Playgroups for Inclusion - Experimental Evaluation and Study of Implementation

Final Report

(Expected submission Date: 31-07-2017)

Playgroups for Inclusion - Experimental Evaluation and Study of Implementation

Final Report

July 2017

Prepared for:

European Commission

DG Employment, Social Affairs and Inclusion

Programme for Employment and Social Solidarity - PROGRESS (2007-2013), under the call VP/2013/012: Call for Proposals for Social Policy Experimentations Supporting Social Investments.

Grant Agreement: VS/2014/0418, Playgroups for Inclusion

Prepared by:

Grupos Aprender, Brincar, Crescer

University of Coimbra, FPCE (Ed. 2), Praça Dom Dinis (Antiga Farmácia Ex-HUC), Gabinete 2.11
3000-143 Coimbra, Portugal

ISCTE – University Institute of Lisbon, DPSO (Ala Autónoma)
Avenida das Forças Armadas, Edifício
1649-026 LISBOA Portugal

Playgroups for Inclusion - Experimental Evaluation and Study of Implementation
Final Report

Authors

Prepared by:

M. Clara Barata

Joana Alexandre

Bruno de Sousa

Catarina Leitão

Vanessa Russo

Contributing Authors:

Pedro Cunha, Hélder Pais e Liliana Marques

Lúcia Santos e Joana de Freitas Luís

Júlia Santos e Marisa Horta

Alexandra Marques

Suggested Citation:

Barata, M.C., Alexandre, J., de Sousa, B., Leitão, C., Russo, V. (2016). Playgroups for Inclusion: Experimental Evaluation and Study of Implementation, Final Report. Portugal: University of Coimbra & ISCTE-IUL.

Acknowledgments

This project is funded by the European Union Programme for Employment and Social Solidarity - PROGRESS (2007-2013) under the grant agreement VP/2013/012/0577. This programme is implemented by the European Commission. It was established to financially support the implementation of the objectives of the European Union in the employment, social affairs and equal opportunities area, and thereby contribute to the achievement of the Europe 2020 Strategy goals in these fields.

We would like to thank all participating families and children. We also want to express our appreciation to the personnel in all public, private and semi-private entities who served as partners in this project. We also wish to thank all the members of the Playgroups for Inclusion implementation team, as well as the members of the consortium Playgroups for Inclusion, including the coordination by the Directorate General of Education (DGE), and consortium partners Fundação Bissaya Barreto (FBB), Calouste Gulbenkian Foundation, ACM-High Commission for Migration, I. P, University of Coimbra and ISCTE - University Institute of Lisbon. Finally, we are thankful to Luisa Canto e Castro Loura and Joaquim Santos at the Ministry of Education (DGEEC) for access to important data and statistical support.

Table of Contents

Figures:

Tables:

Executive Summary

The Playgroups for Inclusion or "Grupos Aprender, Brincar, Crescer" (GABC, Groups where children Learn, Play and Grow) was a pilot project which aimed to develop, test, and disseminate an innovative educational program for Early Childhood Education and Care (ECEC), targeting children aged 0-4 and their caregivers not participating in any of the currently available ECEC services in Portugal. The Inclusion focus was threefold:

1. Playgroups for inclusion aimed to provide a service for children and caregivers not participating in currently available ECEC services in order to reduce developmental gaps in cognitive and social domains, the likelihood of future school failure and social exclusion during compulsory schooling, while also increasing the quality of caregiving environment and caregivers' mental health;

2. By focusing on children and caregivers belonging to minority groups (such as the Roma), and families that were recently unemployed and underemployed, Playgroups for inclusion aimed to break inter-generational cycles of exclusion and social disadvantage, while increasing social cohesion, participation and intercultural dialogue in the communities.

3. Playgroups for inclusion aimed to serve the unemployed and underemployed by empowering them to actively participate in the development of a new social and education program, and train then to become playgroup facilitators and community mobilizers;

To reach these objectives, the consortium Playgroups for Inclusion designed a comprehensive intersectoral stakeholder approach, with a communication and dissemination plan built in throughout the project for sustainability, buy-in and possible scale-up, if the effectiveness of the program was demonstrated.

Research evidence about playgroups provides indication that such services improve a range of outcomes for children, such as language, cognition, and behavioral skills (Deutscher, Fewell, & Gross, 2006), and for parents, such as enhancement of the quality of parent-child interactions (Evangelou, Brooks, & Smith, 2007), and increased facilitation of children's learning (Hackworth et al., 2013). However, weaknesses in the design of these studies limit the internal and external validity of findings. To date,

Playgroups for Inclusion is the first randomized controlled trial of a playgroups-only program. Playgroups for Inclusion also aimed to be one of the first social program experimentations in education in the country.

In the past few decades there has been substantial development in program evaluation research, as well as methods for statistical data analysis of program impact. In the U.S., the Institute for Educational Sciences funds and supervises large-scale experimental and quasi-experimental trials of educational programs. However, in the majority of European and other countries, causal methods for program evaluation are still seldom applied to educational programs at scale, and most of the empirical evidence developed by the Ministries of Education does not answer the question "What is the true impact (or cause-effect relationship) of an educational program in participants' development and/or achievement?" This empirical problem is due to the fact that a considerable number of program evaluation studies in education use only qualitative data, small samples, and lack a control group which makes these inferences about the effect of the program unlikely.

In current educational research, even when these problems are overcome by collecting adequate quantitative data in groups affected and not affected by the program, and comparing these groups, the fact that participation in most educational programs depends on choices made by parents, teachers, legislators or other stakeholders make it hard to attribute program differences to the actual programs. These choices make the participation in educational programs the product of a process of self-selection, instead of being randomly determined (Murnane and Willett, 2010; Shadish, Cook, and Campbell, 2002). As such, the variation in educational programs is potentially correlated with other determinants of educational achievement, such as teachers' expectations or parents' investment, producing biased results of the effects of the program.

Because ECEC policies have the highest return on investment of all educational policies (Cunha, Heckman, Lochner, & Masterov, 2005), it is essential that we know the true impact of these policies. The best way to solve this empirical problem is to randomly assign participants to ECEC or a control group, and then evaluate their outcomes. In Playgroups for Inclusion, children and caregiver dyads were randomized within each district where Playgroups were implemented (Aveiro, Coimbra, Lisboa, Porto,

Setúbal) to two conditions: the *Playgroups for Inclusion* intervention group, which was offered a 10 month program, and the control group, which received a brief version of the program (3 months) in the following year[1]. In the experimental study, and in accordance to inclusion goals, we explicitly examined whether impacts varied by child ethnicity, and caregiver employment status[2].

The Playgroups for Inclusion consortium also proposed a comprehensive study of program implementation. The study of program implementation is as crucial as a carefully rigorously designed impact evaluation of that same intervention (Durlak & DuPre, 2008). Because the program Playgroups for Inclusion was implemented in 5 different geographical regions, and although the implementation procedures were somewhat standardized, we expected considerable differences in implementation practices across regions. Moreover, the diversity among participating families, among playgroups, and the community settings in which playgroups were housed, were assumed to determine differences in implementation practices, and were taken into account in this study. Through a careful, embedded study of implementation of the program, we could better ensure that observed outcomes were related to program implementation variables. In particular, quality, sustainability and attendance were given attention considering extant literature on their role as key aspects for playgroups effectiveness (see Walker et al., 2011). In complement to the experimental trial, the implementation study also allowed to gather the perspective of multiple stakeholders (caregivers, supervisors, and facilitators) on the program, and across time. Finally, in our study of implementation, and in accordance with the program's inclusion goals, we also examined whether family characteristics broadly predicted implementation patterns.

Study Design:

1. Randomized-controlled trial of children-caregiver dyads to experimental – Phase 1- and Control – Phase 2 - groups.
2. Embedded study of the implementation of the project.

[1] Treatment on the treated analyses were not implemented because families in the control condition were not participating in similar (group) activities, and therefore it was not possible to collect analogous attendance data. In replacement of TOT estimates, we have estimated non-experimental impacts for families and children receiving a higher dosage (i.e. attended more sessions) of the program.

[2] Our initial goals also included the explicit targeting of families belonging to migrant families. Although these were not excluded from the program, they were also not targeted. In consequence, no explicit analysis was made of the effects for these families.

This report aims to present the objectives, participants, procedures, activities and results of the evaluation of the project Playgroups for Inclusion, namely the experimental and implementation studies undertaken by the teams of WorkPackage 2 (WP2) at the University of Coimbra and ISCTE – University Institute of Lisbon (ISCTE-IUL). The WP2 evaluation team was coordinated at the national level by 3 Professors (including the PI), a doctoral-level research assistant for the experimental impacts team, and a master's level research assistant for the implementation study team.

Main objectives of the study:

1. Estimate the true, experimental impact of Playgroups for Inclusion on participants' outcomes;
2. Determine whether impacts vary by child ethnicity, and caregiver employment status.
3. Describe program implementation variables, namely quality, sustainability and attendance.
4. Study the implementation of Playgroups for Inclusion considering the perspective of different stakeholders (caregivers, supervisors, and facilitators).
5. Examine whether family characteristics broadly predicted implementation patterns.

Overall the study aimed to address the following research questions:

What was the true, experimental impact (or cause-effect relationship) of Playgroups for Inclusion in the Home environment and caregiving practices?

What was the experimental program impact in children's cognitive development?

What was the experimental program impact in children's temperament and behavior?

What was the experimental program impact in caregiving goals, values and aspirations?

What was the experimental program impact in caregivers' psychological distress?

What was the experimental program impact in caregivers' labor market participation?

What were the experimental program impacts in community participation outcomes?

Did experimental impacts vary by child ethnicity and caregiver employment status?

What were the initial expectations of participating caregivers, facilitators and supervisors about playgroup sessions?

What were the reasons stated by caregivers to enroll in this new service?

Which program activities were delivered during the 10 months of intervention? To what extent were the activities fully implemented?

To what extent were caregivers' and children's needs met?

What were the experiences for caregivers, facilitators and supervisors during the 10 months of program?

What was the level of caregiver involvement in playgroup?

What were the levels of family attendance, and how were these related to family characteristics?

What were caregivers, facilitators and supervisors' suggestions to improving family attendance?

What were the main perceived barriers to attendance by caregivers? How did these perceived barriers compare to the perceptions of facilitators?

What was the level of quality in playgroup sessions?

What were the perceptions of caregivers and supervisors regarding the quality of playgroups?

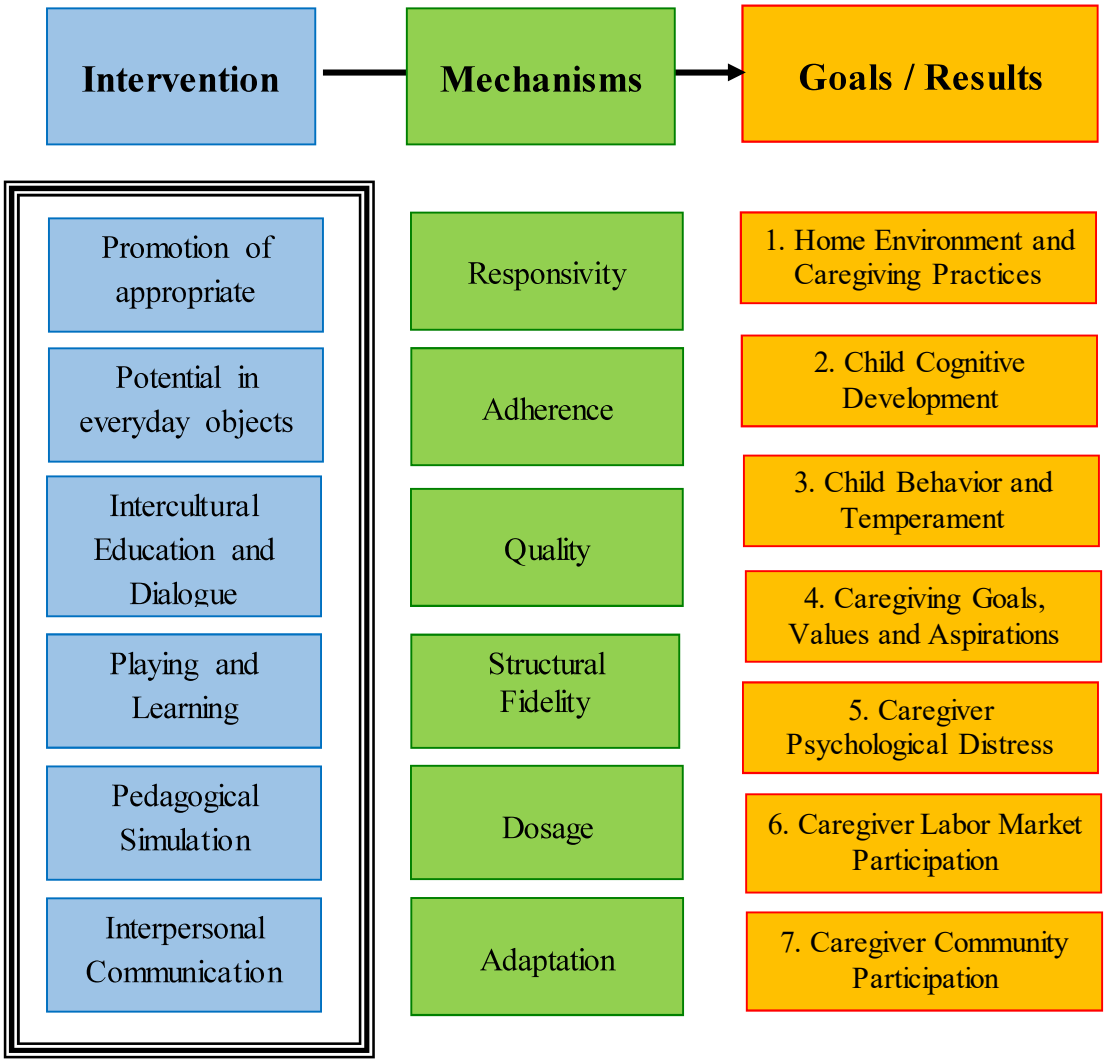
Did implementation patterns vary by family characteristics?

What was the impact of Playgroups for Inclusion for those who received a higher dosage (i.e. attended more sessions) of the program?

The first chapter of the report refers to the overall study goals and purposes. We briefly describe the program Playgroups for Inclusion or "Grupos Aprender, Brincar, Crescer" (GABC), and provide some theoretical and empirical background for the evaluation program proposed by WP2. We also contextualize this intervention in the Portuguese National Mandate for ECEC, and finally, present study objectives and research questions.

The second chapter is dedicated to the experimental evaluation of the program Playgroups for Inclusion. Outcomes for the evaluation of impact of Playgroups for Inclusion were selected according to a carefully designed and frequently reviewed Theory of Change. A Theory of Change is a conceptual tool that allows teams to examine the congruence between the object of study, and the proposed research design(s), evaluation measures, analysis plan, etc. (Anderson, n.d.; Connell & Kubisch, 1998; Weiss, 1995). It is a particularly efficient tool to align expectations of providers and evaluators regarding community interventions (Buitrago, 2015).

Theory of Change for Playgroups for Inclusion



This Theory of Change was initially based on extant literature on playgroups and ECEC services, and then reviewed in accordance to the intentionality discussed by the intervention team in common meetings. In the Playgroups for Inclusion, Theory of Change hypotheses of effect sizes were expressed in the rank order of outcomes, i.e. larger effects were expected for Home environment and Caregiving Practices, Child Cognitive Development, and Child Temperament and Behavior; smaller effects were expected for Caregiving Goals, Values, and Aspirations, and the smallest effects were expected for the remaining three domains. In its final version, the Theory of Change specified impacts of playgroups on a set of expected main and secondary outcomes at the domain and subdomain level organized by domain importance.

To estimate how many participants should be included in the sample in order to have a good chance at detecting the effects of the Playgroups for Inclusion Project, we made a careful power analysis estimating Minimum Detectable Effects (MDE) for child-level outcomes. In specific, and taking into account various assumptions for the type of design, preliminary power analyses to detect child or family-level outcomes indicated that we should focus on recruiting 940 to 950 families in total for the experimental trial. This number divided by 5 districts meant approximately 188 to 190 families across the two conditions (intervention and control) per district, or approximately 19 families across the two conditions per playgroup.

To meet recruitment and project goals, recruitment procedures in the pretest phase included essentially three strategies: 1) recruitment through referrals from entities in the community, government and social support network; 2) direct recruitment of eligible families through a strong and purposeful dissemination strategy; 3) and recruitment by referral by participating families (snowball sampling). This three-pronged strategy was designed to facilitate achieving the proposed sample size, while also reaching out to the proposed target families (i.e. children belonging minority groups, such as the Roma, and children from families that are recently unemployed and/or underemployed). The first strategy resulted in a network of 947 entities that worked directly with families (such as national/regional/local-level policy and decision-makers, social and economic partners, and NGOs). We requested their cooperation in all recruitment efforts. Among the entities/institutions contacted, 94 were community

health centers, 24 were child protection services, and 48 were entities that were part of the social welfare ministry.

Our pretest sample included 415 dyads – caregiver and children aged 0 to 4, not participating in ECEC services. In this sample, 63% of the families were referred by entities, only 30% of caregivers were employment, 14% of children were Roma or Romani, 8% were of "Other" ethnicity[3], and only 4% (of the total number of children) had attended some form of ECEC prior to enrollment in the program. The diversity of the sample speaks to the success of the comprehensive stakeholder approach, and initial communication and dissemination plan. However, this indicated that the study was considerably underpowered since the pretest.

After the pretest, 225 families were randomly assigned to the intervention condition (10 months of Playgroups for Inclusion) and 190 to the control condition (3 months of the same program one year later). There were no significant differences between the intervention and control conditions in all but one of the pretest assessments and demographic characteristics. These results suggest that the randomization was done with integrity, and that the pretest sample could provide the necessary confidence in the validity of the impact estimates.

Study Sample – Experimental Trial

Power estimates required 940 to 950 families in total

A three-pronged strategy for recruitment was used, including a comprehensive stakeholder approach with 947 entities working with families, and a strong and purposeful communication and dissemination plan

Pretest sample was diverse and included 415 dyads – caregiver and children

225 families were assigned to the intervention condition and 190 to the control condition

Posttest (final) sample of 257

Final power estimates indicated detectable MDEs in the range of 0.31 to 0.43 (medium).

Of the pretest sample, 62% (N=257, 115 control families and 142 intervention families) were assessed at posttest, which further limited the statistical power to examine experimental impacts (i.e. final power estimates indicated detectable MDEs in the medium range of 0.31 to 0.43, which is seldom found in ECEC program impacts).

[3] Children in the "Other" ethnicity group were reported to be Black, Negro, "Mestiço", Arab, Muslim, Latino, Hispanic, and others. Frequency of each category was very low. In order to avoid identification of participants, and in keeping with ethical standards and requirement, these ethnic profiles were compiled in one single group with a neutral label.

Methods – Experimental Study

Experimental estimates use multilevel models that account for pretest levels and covariates with random effects for playgroups and fixed effects for districts; robustness checks with OLS regression using Huber-White corrections

Adjusted impact estimates for attrition using nonresponse propensity weights.

Pooled data across the five districts.

Attriters and non-attriters had similar proportions of families in the intervention group. Moreover, intervention and control groups included in the posttest sample were still equivalent in pretest assessments and demographic characteristics. However, non-attriters were significantly different from attriters on 5 out of 18 (33%) pretest variables, namely non-attriters had a significantly larger percentage of families directly recruited (41%), caregivers had higher rates of employment, higher household incomes, and higher levels of education than attriters (28%), making the posttest sample more affluent, and the estimates of impact potentially biased.

To estimate experimental impacts, we compared outcomes for 115 control to 142 intervention families, caregivers and children using multilevel linear regression models that account for pretest levels and other covariates, as well as for the nesting of caregivers and children within playgroups by using playgroup random effects (Bloom, Richburg-Hayes, & Black, 2007; Hedges & Hedberg, 2007; Raudenbush, Martinez, & Spybrook, 2007) and district fixed effects. As a further robustness check, we also used an alternative model specification – ordinary least squares regression with correction of the standard errors (Huber-White) for playgroup clustering. We used nonresponse probability weights to account for attrition, and reduce nonresponse bias. For all analyses, we pooled data across the five districts.

Our results are summarized by presenting the effect size for each main and secondary outcome domain. Effect sizes indicate the magnitude of the Playgroup for Inclusion effect regardless of the instrument or method used, and are of prime interest for policy decisions.

Results of the experimental trial so far indicate largely non-significant mixed experimental impacts of the intervention for participating children, caregivers and families. Effect sizes for main impact domains for children and caregiver outcomes

were small (-0.150to 0.305). Effect sizes for secondary impacts were small to medium (-0.396 to 0.270).

Effect Sizes and Statistical Significance for Outcomes of the Experimental Evaluation of Playgroups for Inclusion

- Main outcomes -

Outcome	Effect Size	Sig.
Home environment and Caregiving Practices		
Responsiveness	0.086	0.537
Acceptance	-0.103	0.494
Involvement (under 24mo)	0.084	0.589
Child Cognitive Development		
Hearing and Language / Language (C)	0.028	0.597
Performance (E)	0.121	0.001**
Child Temperament and Behavior		
Negative Affect	0.104[a]	0.508
Effortful Control	0.114	0.499
Caregiving Goals, Values and Aspirations		
Connectedness Goals	0.305	0.020*
Childcare Values	-0.045	0.780
Expectations	-0.147	0.205
Aspirations	-0.135	0.209
Labor Market Participation		
Active Job Search	-0.150	0.164
Entrepreneurship	-0.086	0.422

Note. ~ p<.10, * p<.05, ** p<.01, *** p<.001. The effect size was computed by dividing the estimated adjusted difference between groups by the standard deviation of the indicator for the comparison group. Non-shaded lines indicate main domains of impact, and shaded lines indicate secondary domains according to the Theory of

Effect Sizes and Statistical Significance for Outcomes of the Experimental Evaluation of Playgroups for Inclusion

- Secondary Outcomes -

Outcome	Effect Size	Sig.
Home environment and Caregiving Practices		
Academic Stimulation (above 24mo)	0.203	0.257
Child Cognitive Development		
Locomotor (A)	0.002	0.971
Practical Reasoning (F)	-0.206	0.052~
Child Temperament and Behavior		
Duration of Orienting	0.270	0.303
Distress to Limitations	0.000[a]	0.999
Fear	0.396[a]	0.369
Caregiving Goals, Values and Aspirations		
Achievement Goals	-0.226	0.098~
Self-Maximization Goals	-0.044	0.694
Education Values	-0.029	0.825
Work Values	0.061	0.705
Psychological Distress		
Psychological Distress (K6)	0.264[a]	0.040*
Psychological Distress (K6) - clinical level	0.174[a]	0.164
Labor Market Participation		
Job Training	-0.068	0.623
Community Participation		
Perceived competence	-0.062	0.540
Critical Awareness	-0.060	0.593
Quantity of Intergroup contact	-0.084	0.444
Quality of Intergroup contact	0.222	0.140
Quantity of international friends	-0.025	0.868
Frequency of contact with international friends	-0.232	0.127

Note. ~ p<.10, * p<.05, ** p<.01, *** p<.001. The effect size was computed by dividing the estimated adjusted difference between groups by the standard deviation of the indicator for the comparison group. Non-shaded lines indicate main domains of impact, and shaded lines indicate secondary domains according to the Theory of Change. [a]Positive effects indicate benefits for the GABC families or children, expect where indicated.

Accounting for the multilevel structure of the data, pretest scores, child and caregiver characteristics, **three significant program effects were detected** as follows:

Playgroup children scored significantly higher (1.5 points on an observed range of 25 to 74 points) than Control children in **Performance**, a measure of children's

developing ability to reason through manual and visuospatial problems, including speed of working and precision, as measured by the Griffiths Mental Development Scales.

Playgroup caregivers were (15%) significantly more likely than Control caregivers to endorse **Connectedness Goals**, i.e. qualities that emphasized children's ability to maintain positive, harmonious relationships with others.

Playgroup caregivers self-reported significantly higher levels (1.1 points on an observed range of 0 to 22 points) than Control caregivers of **Psychological Distress** in Ron Kessler's K6.

In addition, accounting for the multilevel structure of the data, and child and caregiver characteristics, **two trend-level program effects were detected** as follows:

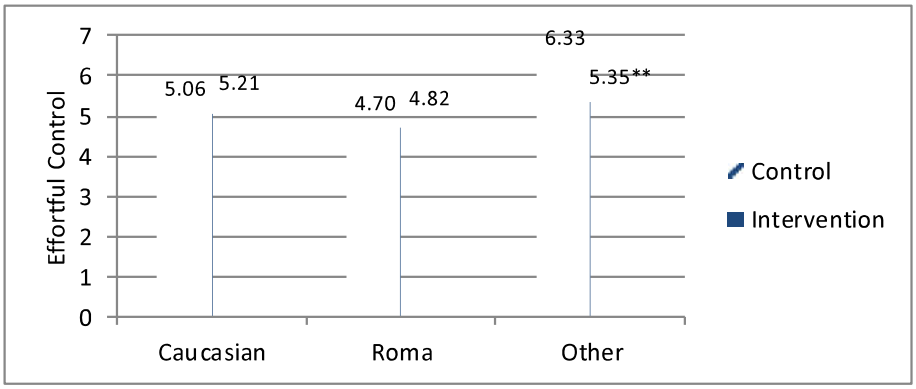
Playgroup children scored lower at trend level (1.2 points on an observed range of 0 to 25 points) than Control children in **Practical Reasoning**, a measure of children's ability to solve practical problems, understanding of basic mathematical concepts and understanding of moral issues, as measured by the Griffiths Mental Development Scales.

Playgroup caregivers were (10%) less likely at trend level than Control caregivers to endorse **Achievement Goals**, i.e. qualities related to children's academic performance or intellectual development.

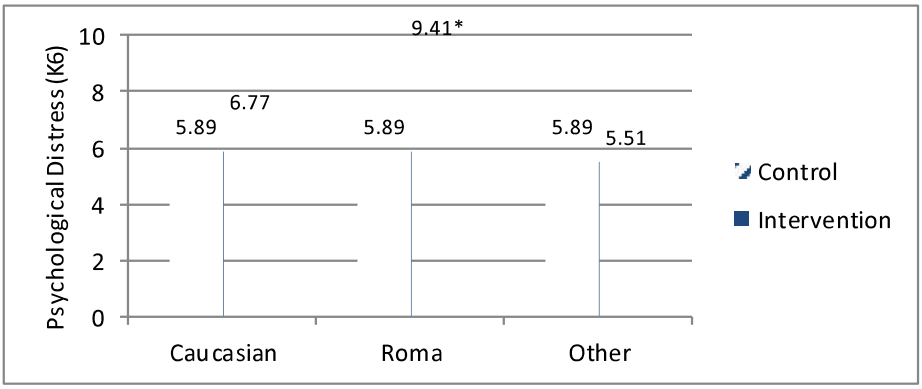
In terms of subgroup effects, and accounting for the multilevel structure of the data, pretest scores, child and caregiver characteristics, we found that **program effects varied significantly by subgroup (ethnicity and employment status) for three subdomains**, as follows:

Playgroup children in the "Other" Ethnicity scored lower (-1.1 points in an observed range of 2.4 to 6.6) in **Effortful Control**, a caregiver report measure of the child's ability to inhibit poor behavior, and mobilize attention resources to regulate behavior

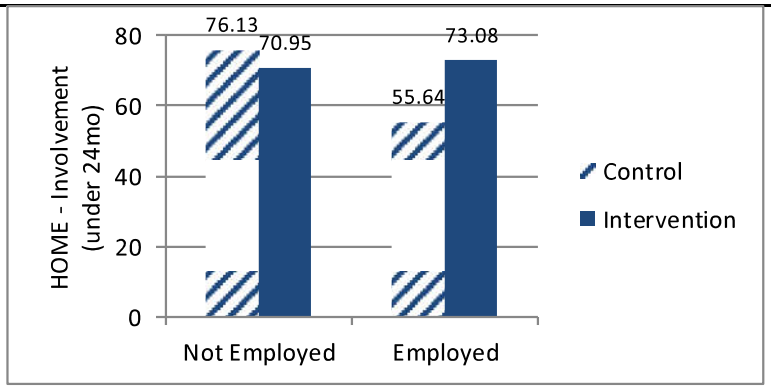
and emotions, when compared to their Control counterparts of the same ethnicity group, while Caucasian and Roma children saw (albeit small and non-significance) increases to their Effortful Control skills.



Roma caregivers in the intervention groups presented significantly higher levels (2.6 points in an observed range of 0 to 22) of Psychological Distress, a general term that is used to describe unpleasant feelings or emotions that impact level of functioning, than their Control counterparts of the same ethnicity, particularly in comparison to those of intervention and control Caucasian caregivers.

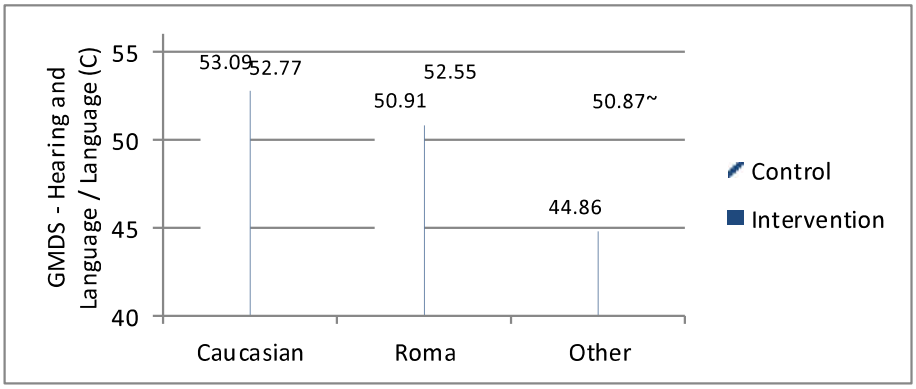


Playgroup caregivers that were employed presented a positive difference (17.44% in an observed range of 0 to 100) in Involvement, a measure of how the caregiver interacts physically with the child as measured by the HOME, to their Control counterparts that were also employed, than unemployed caregivers.

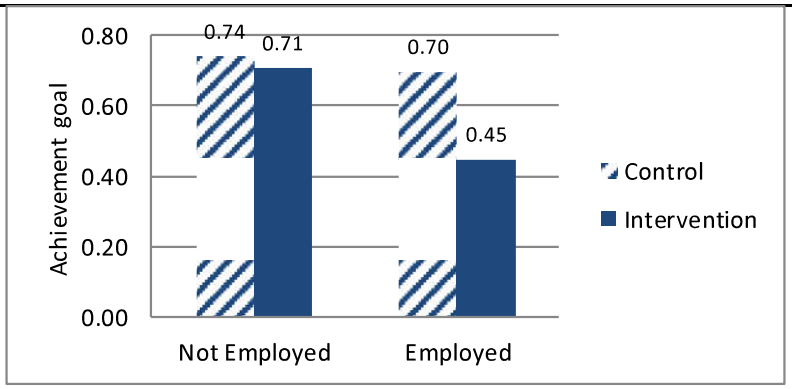


In addition, accounting for the multilevel structure of the data, and child and caregiver characteristics, **two trend-level program effects were detected** as follows:

Playgroup children in the "Other" Ethnicity group presented higher scores (6.3 in an observed range of 17 to 85) in Hearing and Language skills, a measure of receptive (comprehension) and expressive language, than their Control counterparts of the same ethnicity, particularly in comparison to Caucasian children.



Playgroup caregivers that were employed were less likely (25% in an observed range of 0 to 100%) to endorse Achievement goals, qualities related to children's academic performance or intellectual development, to a lesser degree than their Control counterparts that were also employed, particularly in comparison to unemployed caregivers.



Focusing first on the main experimental impact domains, positive (significant and non-significant) effects were more frequent in the three highest ranked domains - Quality of the Home Environment, Child Cognitive Development and Child Temperament and Behavior – as expected. These impacts were stronger in the following skills:

- the extent of responsiveness and involvement of the caregiver in learning and stimulating the child,
- the child's developing ability to reason through manual and visuospatial problems, use language for comprehension, inhibit poor behavior, and focus on task.

Although there was seldom evidence of program effects varying by subgroup (ethnicity and employment status), we highlight benefits to Playgroup Employed caregivers in Involvement. These results overall on main outcomes seem to indicate a focus of the intervention on the promotion of development and the precursors of learning.

We also note that children in the "Other" Ethnicity scored lower in Effortful Control than their control counterparts, while Caucasian and Roma children saw (albeit small and non-significance) increases to their skills. However, Playgroup children in the "Other" Ethnicity group seemed to gather more benefits receptive (comprehension) and expressive language, than their Control counterparts of the same ethnicity, which is not surprising given that the diversity of the category may reflect some diversity in home language.

This focus on development and learning was shared with an intentionality to promote socialization between the children, which resulted in a positive and significant increase in caregivers' endorsement of children's ability to maintain positive, harmonious relationships with others. This focus was well-aligned with caregivers´ motivations to enroll in playgroups (as described below in the study evaluation section)

It is also important to note on the secondary domains a strong significant negative result on Psychological distress; in specific, Playgroup caregivers reported higher frequencies of unpleasant feelings or emotions that impact their overall level of functioning. This effect was particularly pervasive for Roma Playgroup caregivers, and should be examined in further detail in the future.

We also note that positive effects were more frequent in the direct observation measures, applied by research assistants after careful training, rather than in the caregiver self-report measures. Larger (non-significant) negative effect sizes were found in caregiver self-report measures, indicating perhaps a compensatory effect of the control Families (i.e. John Henry effect). In the future, randomization trials should focus on using a smaller number of instruments to measure impact that rely mostly in standardized measures of impact, or if interview-based, should account for social desirability. Chapter 2 includes a detailed discussion of these findings and a set of recommendations for program design and evolution.

Methods– Monitoring

Multi-informant, mixed-methods approach.

Questionnaires, individual interviews and focus groups.

Development and application of new measure of Playgroup Quality – the PERS or Playgroups Environment Rating Scale.

Chapter 3 refers to the implementation study. The chapter contemplates a detailed description of the implementation study, including procedures, measures developed and the characterization of the randomized implementation study subsample.

From the 25 playgroups operating in December 2015 in five districts, we randomly selected 13 playgroups (corresponding to 103 caregivers and children), stratified by district, for the purpose of the study of implementation[4]. These

[4] At time 2, only 12 out of the 13 GABCs randomly selected at time 1 were monitored because one of the GABCs ended before the second wave of data collection.

playgroups were studied during the implementation time period. Comparisons between the implementation study subsample and the full intervention sample demonstrated no statistically significant differences. These comparisons suggest that the results documented for the implementation study subsample of families are representative of the experiences of the full intervention sample.

Study Sample – Implementation Study

13 randomly selected playgroups

103 caregivers and children

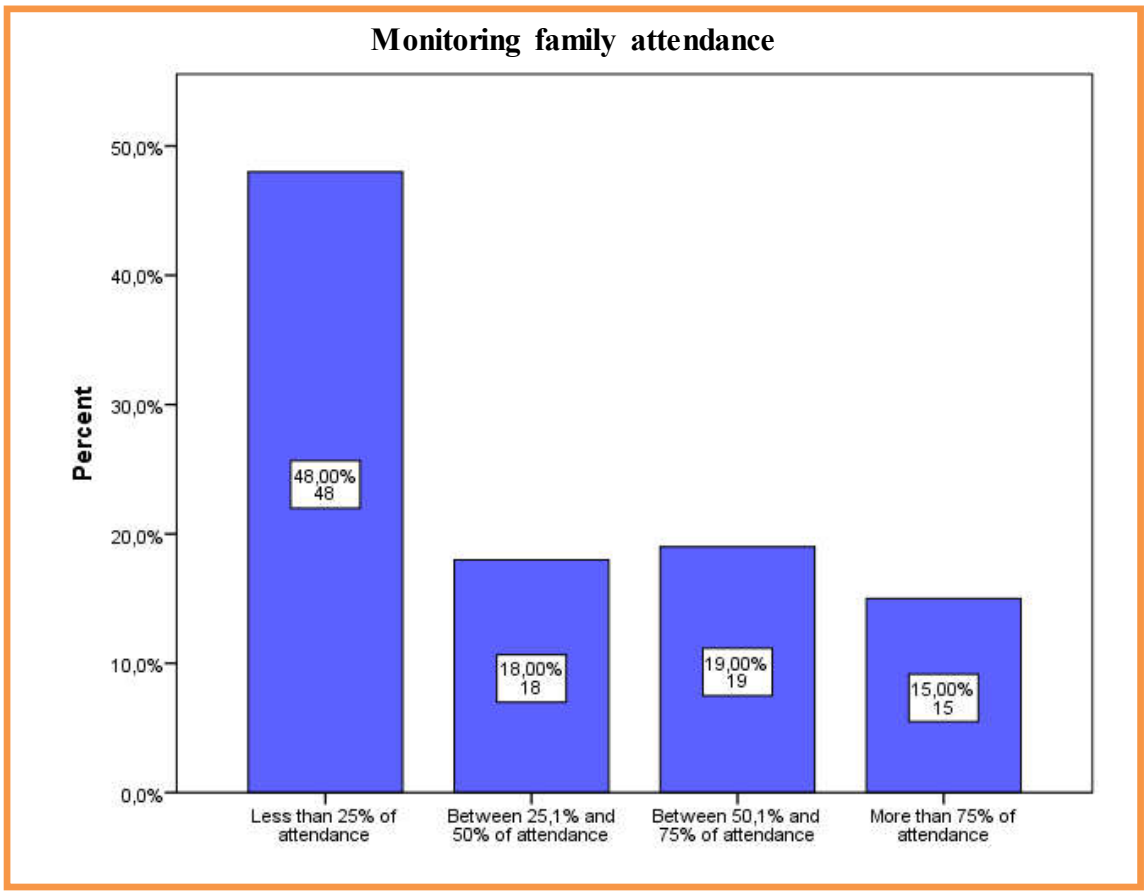
5 supervisors and 14 facilitators

Two waves of data: time 1 (T1) – December 2015/January 2016 and February – and time 2 (T2) –

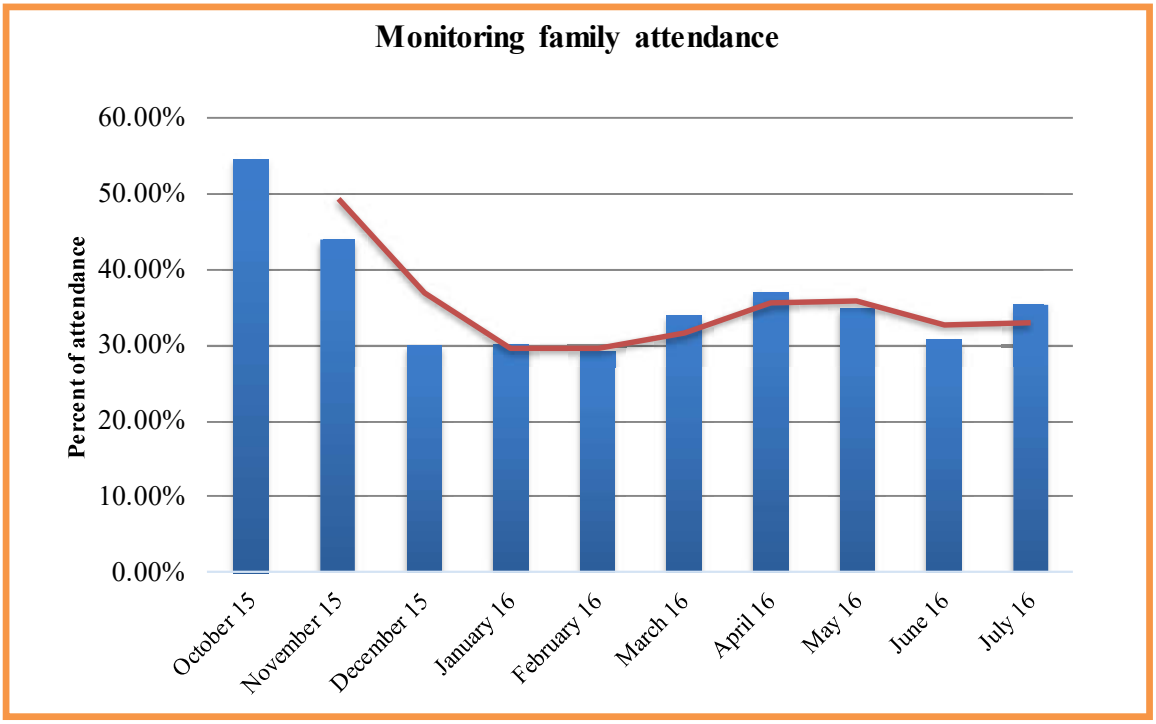
Implementation study data were collected in two waves: time 1 (T1) – December 2015/January 2016 and February (only for the phone interviews) – and time 2 (T2) – May and June 2016[5]. In order to capture a more comprehensive analysis of the implementation process we used a multi-informant, mixed-methods approach. We developed questionnaires that were distributed by email to playgroup staff, i.e., supervisors (N = 5, monthly questionnaire) and facilitators (N = 14, two rounds). We also conducted two individual interviews with the supervisors (T1 and T2), in-person focus groups with caregivers (n= 31 at T1; n=28 at T1) and individual in-depth phone interviews (n=49 at T1, n=29 at T2). We measured playgroup quality twice: at baseline (one month after the beginning of the implementation), and one month before the end of implementation through direct observation of playgroups (12 playgroup at T1 and T2).

Main results of the implementation study indicate that caregivers' attendance was quite low across the 10 months of implementation (38%) and that almost half of the subsample only attended 25% of the sessions. This average attendance rate is, however, similar to attendance rates at comparable programs that target high-risk groups (Baker, Arnold, & Meagher, 2011; Nicholson, Berthelsen, & Vogel, 2008).

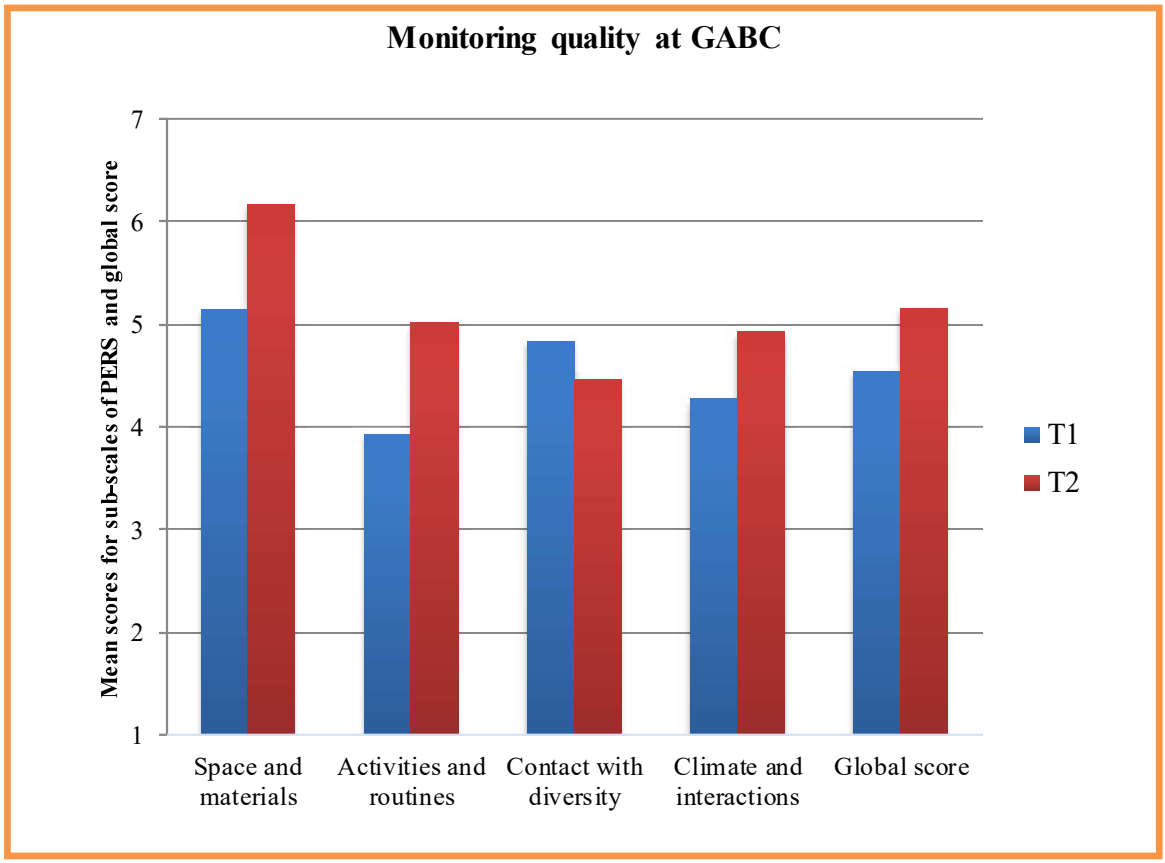
[5] The same measures and procedures were used at Time 1 and Time 2.



Moreover, the average attendance in the first two months (October and November, 2015) of playgroup implementation was near to 50%, a level of attendance similar to other playgroups with two to four months of implementation (Berthelsen et al., 2012).

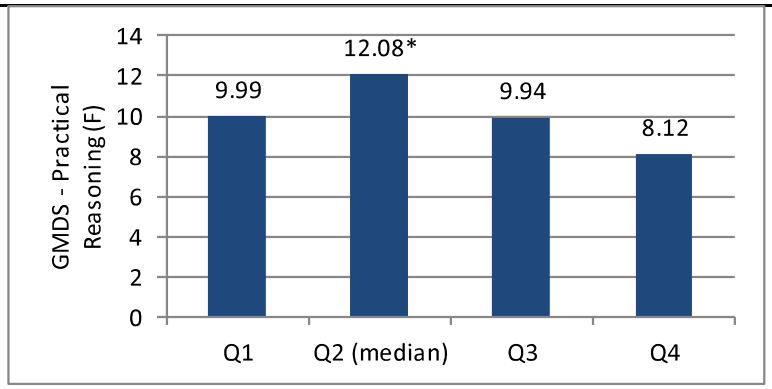


Quality was measured with a new measure – PERS or Playgroups Environment Rating Scale - specifically developed for this study which focusses on the assessment of the various dyadic interactions taking place in playgroups (e.g., facilitators-caregivers; facilitators-children; caregivers-children; children-children), as well as assessment of climate, space and materials, activities and routines, and contact with diversity. PERS assessments indicated that the level of playgroup quality was good in all dimensions at the beginning of the project (average quality was 4.54 with a standard deviation of 0.64at T1, on a scale from 1 to 7, with a medium point at 3.5) and significantly increased during the implementation period (average quality was 5.14 with a standard deviation of 0.61 at T2). Increases were statistically significant in Space and Materials (e.g. playgroup sessions improved opportunities for contact with everyday materials, disposable materials and nature), Activities and Routines (e.g. increased number of contacts with nature, tours at the local community and outdoor play), and Climate and Interactions (e.g. participants enjoyed enhanced complicity, and there was an observable well-established relationship between caregivers, and between facilitators and caregivers).

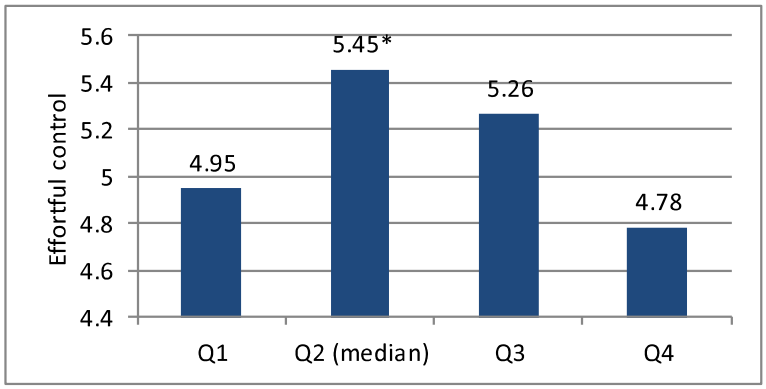


Additional analysis examining non-experimental impacts for different levels of program dosage generally demonstrated benefits in child and caregiver outcomes for those who attended 25.1% to 50% of the sessions, when compared to those who only attended 25% or fewer of the sessions. Accounting for the multilevel structure of the data, pretest scores, child and caregiver characteristics, we highlight the following **significant dosage effects**:

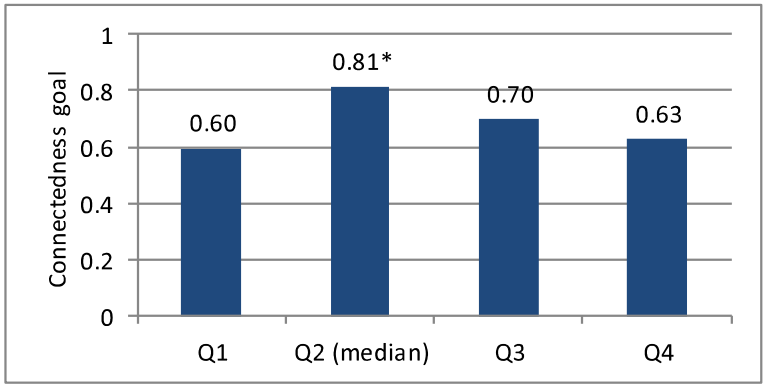
Playgroup children who attended 25.1% to 50% of the sessions (Quartile 2) were 2.087 points significantly (at trend level) higher in Practical Reasoning, than children who only attended 25% or fewer of the sessions (Quartile 1).



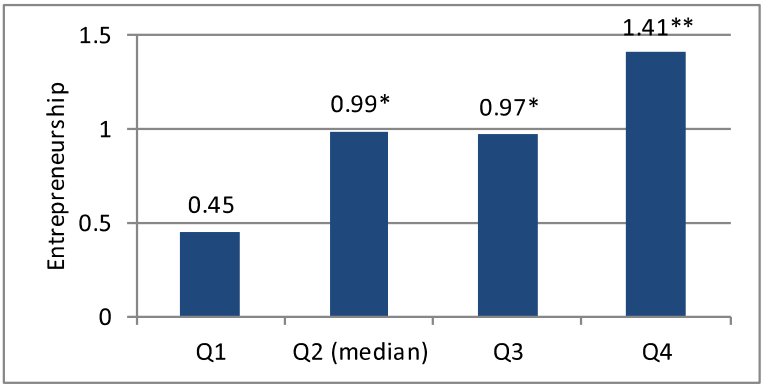
Playgroup caregivers who attended 25.1% to 50% of the sessions (Quartile 2) were 0.507 points significantly higher in Effortful Control than caregivers who only attended 25% or fewer of the sessions (Quartile 1).



Playgroup caregivers who attended 25.1% to 50% of the sessions (Quartile 2) were 0.232 points significantly more likely to endorse Connectedness Goals than caregivers who only attended 25% or fewer of the sessions (Quartile 1).

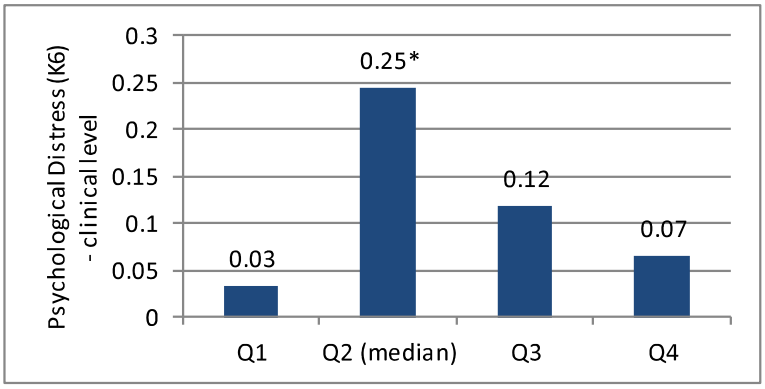


Playgroup caregivers who attended 25.1% to 50% of the sessions (Quartile 2) were 0.537 points significantly more likely to report involvement in Entrepreneurship activities than caregivers who only attended 25% or fewer of the sessions (Quartile 1). Dosage effects were accentuated for higher levels of attendance.

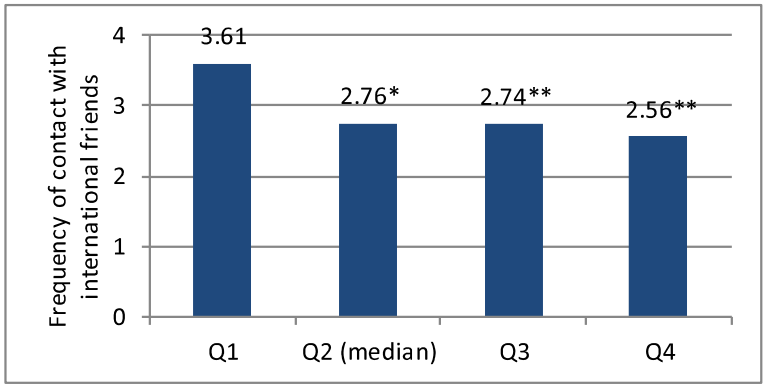


Also of note are some significant effects of dosage in the opposite direction than expected, as follows:

Playgroup caregivers who attended 25.1% to 50% of the sessions (Quartile 2) were 21.1% significantly more likely to score on the clinical level of Psychological Distress, than caregivers who only attended 25% or fewer of the sessions (Quartile 1).



Playgroup caregivers who attended 25.1% to 50% of the sessions (Quartile 2) reported a frequency of contact with international friends 0.851 points significantly lower than caregivers who only attended 25% or fewer of the sessions (Quartile 1). Dosage effects were accentuated for higher levels of attendance.



Although dosage results seem to indicate effects in the expected direction – higher attendance, higher dosage, more positive results - it is important to bear in mind that these results are non-experimental and prone to self-selection bias. There is considerable evidence that participation and attendance to programs are correlated with parents' investment, as well as other determinants of development and learning, such as expectations of others (teachers, grandparents, etc).

To inspect sources of self-selection, we examined differences in pretest characteristics and demographics between families, caregivers and children that attended fewer than 25% of the available sessions and caregivers that attended more than 25% of the available sessions. We found very little evidence of differences, in general low attendees were younger and came from households with lower income than higher attendees, but these differences were not significant. We only found one significant difference between caregivers that attended fewer than 25% of the available sessions and caregivers who attended to more than 75% (the first group was less likely to have completed secondary education). This signifies that sources was self-selection were mostly unobserved, or potentially not so strong for the distinction between these two groups.

Initial reflections from the discussion of implementation study data include the following recommendations:

• Playgroups should have **a shorter period of implementation** (two to four months, for example) that guarantees averages levels of attendance for the participating families;
• It is important to **provide incentives for attendance** to families who might not otherwise engage on an ongoing basis with the playgroup sessions (for example, Roma families, younger caregivers and families with lower household incomes), as disadvantaged families are more likely to have more benefits from attending a playgroup when the dosage is appropriate;
• It is important to **take into account the interests and needs of the participating families**, as the degree of structure in sessions preferred by caregivers varied greatly across the participants
• In order for playgroups to achieve the observed level of quality, , **playgroups must have trained facilitators**, and the **quality of the environment must be monitored** with a valid observation tool, such as the PERS.