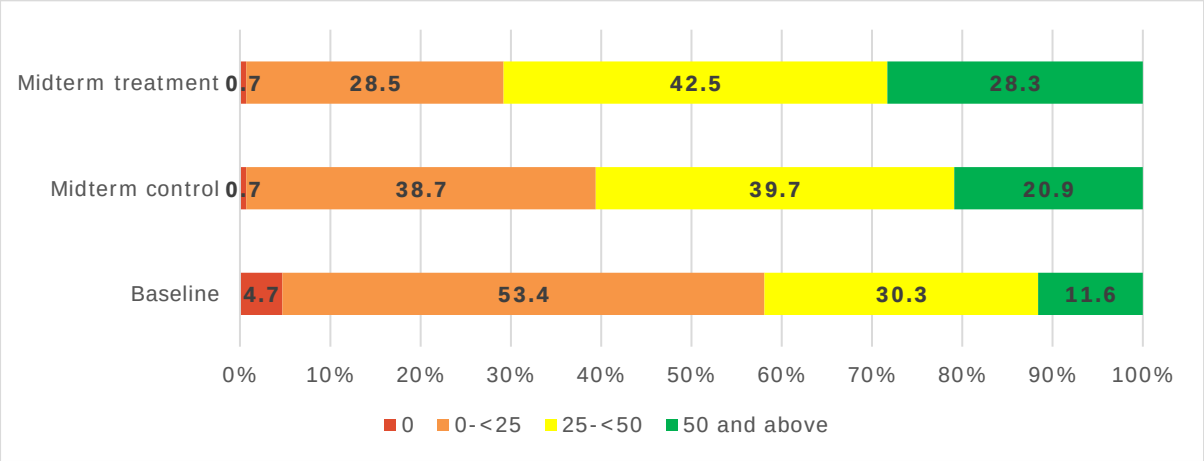
1. 0%
2. below 25%
3. between 25% and 50%
4. more than 50%.

We tentatively set the benchmark for the overall index to 50 and, as such, children attaining this benchmark would be said to be 'School Ready'. This choice of benchmark, however, will require MoE confirmation prior to the Tayari endline assessment. Note that we based this analysis on the longitudinal sample (children who were assessed at both baseline and midterm).

Figure 7 shows the proportion of children reaching various cut-off scores on school readiness. As indicated, there was a positive change in the trend in the School Readiness Index scores. Analysis showed that the treatment group recorded larger gains than the control group. Despite both groups recording a reduced proportion of children scoring 0% – from 4.7% to 0.7% – the proportion of children scoring 0% for the two treatment groups remained the same (at 0.7%). However, other benchmark bands showed that children reaching an index value of less than 25% were fewer in the treatment group (28.5%) as compared to those in the control group (38.7%). In the same vein, Figure 7 shows that in the treatment group, almost 70% of children had attained an index value of 25% or above, as compared to approximately 60% for the control group.

Figure 7 also shows that at baseline, 11.6% of the children reached the desired benchmark of 50 and above. At midterm, this proportion rose to 20.9% for the control group and 28.3% for the treatment group, depicting a score that was 7.4 percentage points higher for the treatment as compared to the control group.

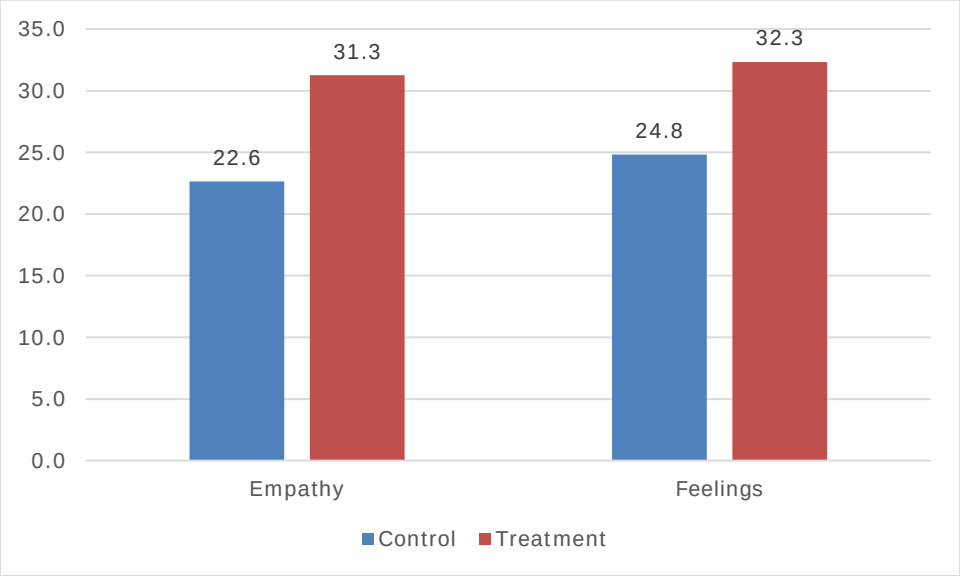
Figure 7. Benchmarks, by treatment group (%)



5.4 Impact by Subtasks

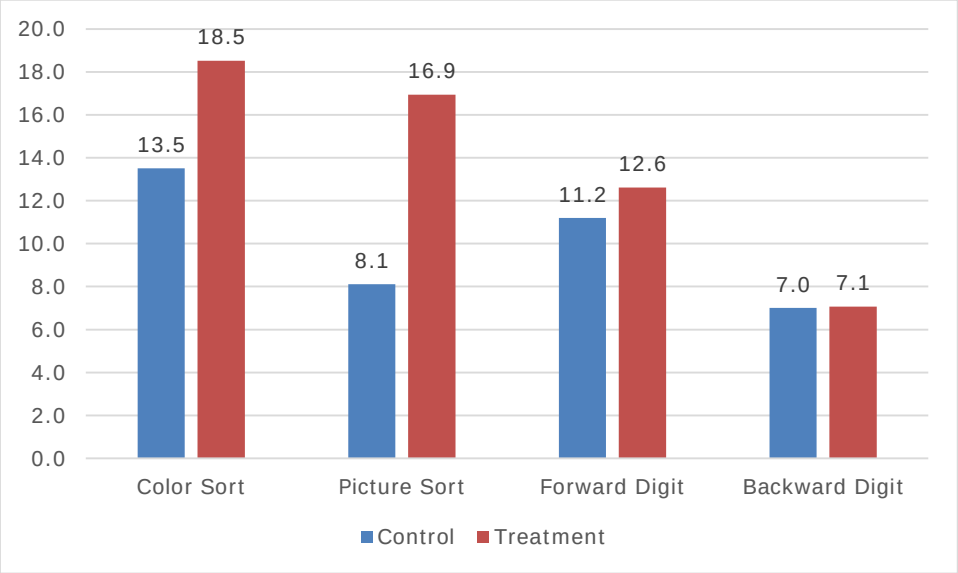
In the sections that follow, we present our findings on the longitudinal gains experienced by children in the Tayari sample. These are percentage correct scores for various categories of items, starting with the two socio-emotional tasks: empathy and feelings. We found, as shown in ***Figure 8,*** that children increased their empathy scores by 8.7 percentage points more in treatment schools (*p*-value .04) than in control schools. Similarly, the feelings scores increased by 7.5 percentage points more in treatment than control schools (*p*-value .02). These positive results were somewhat surprising, given that improved socio-emotional skills were not directly targeted by the programme but apparently can be supported in a more interactive classroom environment.

Figure 8. Gains in socio-emotional scores (%)



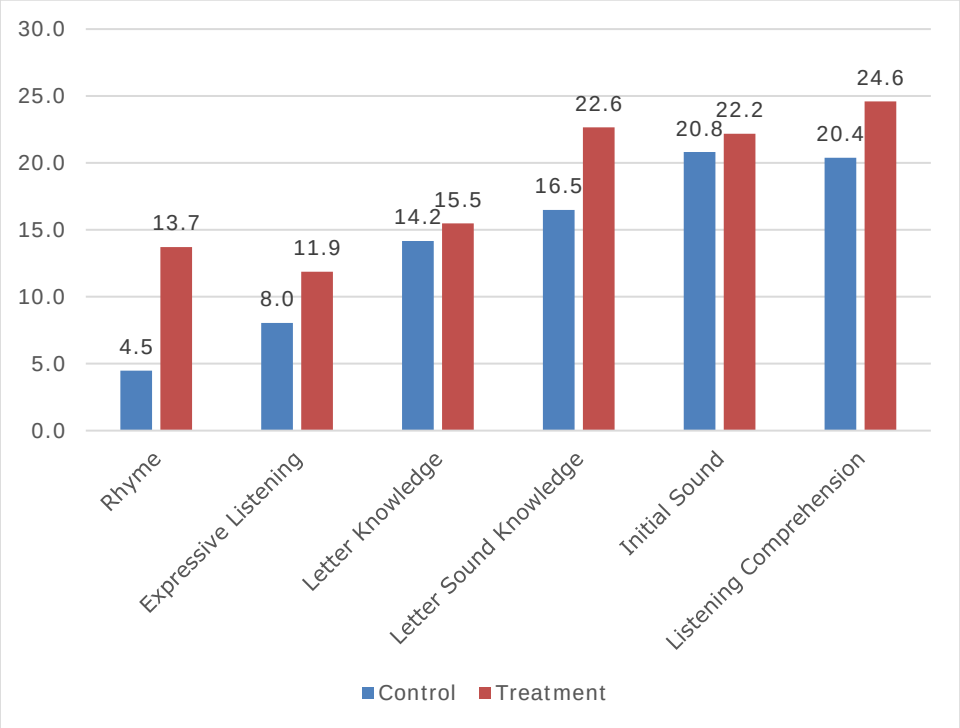
The Tayari longitudinal data presented an opportunity to investigate whether children in Tayari schools increased in their executive functioning and impulse control abilities (see ***Figure 9***). These skills are theorised to be related to learning outcomes in positive ways, and we wanted to understand whether the Tusome intervention had an impact on these skills even from pre-primary education. We found that the gains in colour sorting were 5.0 percentage points higher in treatment schools (*p*-value .22), which was a statistically insignificant difference. On the other hand, we found that on the picture sort task, which was Tayari's closest approximation to an executive functioning task, the treatment group had an increase of 8.8 percentage points more than the control group at the midterm (*p*-value .03). This increase seems meaningful at 0.22 SD, but we were somewhat surprised to see it, without initial explicit instruction on executive functioning being part of the project. For forward digit span, we saw no difference between treatment and control (*p*-value .63); similarly, for backward digit span, which was the closest approximation to an impulse control measure, there was no difference by treatment (*p*-value .97).

Figure 9. Gains in executive functioning (%)



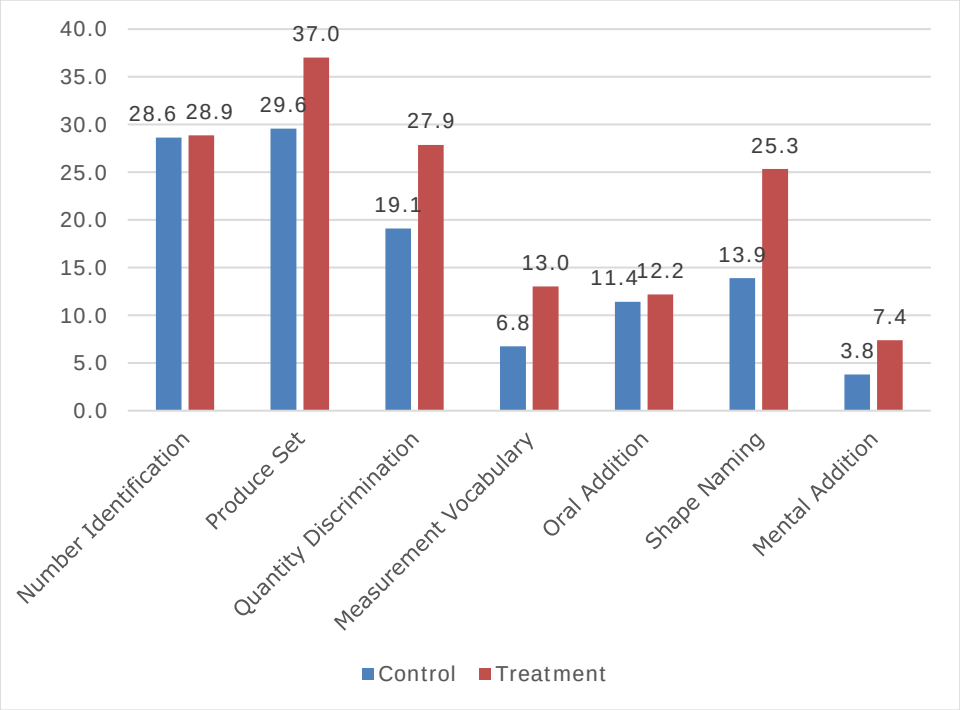
The language skills assessed in Tayari were quite broad (refer to ***Figure 10***). We found that for rhyme, the Tayari causal effect was a 9.2 percentage point increase in school readiness (*p*-value <.01). The expressive listening impact was 3.9 percentage points, statistically significant at the .10 level (*p*-value .08). Interestingly, the letter knowledge gain was not statistically significant (*p*-value .62), meaning that the control schools were also teaching letter names. Letter sounds had a meaningful increase of 6.1 percentage points, which was statistically significant (*p*-value .03). A phonemic awareness task investigating initial sound skills was not statistically significant (*p*-value .56), and neither was listening comprehension (*p*-value .19). This shows that several of the key tasks that Tayari was organised to support were increasing as a result of Tayari's language development focus, but not all of the tasks of interest responded initially to Tayari's intervention.

Figure 10. Gains in language skills (%)



Gains in numeracy skills in the Tayari data set showed some interesting comparisons ***(Figure 11).*** There was no effect on number identification (*p*-value .89). This is a skill that is already taught in existing Kenyan ECDE centres, so the Tayari programme had no differential effect on simply identifying numbers. The number sense skill called producing a set did show a causal effect of 7.4 percentage points (*p*-value .02). Measuring vocabulary had no statistically significant difference (0.19 SD), nor did oral addition (0.74 SD). The applied skill of shape naming had a meaningful impact of 11.4 percentage points on outcomes (*p*-value <.001), while mental addition's modest gain for Tayari was statistically insignificant (*p*-value .20).

Figure 11. Gains in numeracy (%)



5.5 Predictors' relationships with school readiness

The discussion of the findings above focused on the various elements of the school readiness impacts within the Tayari programme. In this section we present the relationships identified between various student level predictors and Tayari outcomes. These are parameter estimates from regression models where we held treatment constant and tried to determine the isolated relationship between these areas and Tayari school readiness. Note that these relationships are not causal but relational.

Detailed information on impact by subtask, class level and gender appears in ***Appendix B*** of this report.

We included several models with a variety of predictors, but ***Figure 12*** below presents only the relationships that were statistically significant. We found negative relationships between having polio (-22.2%) and Tayari school readiness (*p*-value <.001), speaking a mother tongue at home (-14.1%) and school readiness (*p*-value <.001),[10] having a hearing impairment (-12.1%) and school readiness (*p*-value <.01), and having learning difficulties (-4.8%) and school readiness at the .10 level (*p*-value .09). That these student characteristics were negatively related to learning outcomes is unsurprising; what is disturbing is the magnitude of the relationships. This suggest that urgent attention is required on how Kenya's ECDE system responds to those who have hearing problems, physical diseases, and learning disabilities so that they are able to have a fair chance to learn the basics that ECDE provides. Currently, the disadvantages that these

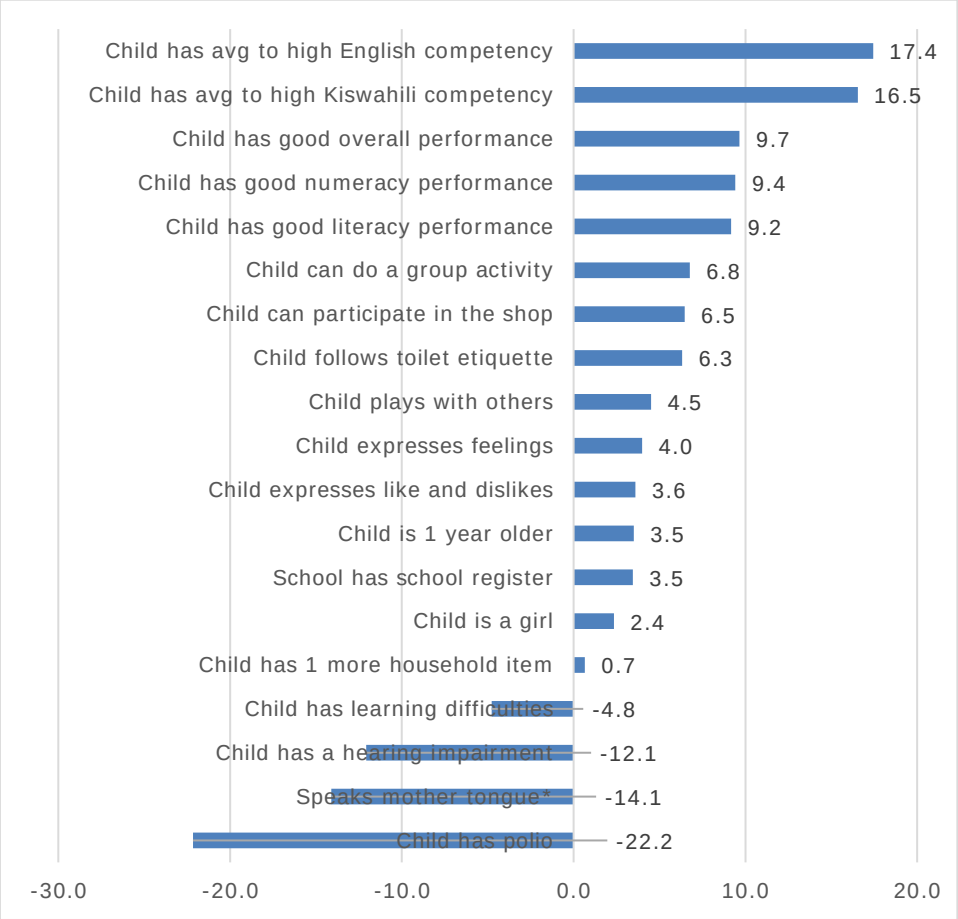
[10] The mother tongue regression model was fit only on the control schools and controlled for the county in which the student lived.

children have are actually as large as, or larger than, the impact of the entire Tayari programme.

Several other indicators suggest higher school readiness. Each additional household item (from the socioeconomic measure) was related to 0.7 percentage points higher school readiness outcomes (*p*-value .06). Being one SD above the mean in socio-economic status meant that those children were 1.4 percentage points more ready for school, even controlling for the Tayari effect. Girls performed 2.4 percentage points better on school readiness (*p*-value <.01), which confirmed consistent findings in Kenya for girls outperforming boys in primary school (Piper & Mugenda, 2012). For every one year older children were than the mean age for a grade, they performed 3.5 percentage points higher (p-value .01). This finding is somewhat surprising and more research should be undertaken to understand why older children did better. Many interpersonal skills were related to school readiness, including 6.8 percentage points higher if the child had the ability to do group activities (*p*-value <.001), 6.5 percentage points higher if the child could participate in buying and selling activities at the shop (*p*-value <.001), 4.5 percentage points higher if the child played with other children (*p*-value <.001), 4.0 percentage points higher if the child expressed feelings (*p*-value <.001) and 3.6 percentage points higher if the child could express likes and dislikes (*p*-value <.001). Interestingly, given another treatment group's focus on health activities at the school, as well as the focus of all of Tayari materials on health behaviours, we found that children who followed toilet etiquette had 6.3 percentage points higher school readiness (*p*-value <.001).

These results also showed a close correlation with parental indications of children's success. For those that had good overall performance, good numeracy performance and good literacy performance, their school readiness scores were 9.7, 9.4 and 9.2 percentage points higher (*p*-value <.001, <.001, <.001). Even stronger effects were found for the language skills of children, with 17.4 percentage points higher school readiness if the children were identified as having average to high English competency (p-value <.001), and 16.5 percentage points higher school readiness if the child was identified with average to high Kiswahili competency (*p*-value <.001). This suggests that what drives success in Kenya's ECD programmes is prior language skill, which is somewhat disappointing given the belief that ECD can ameliorate home background differences. It might be that the skills that children begin with are so large that the 5.0 percentage points impact of Tayari after one year could not do the entire job required to support children who entered ECD significantly behind in language and other academic skills.

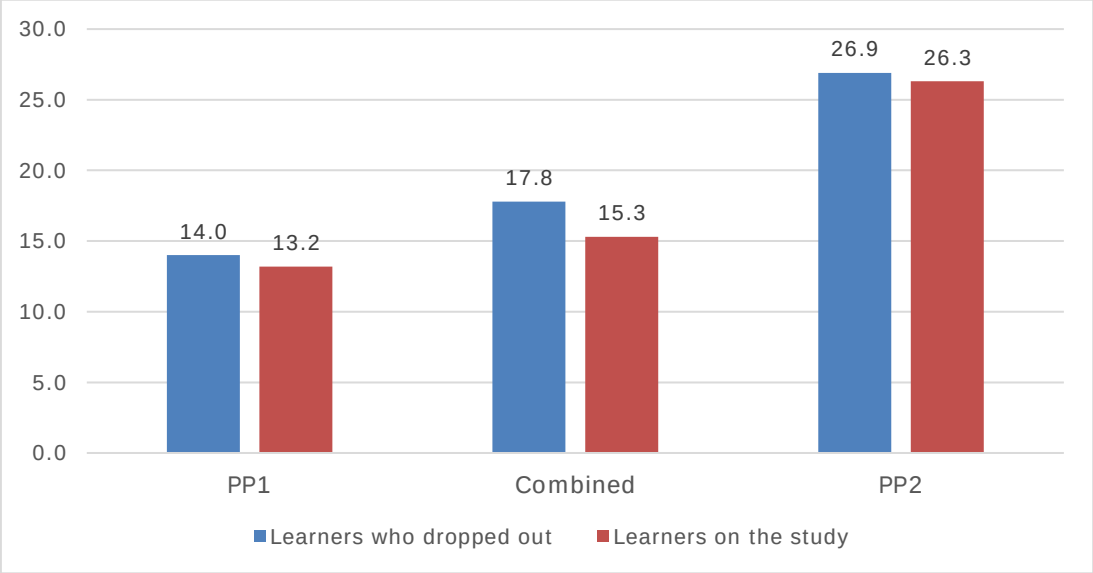
Figure 12. Relationships with school readiness



5.6 Effect of Attrition on the Sample

As discussed in Section 3.8 above, 11.2% of learners assessed at baseline dropped out of the study by midterm. Below, we analyse the difference between the sample of learners who dropped out against those who remained. As shown in ***Figure 13,*** however, there was no statistically significant difference between the dropped-out and the retained learners for the PP1, PP2, or Combined samples at baseline. At PP1, those who dropped out had an average School Readiness Index score of 14.0 while among those who remained, the index value was 13.2 (p=0.40). Similarly, learners who dropped out at PP2 had an average index score of 26.9 compared to 26.3 recorded for those who remained (p=0.69). For the combined group of ECDE centres, on average, learners who dropped out had an average index score of 17.8 as compared to 15.3 for those who remained (p=0.06).

Figure 13. School Readiness Index scores by attrition status (%)



6. RECOMMENDATIONS AND CONCLUSIONS

Based on the encouraging findings from the Tayari midterm assessment, we present the following recommendations:

- **Consider the magnitude of the first-year results from Tayari.** The results presented in this report show effects that are commensurate with the impact of PRIMR after its first year, and larger than many or most other large-scale ECDE improvement programmes. This finding suggests that Tayari has the potential to be a successful intervention at scale.
- **Continue to implement pilots at medium scale, through government mechanisms.** The Tayari design ensures that implementation of the programme occurs through the government systems at both the national and county levels. The training, the coaching and the support to Tayari are being implemented by Kenyan officials and field-based officers. Thus, the Tayari results are likely indicative of what would happen if the programme were scaled up to additional sites.
- **Consider scaling up the Tayari intervention.** While we recommend waiting until the endline assessment in October 2017 to determine whether these encouraging results continue, we suggest that Kenya begin now to consider scaling up the Tayari programme more broadly based on the encouraging findings.
- **Improve School Readiness measures.** The reliability statistics presented in this report show that the Tayari assessment tools meet the accepted research standards for reliability and discrimination. Given that few school readiness tools have been utilised in developing countries, it is particularly important for the Tayari team to work on refining and improving the new set of tools that Tayari is experimenting with from the MELQO activities funded by CIFF. In particular, some of the executive function and socio-emotional items did not

perform as well as hoped, and additional work should be undertaken to better assess these portions of school readiness.

- **Review the selection of and results from APHRC public control schools.** While these longitudinal results were largely similar to APHRC's results, children in the APHRC public control schools improved their school readiness results by significantly more than the control children in the longitudinal samples presented here. The results for the APHRC external evaluation control schools also increased by more than those for the APBET control schools, which – based on APHRC's work in the sector and RTI's work in APBET schools – is not consistent with results in any other study over the past several years. In order to obtain a fair analysis of Tayari's impact in public schools, these odd results should be carefully considered, and potentially additional schools included in the control sample to determine whether these results are consistent at the endline assessment. In addition, given the differences in scoring on the school readiness index for the APHRC external evaluation and the longitudinal study, we suggest joint training of assessors going forward for the January 2017 and October 2017 assessments.

REFERENCES

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology, 31*(3), 597–606. https://doi.org/10.1177/0038038597031003015

Aron, A., Coups, E. J., & Aron, E. N. (2013). *Statistics for psychology* (6th ed.). Boston, MA: Pearson.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https:/dx.doi.org/10.1007/BF02310555

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ, 314,* 572. https://doi.org/10.1136/bmj.314.7080.572

Kwayumba, D., & Piper, B. (2016). *Baseline assessment report for the Tayari Early Childhood Development and Education Programme.* Prepared for the Children's Investment Fund Foundation, Contract No. 338 (RTI Project No. 0214471). Research Triangle Park, NC: RTI International.

Ministry of Education, Science and Technology, Republic of Kenya. (2014). *2014 basic education statistical booklet.* Produced with assistance from UNICEF. https://www.kicd.ac.ke/images/ICT/2014BasicEducationStatisticalBooklet.pdf

Murunga, J. (2015). Devolving early childhood development education in Kenya: Policy challenges and opportunities. *International Journal of Education and Research, 3*(2), 611–620. http://www.ijern.com/journal/2015/February-2015/49.pdf

Ngaruiya, S. (2006). *Pre-school education and school readiness: Kenya's experience.* Conference paper presented at the 2006 Biennale of the Association for the Development of Education in Africa (ADEA), Session C3-2, Preparing Children for School: The Role of Holistic Practices in the Early Years of Learning. http://www.adeanet.org/clearinghouse/en/pre-school-education-and-school-readiness-kenyas-experience

Pence, A. (2011). Early childhood care and development research in Africa: Historical, conceptual, and structural challenges. *Child Development Perspectives, 5*(2), 112–118. https://doi.org/10.1111/j.1750-8606.2011.00165.x

Piper, B., Kwayumba, D. N., Oyanga, A., & Jepkemei, E. (2015). *The Primary Math and Reading (PRIMR) Initiative DFID/Kenya Rural Expansion Programme, Bungoma and Machakos endline study.* Prepared for DFID/Kenya under contract 202657-108. Research Triangle park, NC: RTI International. https://globalreadingnetwork.net/eddata/primary-math-and-reading-primr-initiative-dfidkenya-rural-expansion-programme-bungoma-and-0

Piper, B., & Mugenda, A. (2012). *The Primary Math and Reading (PRIMR) Initiative: Baseline report.* Prepared under the USAID Education Data for Decision Making (EdData II) project, Task Order No. AID-623-M-11-00001 (RTI Task 13). Research Triangle Park, NC: RTI International. Retrieved from http://pdf.usaid.gov/pdf_docs/pa00hx75.pdf

Piper, B., & Mugenda, A. (2016). *USAID/Kenya Primary Math and Reading (PRIMR) Initiative: Endline impact evaluation, revised edition.* Prepared under the United States Agency for international Development (USAID) Education Data for Decision Making (EdData II) project, Task Order No. AID-623-M-11-0001 (RTI Task 13). Research Triangle Park, NC: RTI International. https://globalreadingnetwork.net/eddata/usaidkenya-primary-math-and-reading-primr-initiative-endline-impact-evaluation-revised

Piper, B., Ralaingita, W., Akach, L., & King, S. (2016). Improving mathematics outcomes in Kenya: The impact of the Primary Math and Reading (PRIMR) Initiative. *Journal of Development Effectiveness.* Published online 17 Mar 2016. http://dx.doi.org/10.1080/19439342.2016.1149502

Piper, B., Zuilkowski, S. S., & Mugenda, A. (2014). Improving reading outcomes in Kenya: First-year effects of the PRIMR Initiative. *International Journal of Educational Development, 37,* 11–21. https://doi.org/10.1016/j.ijedudev.2014.02.006

Republic of Kenya. (2006). *National Early Childhood Development Policy Framework 2006.* Nairobi: Ministry of Education. http://planipolis.iiep.unesco.org/upload/Kenya/KenyaECDPolicyFramework.pdf

Republic of Kenya. (2012a). *Kenya Population and Housing Census 2009: Analytical report on education, Vol. IX.* Nairobi: Kenya National Bureau of Statistics. http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&id=109:population-and-housing-census-2009&Itemid=599

Republic of Kenya (2012b). *Sessional Paper No. 14 of 2012. A policy framework for education and training. Reforming education and training in Kenya.* Nairobi: Ministry of Education and Ministry of Higher Education, Science Technology. http://planipolis.iiep.unesco.org/upload/Kenya/Kenya_Policy_framework_for_education_and_training_Sessional_Paper_July_5,_2012.pdf

Republic of Kenya. (2013). *Basic Education Act of 2013.* Nairobi: Parliament of Kenya. http://www.education.go.ke/index.php/downloads/file/122-basic-education-act-2013

Republic of Kenya. (2014). *National Education Sector Plan (NESP) of 2013–2018. Volume 1: Basic education programme rationale and approaches.* Nairobi: Ministry of Education. http://www.education.go.ke/index.php/downloads/file/83-

national-education-sector-plan-volume-one-basic-education-programme-rationale-and-approach

Republic of Kenya. (2015). *Sessional Paper No. 1 of 2015 on the National Policy on Elimination of Child Labour.* Nairobi: Ministry of Labour, Social Security, and Services. http://www.ilo.org/wcmsp5/groups/public/---africa/---ro-addis_ababa/---ilo-dar_es_salaam/documents/genericdocument/wcms_532338.pdf

Schaie, K. W. (1983). The Seattle Longitudinal Study: A twenty-one year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64–135). New York: Guilford. https://sharepoint.washington.edu/uwsom/sls/Documents/1980/SLS_21_exp_psy_intell.pdf

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

United Nations Children's Fund (UNICEF). (2001). *The state of the world's children 2001. Early childhood.* New York: UNICEF. https://www.unicef.org/sowc/archive/ENGLISH/The%20State%20of%20the%20World's%20Children%202001.pdf

United Nations Children's Fund (UNICEF). (2012). Early childhood education and school readiness [Web page]. New York: UNICEF. https://www.unicef.org/education/bege_61627.html

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2015). *Education for All 2015 national review report, Kenya.* Prepared for the World Education Forum, Incheon, Republic of Korea, 19–22 May 2015. Nairobi: Ministry of Education, Science and Technology, Republic of Kenya. http://unesdoc.unesco.org/images/0023/002316/231637e.pdf

World Health Organization (WHO) & UNICEF. (2012). *Early childhood development and disability: A discussion paper.* Geneva: WHO Press. http://apps.who.int/iris/bitstream/10665/75355/1/9789241504065_eng.pdf

APPENDIX A: PROGRAMME DESIGN DETAILS BY COUNTY

Table A1. Programme design: Nairobi County APBET

Location		2016			2017			2018
Sub-county	Cluster	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Kasarani Embakasi Njiru	Komboni Chokaa Soweto							
Dagoretti Langata Kasarani	Congo Kianda Mwiki							
Kasarani Ruaraka	Baraka Mathare North							
Embakasi Njiru	Gitari Kware Mowlem							
Westlands Kasarani Embakasi	Gichagi Korogocho Tushauriane							
Njiru Dagoretti	Njiru Ngando							

Location		2016			2017			2018
Sub-county	Cluster	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Dagoretti	Gatina							
Langata	Riruta							
Embakasi	Makina							
Kasarani	Tassia							
Starehe	Kariobangi North							
	Kiamaiko							

Table A2. Programme design: Nairobi County

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Westlands Makadara	Parklands Viwandani							
Kamkunji Starehe Westlands	Eastleigh Juja Road Kilimani							
Embakasi	Kayole							
Langata Kasarani	Nairobi West Ruaraka							
Starehe Makadara Kasarani	Central Buruburu Kahawa							
Dagoretti	Waithaka							

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Dagoretti	Riruta							
Kamkunji	Bahati							
Njiru	Dandora							
Langata	Karen							

Table A3. Programme design: Laikipia County

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Laikipia East Laikipia North	Igwamiti Sosian							
Laikipa North Laikipia West	Mugodo East Githiga Marmanet							
Laikipia West	Salama							
Laikipia East	Umande Ngobit							
Laikipia East Laikipia North	Thingitu Nanyuki Segera							
Laikipia West	Ol Moran							
Laikipia West Laikipia North Laikipia East	Rumuruti Township Mukogodo West Tigithi							

Table A4. Programme design: Uasin Gishu County

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Eldoret West Wareng	Kesses Sirikwa							
Eldoret West Wareng	Soy Kapyemit Tulwet							
Eldoret East	Moiben							
Wareng Eldoret East	Kaptagat Pioneer							
Eldoret West Wareng Eldoret East	Moi's Bridge Timboroa Kapsoya							
Eldoret East	Meibek							

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Eldoret East Eldoret West Wareng	Karona Chepkoilel Sugoi Koisagat Cheptiret Kapsiret							

Table A5. Programme design: Siaya County

Location		2016			2017			2018
Sub-county	Zone	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Training and Support	Training and Support + Books / Teachers' Guides	Training and Support + Books / Teachers' Guides + Health	Control
Ugenya Rarieda	West Yimbo North Uyoma							
Rarieda Gem Ugenya	West Uyoma West Gem Central Sakwa							
Ugenya	West Ugenya							
Rarieda	South Uyuma East Asembo							
Alego Usonga Gem	South East Alego Central Alego Yala Township							
Gem Ugenya	North Sakwa							
Ugunja Gem Ugenya Alego Usonga	Ugunja South Gem East Ugenya South Sakwa Ukwala Usonga							

APPENDIX B: ADDITIONAL DETAIL ON GAINS IN SCORES, BY SUBTASK, CLASS AND GENDER

Table B2. Gains in scores by class

Subtask	Class	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
School Readiness Index	PP1	11.3	1.33	15.6	1.36	0.03	14.01	0.309
	PP2	15.1	1.30	20.0	0.93	0.01	14.95	0.325
	Combined	13.5	1.25	20.2	2.08	0.01	14.84	0.449
Rhyme	PP1	6.4	0.03	17.4	0.03	0.02	0.31	0.353
	PP2	3.7	0.02	12.1	0.03	0.03	0.31	0.270
	Combined	2.5	0.02	9.9	0.03	0.08	0.31	0.236
Expressive listening	PP1	10.3	0.03	15.2	0.02	0.19	0.33	0.150
	PP2	5.5	0.02	9.9	0.03	0.20	0.38	0.113
	Combined	9.9	0.03	10.1	0.02	0.96	0.29	0.006
Letter-name knowledge	PP1	12.8	0.03	15.1	0.02	0.54	0.26	0.085
	PP2	15.1	0.03	13.8	0.02	0.74	0.35	-0.037
	Combined	14.6	0.03	21.7	0.04	0.15	0.29	0.241

Subtask	Class	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
Letter-sound knowledge	PP1	15.6	0.03	17.2	0.03	0.75	0.28	0.058
	PP2	18.4	0.03	25.5	0.03	0.12	0.40	0.177
	Combined	13.4	0.03	26.9	0.04	0.01	0.30	0.451
Initial sound identification	PP1	12.0	0.02	13.9	0.02	0.48	0.26	0.070
	PP2	30.4	0.02	28.5	0.02	0.60	0.38	-0.050
	Combined	13.9	0.04	22.6	0.03	0.11	0.31	0.274
Listening comprehension	PP1	19.5	0.03	27.5	0.02	0.03	0.32	0.254
	PP2	22.4	0.03	23.5	0.03	0.81	0.33	0.035
	Combined	17.0	0.04	20.9	0.04	0.47	0.30	0.129
Sequencing	PP1	11.7	0.05	17.8	0.05	0.41	0.52	0.119
	PP2	17.3	0.05	28.4	0.05	0.16	0.56	0.198
	Combined	6.6	0.06	19.4	0.04	0.09	0.45	0.285
Shape naming	PP1	13.1	0.03	21.9	0.04	0.06	0.40	0.220
	PP2	15.0	0.03	28.2	0.03	0.00	0.46	0.288
	Combined	12.9	0.02	24.6	0.04	0.01	0.40	0.292
Producing a set	PP1	33.9	0.05	40.5	0.03	0.28	0.40	0.163
	PP2	25.3	0.03	31.0	0.04	0.23	0.38	0.150
	Combined	31.6	0.04	47.2	0.05	0.03	0.43	0.363

Subtask	Class	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
Quantity discrimination	PP1	18.6	0.03	20.1	0.03	0.75	0.44	0.034
	PP2	17.0	0.03	31.3	0.03	0.00	0.40	0.355
	Combined	25.6	0.03	35.7	0.05	0.09	0.44	0.230
Mental addition and subtraction	PP1	3.3	0.02	7.1	0.03	0.29	0.30	0.127
	PP2	3.9	0.03	6.2	0.03	0.62	0.40	0.058
	Combined	4.4	0.02	11.8	0.02	0.02	0.33	0.224
Measurement vocabulary	PP1	4.2	0.03	13.2	0.04	0.09	0.46	0.194
	PP2	3.7	0.04	9.5	0.05	0.37	0.45	0.128
	Combined	19.8	0.05	23.4	0.05	0.61	0.46	0.077
Oral addition	PP1	7.9	0.02	5.5	0.02	0.42	0.26	-0.094
	PP2	14.0	0.03	17.5	0.03	0.40	0.36	0.097
	Combined	11.8	0.02	11.5	0.02	0.93	0.31	-0.008
Showing empathy	PP1	21.8	0.03	33.4	0.05	0.09	0.47	0.248
	PP2	24.8	0.04	31.6	0.03	0.20	0.44	0.156
	Combined	18.8	0.04	24.9	0.03	0.21	0.41	0.146
Understanding feelings	PP1	24.2	0.04	32.2	0.04	0.24	0.52	0.155
	PP2	26.0	0.03	33.2	0.03	0.06	0.47	0.154
	Combined	23.0	0.04	29.9	0.03	0.19	0.47	0.146

Subtask	Class	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
Colour sort	PP1	14.1	0.02	26.5	0.03	0.00	0.37	0.336
	PP2	9.7	0.04	11.8	0.03	0.71	0.31	0.069
	Combined	22.2	0.03	20.3	0.03	0.64	0.40	-0.049
Picture sort	PP1	3.6	0.02	17.1	0.04	0.01	0.42	0.320
	PP2	9.0	0.02	16.5	0.04	0.19	0.38	0.198
	Combined	15.0	0.03	17.8	0.03	0.56	0.41	0.071
Forward digit span	PP1	8.9	0.02	11.9	0.02	0.37	0.31	0.095
	PP2	11.2	0.03	11.6	0.04	0.92	0.31	0.016
	Combined	15.9	0.04	17.2	0.05	0.83	0.37	0.035
Backward digit span	PP1	2.7	0.01	4.2	0.02	0.40	0.13	0.119
	PP2	11.8	0.01	10.1	0.02	0.56	0.25	-0.068
	Combined	3.4	0.01	4.6	0.02	0.56	0.16	0.073

Table B2. Gains in scores by gender

Subtask	Gender	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
School Readiness Index	Boy	12.9	1.18	17.9	0.94	0.00	14.63	0.345
	Girl	14.1	0.82	18.9	1.03	0.00	14.80	0.327
Rhyme	Boy	3.0	0.02	13.4	0.02	0.00	0.31	0.330
	Girl	6.1	0.02	14.0	0.03	0.03	0.31	0.256
Expressive listening	Boy	8.9	0.02	12.8	0.02	0.16	0.34	0.110
	Girl	7.0	0.02	10.9	0.02	0.18	0.36	0.109
Letter-name knowledge	Boy	13.7	0.02	15.4	0.01	0.47	0.30	0.059
	Girl	14.7	0.03	15.5	0.02	0.83	0.32	0.024
Letter-sound knowledge	Boy	14.8	0.02	21.2	0.02	0.06	0.32	0.202
	Girl	18.5	0.02	24.2	0.02	0.07	0.37	0.155
Initial sound identification	Boy	18.3	0.02	20.0	0.02	0.57	0.32	0.052
	Girl	23.7	0.02	24.6	0.01	0.74	0.35	0.024
Listening comprehension	Boy	20.9	0.02	26.4	0.02	0.12	0.32	0.168
	Girl	19.8	0.02	22.7	0.02	0.38	0.32	0.091
Sequencing	Boy	13.2	0.04	24.7	0.04	0.08	0.53	0.218
	Girl	13.5	0.04	21.5	0.05	0.26	0.53	0.148

Subtask	Gender	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
Shape naming	Boy	15.4	0.02	25.1	0.02	0.00	0.42	0.227
	Girl	12.2	0.02	25.6	0.03	0.00	0.44	0.307
Producing a set	Boy	29.1	0.02	36.6	0.03	0.05	0.40	0.188
	Girl	30.1	0.02	37.4	0.03	0.05	0.40	0.184
Quantity discrimination	Boy	19.7	0.03	28.5	0.02	0.02	0.41	0.214
	Girl	18.5	0.02	27.1	0.03	0.05	0.43	0.200
Mental addition and subtraction	Boy	4.1	0.02	6.4	0.02	0.44	0.34	0.066
	Girl	3.4	0.02	8.5	0.03	0.15	0.37	0.139
Measurement vocabulary	Boy	8.5	0.03	11.1	0.04	0.59	0.47	0.057
	Girl	4.8	0.03	15.1	0.04	0.07	0.45	0.226
Oral addition	Boy	9.6	0.02	12.7	0.02	0.28	0.33	0.094
	Girl	13.4	0.02	11.6	0.01	0.44	0.32	-0.057
Showing empathy	Boy	23.6	0.03	31.6	0.03	0.14	0.47	0.170
	Girl	21.6	0.03	30.9	0.03	0.04	0.43	0.220
Understanding feelings	Boy	26.3	0.04	32.3	0.03	0.17	0.48	0.124
	Girl	23.2	0.03	32.4	0.03	0.05	0.49	0.188
Colour sort	Boy	15.1	0.03	17.1	0.03	0.70	0.36	0.056
	Girl	11.7	0.02	20.0	0.03	0.02	0.34	0.247

Subtask	Gender	Control		Treatment		Programme impact		
		Gain	Std. error	Gain	Std. error	p-value	Pooled SD	Effect size
Picture sort	Boy	8.5	0.02	16.4	0.04	0.09	0.40	0.199
	Girl	7.7	0.02	17.5	0.03	0.02	0.40	0.244
Forward digit span	Boy	11.9	0.02	12.0	0.03	0.98	0.33	0.002
	Girl	10.4	0.02	13.3	0.02	0.34	0.32	0.091
Backward digit span	Boy	5.2	0.01	5.8	0.02	0.79	0.19	0.027
	Girl	9.0	0.01	8.5	0.02	0.84	0.22	-0.024