CENTRE FOR EFFECTIVE EDUCATION

Queen's University Belfast

Centre for Effective Education
School of Education
Queen's University Belfast
www.qub.ac.uk/cee

An Evaluation of the Effects of the Eager and Able to Learn Programme on Outcomes for 2-3 Year Olds in Early Years Settings

Carol McGuinness, Angela Eakin and Paul Connolly

November 2012

early years
the organisation for young children
the coral initiative

A CORAL Initiative: Aiming to better understand how programmes delivered by Early Years are improving long-term outcomes for children, families and communities.

An Evaluation of the Effects of the Eager and Able to Learn Programme on Outcomes for 2-3 Years Olds in Early Years Settings

Carol McGuinness[1], Angela Eakin[2] and Paul Connolly[2]

[1]School of Psychology, Queen's University Belfast
[2]Centre for Effective Education, Queen's University Belfast

November 2012

Additional Contributors

The following members of the Research Team also made significant contributions to the research upon which this report is based: Celine McStravick, Frances Molyneaux, Cheryl Gamble, Joanne McDowell, Teresa Geraghty and Ruth Sinclair from the National Children's Bureau; Sarah Miller from Queen's University Belfast; and Barbara McConnell and Glenda Walsh from Stranmillis University College.

How to Cite this Report

Any citation of this report should use the following reference:

McGuinness, C., Eakin, A. and Connolly, P. (2012) *An Evaluation of the Eager and Able to Learn Programme for 2-3 year olds in early years settings.* Belfast: Centre for Effective Education, Queen's University Belfast.

Available for download from http://www.qub.ac.uk/cee/

School of Education, Queen's University Belfast, 69-71 University Street, Belfast BT7 1HL, Northern Ireland
Ph: +44 (0)28 9097 5051 | Fax: +44 (0)28 9097 5066 | Email: cee@qub.ac.uk | Web: www.qub.ac.uk

Contents

Executive Summary

Eager and Able to Learn (EAL) is a new pilot programme designed by *Early Years-the organisation for young children* in Northern Ireland, and targeted at 2-3 year-old children in early years settings. It aims to improve young children's eagerness and ability to learn through enhancing their physical, social, emotional, and linguistic development. The programme places a particular emphasis on physical movement, on the physical design of early childhood programme settings, and on relationships - the practitioner/child relationship, the parent/child relationship and the partnership between the parent and the practitioner - to support young children's development. The theory of change underpinning the programme is that movement provides a natural context for children of this age to develop. The programme has a group-based element, which involves a series of developmental movement and play activities, and a home-based element including home visits, which encourages parents to explore play activities with their children in the home environment.

A Senior Early Years Specialist (SEYS) was assigned to each setting to provide: (1) initial training in programme implementation for practitioners; (2) a series of support visits and cluster sessions for practitioners throughout the year; and (3) workshops for parents of children who participated in the programme. In addition, practitioners were given a service design manual to guide them through the delivery of all aspects of the programme. A home learning package for parents was provided.

The Centre for Effective Education at Queen's University Belfast, in collaboration with the National Children's Bureau (NCB) Northern Ireland and Stranmillis University College, were commissioned by Early Years the Organisation for Young Children to undertake a rigorous and independent evaluation of the Eager and Able to Learn programme. The evaluation took the form of a cluster trial using a partial-cross-over design, led by the Centre for Effective Education with the School of Psychology at Queen's, and a fidelity implementation study, led by NCB.

The evaluation studies were preceded by a baseline study in 90 settings. A pilot evaluation in a small number of settings and in-depth qualitative case studies (led by NCB NI and Stranmillis University College) informed the subsequent design of the programme and the design of the final evaluation

and fidelity implementation studies.[1] The findings from the baseline surveys and both elements of the evaluation are presented in three detailed reports that are freely available to download.

This report presents the findings of the partial cross-over design evaluation.

Methodology

The evaluation consisted of a cluster trial using a partial cross-over design involving 28 early years setting and 454 children aged 2-3 years. 18 Day Care and 10 Sure Start settings participated. The trial was conducted over two years from September 2008 to June 2010. In Year One, from September 2008 to June 2009, the settings continued with their usual programme of activities and the cohort of 2-3 year olds attending the settings during that year acted as a control group. In Year Two, from September 2009 to June 2010, the same settings introduced the EAL programme and the next cohort of 2-3 year olds who attended the settings acted as the intervention group. The study therefore used a partial cross-over design, with each setting acting as its own control. In addition, 180 practitioners and 390 parents participated.

All data collection was undertaken by a team of highly trained fieldworkers and co-ordinated by the research team. The data were analysed using multi-level modeling in order to account for the clustering of children, parents and practitioners in settings, controlling for pre-test scores on the relevant variables.

Outcomes

For the purposes of this present evaluation, an outcome is defined as a real and discernible change in attitudes and/or awareness that has occurred as a direct result of taking part in the Eager and Able to Learn programme. The following outcomes were identified by Early Years the Organisation for Young Children for the purposes of the present evaluation:

[1] Molyneaux, F., Walsh, G., McConnell, B. and McGuinness, C. (2012) The *Eager and Able to Learn Pilot Year: Lessons learned – implementation and evaluation of the pilot programme*.

Child Outcomes

- Improved language and communication skills: increased vocabulary and increased ability to use vocabulary in context.
- Improved social/emotional skills and behaviours: increased independence and self-help skills.
- Improved ability to think and solve problems.
- Improved levels of involvement: increased levels of concentration, persistence and precision
- Improved levels of physical movement: improved gross, fine and sensory motor development.

Setting outcomes

- Improved quality of the learning environment and experiences in the early years settings for two to three year olds.

Practitioner and Parental Outcomes

- Increased recognition of the importance, and the different purposes, of play, in the development of two-year-old children; and increased frequency in providing different types of play opportunities, both indoors and outdoors.
- Increased responsiveness in practitioners' interactions and engagement with two-year-old children in order to support their communication, social, emotional, physical and cognitive development needs.
- Increased recognition of the importance of movement for two-year-old development and how it can be related to wider developmental goals (e.g. language, cognitive, social-emotional, as well as motor development).
- Increased recognition of the importance of working in partnership with parents around the developmental needs of two-year-old children, increased opportunities to communicate with parents, and increased satisfaction with the communication.

Measures

The children's outcomes were assessed using the Bayley Scales for Infant and Toddler Development, 3rd Edition (2006), commonly known as Bayley III. Five domains of children's development were assessed by trained fieldworkers observing the children as they completed play-based tasks in cognitive development, receptive and expressive communication, fine and gross motor movement. Two other domains were assessed by the practitioners who rated the children's social-emotional

development and adaptive behavior, including communication, functional academics (emergent literacy), self-direction, play and leisure, and social interaction. All scales have high reliability and validity.

The Early Childhood Environmental Rating Scale-Revised Edition (ECERS-R, Harms, Clifford & Cryer, 2005) was the observation instrument used to assess the quality in the group-based settings. ECERS-R has seven sub-scales, each dedicated to a different aspect of early childhood practice. The scale is recommended for use with children aged between 2 ½ to 5 years and it was supplemented with subscale indicators from its sister scale, ITERS-R (The Infant-Toddler Environmental Rating Scale, Revised Edition, Cryer, Harms, & Riley, 2003) designed for younger children aged between 1 month and 30 months.

Survey questionnaires were specially designed for early years practitioners and parents to elicit information about their knowledge, attitude and behaviors related to the developmental needs of 2-3 year olds.

Findings

Child Outcomes

The EAL pilot programme had statistically significant main effects on 3/11 developmental domains as assessed by Bayley lll and produced a surprising pattern of findings. The children's social and emotional development was positively affected by the programme (effect size=+.30), while their cognitive development was negatively affected (effect size= -.29). There were also some smaller positive effect sizes on other developmental domains that were consistent with the positive social emotional effect but which did not reach statistical significance - communication (effect size=+.17), social skills(effect size=+17) and self-direction (effect size=+.13). The strongest negative effect was on emergent literacy activities, called functional pre-academics (effect size=-.29). However, exploratory analysis revealed that this polarising pattern of the findings was statistically related to the children's pre-test scores; the observed effects, both positive and negative, were seen more strongly in children with high pre-test scores than those with lower pre-test scores.

The effects were consistent across the whole sample of children and no differences were found between: boys and girls; settings with different management types; or settings located in rural/urban areas. Overall, the programme was delivered with high fidelity and whatever variation existed had no discernible effects on the majority of the outcomes, with the exception of receptive and expressive language where it had a significant positive effect. The number of hours the children spent in a setting had a significant negative effect on their self-direction scores. Surprisingly, the quality of the settings showed no statistically significant positive effects and one significant effect was in a negative direction.

Setting Outcomes

Participating in EAL improved the *average* quality for settings. Not all settings improved; 18 improved, 7 got poorer, and 3 remained the same.

The most positive and statistically significant effects of EAL were on the subscales related to *interactions* between children and staff, *interactions* between the children, *interactions* between parents and staff and between the staff themselves. Settings were already scoring relatively high on these aspects of practice, receiving ratings between 4 and 5. Nevertheless, EAL provided an additional boost, resulting in the average ratings moving beyond 5, and 20% of settings moved into the excellent range. The EAL programme consisted of increased contact with parents through workshops and home visits, and these clearly contributed to the improved quality ratings with regard to parents and staff.

Parental and Practitioner Outcomes

Participating in the EAL pilot programme had significantly positive effects on practitioners' and parents' beliefs, attitudes and self-reported behaviours. Although significant effects emerged for only a limited number of survey questionnaire items, the effect sizes were often large, ranging from .68 to .19, and almost entirely in the direction expected by the aims and goals of the programme. For example, with regard to providing new and different opportunities and materials for play, EAL practitioners reported that they were using more 'everyday' materials' (e.g., pots, pans, crumpled papers) and props to help with movement games (e.g., scarves, balls, hoops). They also reported less frequent use of 'books and story-telling' and 'number games'. With regard to their interaction with children, the EAL practitioners were much less likely to adopt 'harsh and controlling' interaction styles and more likely to explain the reasons for things in order to encourage the children to think for themselves. Consistent with the new partnership arrangements with parents during the EAL year, practitioners were more positive about how their setting worked with parents, and less doubtful than they were during the control year about the contribution that parents can make to support their children's learning in the setting.

EAL parents also reported playing with their children in different types of ways – with song and dance, and using different materials. EAL parents showed a sharper recognition than Control parents about the relevance of play to different forms of learning both in the present and in the future. EAL parents were also substantially more satisfied than Control parents with communicating and sharing views with staff in the early years settings and also reported more help with materials and training for promoting their child's development. However, in terms of overall self-efficacy and responsiveness as measured by the Parental Self-Efficacy Scale, there was very little difference between the two groups of parents, and Control parents scored marginally higher then EAL parents on expressing emotion and affection to their children, though the scores were high for both groups.

Specific Recommendations for the EAL programme

The impact of the EAL programme on the children's development produced a surprising and unexpected pattern of results, with positive effects on the social emotional development and negative effects on cognitive and emergent literacy outcomes. This 'polarising' effect was more noticeable in those subgroups of children who were more developmentally advanced when they joined the programme. The absence of an effect on gross motor development was also surprising given the emphasis on the movement activities and the high fidelity implementation in this component of the programme.

The positive social emotional impact (from Bayleys) is consistent with the observed positive boost on the rated quality of the EAL settings compared to the control settings, especially on interactions between staff/child and child/child (ECERs-R), and on the practitioners' reports post-EAL that they were interacting with children in a more positive way (practitioners' survey questionnaires). Parents also appeared to learn more about the role of play in children's development and to experiment with different types of play. They were also more positive about their own interactions with the early years settings (parents' questionnaires). The findings from the Fidelity Study (Geraghty et al., 2012) show that the large majority of settings implemented the programme with high fidelity and that the programme was warmly welcomed by the vast majority of practitioners, setting managers and parents.

However, in the light of the mixed findings on child outcomes, the content of the programme needs to be re-evaluated to ensure that the positive child outcomes can be maintained and the negative impact minimised.

From the EAL evaluation, with regard to the development of the programme

- Given the unusual pattern of findings for children's outcomes, Early Years should re-evaluate the content of the EAL programme to ensure that the positive impacts on children, quality of settings, practitioners and parents are maintained and the negative impacts are minimised or turned around.
- Specifically, the dominance of the movement activities in terms of time allotted should be reassessed to create a more balanced programme that focuses directly on socio-emotional development, language, movement and conceptual development.
- The focus on high quality interactions between adults and children should be maintained and enhanced in any future programmes.

- The focus on partnerships between settings and parents should be maintained and enhanced, following the advice from the Fidelity Implementation Study on involving parents and on managing home visits.
- Fidelity monitoring should be part of any future roll-out of the programme.

General Recommendations for Policy and Research

A focus on provision for 2-3 year olds has emerged only recently as a national priority, with the launching of the 2-year old Sure Start programmes in England, Wales and Northern Ireland. Previously, both policy and research had focussed on 3-4 years in the pre-school year (e.g., the EPPE and EPPNI longitudinal research studies and the expansion of free pre-school places). The research base on what we know about the impact of provision for two years in the UK is at a very early stage. For example, the National Evaluation of Neighbourhood Nurseries (Neighbourhood Nurseries Initiative (NNI) Research Team, 2007) in England and the evaluation of the Early Education Pilot for Two Year Old Children (Smith et al., 2009) in England, both focussed on disadvantaged children. The current studies contribute substantially to the research base in Northern Ireland. From a research/policy perspective, it is important that, as well as evaluating the impact of specific programmes, participating in early years provisions (of whatever kind) is included as part of current and any future longitudinal cohort tracking (e.g., the Northern Ireland Millennium Cohort and any future cohort studies in Northern Ireland).

A consistent finding across many pre-school studies is the **importance of the quality** of the settings for early years outcomes. This point has been confirmed again in the pilot evaluation for two-year olds in England, where positive outcomes for children were reported only for those who attended the very highest quality settings (Smith et al., 2009, chapter 7). The average rated quality of the early years settings in the current study deserves immediate attention.

The EAL trial is one evaluation of an innovative pilot programme that focussed on developmental movement experiences as a potential approach for accelerating more general development. Although the findings from the EAL evaluation on child outcomes are surprising, it is important that research continues on the relationship **between different kinds of movement development as a potential approach for early years intervention.**

It is important to appreciate the scale and scope of these early years studies for Northern Ireland and to understand the logistical demands of running research studies on this scale with 2-3 year old children. They need careful consideration so that **policy decisions are research informed** and are appropriately **benchmarked with international developments**.

Acknowledgements

This research was commissioned by Early Years the organisation for Young Children as part of it's CORAL Project with an aim of better understanding how programmes delivered by Early Years are improving long term outcomes for children, families and communities. Early Years have been supported by the Atlantic Philanthropies in both developing evidence based programmes and in commissioning this evaluation. Early Years has also been supported by an International Advisory group in developing and implementing the programmatic content and in overseeing the development of the research programme.

The research team is indebted to the managers and practitioners in the early years settings, the children and parents who participated in the study. We also appreciate the cooperation and ongoing support from Early Years during the evaluation.

1. Introduction

Eager and Able to Learn (EAL) is a new pilot programme designed by *Early Years-the organisation for young children* in Northern Ireland, and targeted at 2-3 year-old children in early years settings. It aims to improve young children's eagerness and ability to learn through enhancing their physical, social, emotional, and linguistic development. The programme places a particular emphasis on physical movement, on the physical design of early childhood programme settings, and on relationships - the practitioner/child relationship, the parent/child relationship and the partnership between the parent and the practitioner - to support young children's development. The theory of change underpinning the programme is that movement provides a natural context for children of this age to develop. The programme has a group-based element, which involves a series of developmental movement and play activities, and a home-based element including home visits, which encourages parents to explore play activities with their children in the home environment.

A Senior Early Years Specialist (SEYS) was assigned to each setting to provide: (1) initial training in programme implementation for practitioners; (2) a series of support visits and cluster sessions for practitioners throughout the year; and (3) workshops for parents of children who participated in the programme. In addition, practitioners were given a service design manual to guide them through the delivery of all aspects of the programme. A home learning package for parents was provided.

The evaluation of the programme took the form of a cluster trial using a partial-cross-over design, led by the Centre for Effective Education with the School of Psychology, and a fidelity implementation study, led by the National Children's Bureau in Northern Ireland. The findings from the fidelity implementation study are reported separately (Geraghty, Molyneaux & Dunne, 2012).

This report presents the findings of the cluster trial evaluation into the effectiveness of the EAL programme in improving outcomes for children, their parents and the early years practitioners/settings involved in delivering EAL. The trial involved 454 children aged 2-3 years, in 28 settings located across Northern Ireland. The trial was conducted over two years from September 2008 to June 2010. In Year One, from September 2008 to June 2009, the settings continued with their usual programme of activities and the cohort of 2-3 year olds attending the settings during that year acted as a control group. In Year Two, from September 2009 to June 2010, the same settings introduced the EAL programme and the next cohort of 2-3 year olds who attended the settings acted as the intervention group. The study therefore used a partial cross-over design, with each setting acting as its own control.

The trial evaluation was preceded by a baseline study conducted between September 2008 and January 2009 (McGuinness, Connolly, Eakin & Miller, 2012), a pilot evaluation of the EAL intervention in 7 settings from Sept 2008-June 2009 and in-depth qualitative case studies in 4 settings during the pilot year (Molyneaux, Walsh, McConnell & McGuinness, 2012).

2. Methodology

This section outlines the methodology used for the present evaluation. It begins by describing the specific outcomes, as agreed with Early Years, that were tested and that provide the focus for the present evaluation. These outcomes reflect the core aims and objectives of the programme. The section concludes with an explanation of the approach used to analyse the data.

2.1 Outcomes

For the purposes of this present evaluation, an outcome is defined as a real and discernible change in the developmental status of the children, and in the attitudes and self-reported behaviours of the parents and practitioners that has occurred as a direct result of taking part in the EAL programme.

2.1.1 Child Outcomes

- Improved language and communication skills: increased vocabulary and increased ability to use vocabulary in context.
- Improved social/emotional skills and behaviours: increased independence and self-help skills.
- Improved ability to think and solve problems.
- Improved levels of involvement: increased levels of concentration, persistence and precision
- Improved levels of physical movement: improved gross, fine and sensory motor development.

2.1.2 Setting outcomes

- Improved quality of the learning environment and experiences in the early years settings for two to three year olds.

2.1.3 Practitioner Outcomes

- Increased recognition of the importance, and the different purposes, of play, in the development of two-year-old children; and increased frequency in providing different types of play opportunities, both indoors and outdoors.
- Increased responsiveness in practitioners' interactions and engagement with two-year-old children in order to support their communication, social, emotional, physical and cognitive development needs.
- Increased recognition of the importance of movement for two-year-old development and how it can be related to wider developmental goals (e.g. language, cognitive, social-emotional, as well as motor development).
- Increased recognition of the importance of working in partnership with parents around the developmental needs of two-year-old children, increased opportunities to communicate with parents, and increased satisfaction with the communication.

2.1.4 Parental Outcomes

- Increased recognition of the importance of play in the development of their two-year-old children; and increased frequency in providing different types of play opportunities, both indoors and outdoors.
- Increased responsiveness in parents' interactions and engagement with two-year-old children in order to support their communication, social, emotional, physical and cognitive development needs.
- Increased recognition of the importance of movement for two-year-old development and how it can be related to wider developmental goals (e.g. language, cognitive, social-emotional, as well as motor development).
- Increased recognition of the importance of working in partnership with practitioners around the developmental needs of their two-year-old children, increased opportunities to communicate with practitioners, and increased satisfaction with the communication.

2.2 Design

The evaluation consisted of a cluster trial using a partial cross-over design involving 28 early years settings for 2-3 year old and was conducted over two years. The 28 settings who participated in the EAL were a sub-sample of the original 90 settings who had participated in the baseline study. The 90 settings were then randomly allocated to a control condition (N=45) and an intervention condition (N=45) as part of a cluster randomized control trial which was planned for 2008-2009. For various reasons, the randomized control trial did not proceed as planned. Instead, it was decided to run an

evaluation trial over 2 years, where the children in those settings who had been randomly allocated to the control condition would continue to constitute the 'control condition' (2008-2009) and the children in the *same settings* would receive the EAL intervention the following year (2009-2010). The study therefore used a *partial* cross-over design, with each setting acting as its own control. During the control year in 2008-2009, 44 settings (one of the original 45 settings had closed) were invited to participate in the evaluation for the following year. 38 settings agreed – 28 Day Care and 10 Sure Start settings. Subsequently, and prior to the commencement of EAL training in September 2009, a further 8 Day Care settings withdraw, due to predicted difficulties with releasing staff for training. Another Day Care setting withdrew at the pre-test intervention point (October 2009) and in a further Day Care setting, no children in the intervention year were in the appropriate age range. Thus, 28 settings participated fully in both the control and intervention years – 18 Day Care and 10 Sure Start settings. Compared to the original sample of 44 settings, where the ratio of Day Care to Sure Start settings was 3:1, the final ratio was approximately 2:1.

During the control year the settings continued with their normal curriculum and practices. Less than half the settings claimed that they were following any structured or prescribed programme during this period. Structured programmes that were mentioned included High Scope (1), Reggio (1), 2-year old programme (3), This is Me (2), Early Years Curriculum (1), Birth-3 Matters (1). In addition, a small number of settings were seeking High Scope accreditation (2), or approval as an All Ireland Centre of Excellence (3), or were having an ETI inspection during the year (1).

During the intervention year, practitioners in the settings were trained by Early Years to deliver the EAL programme which was then implemented from October 2009 to June 2010. During the intervention year, the settings reported that they did not give up or stop doing any other activity in order to implement the EAL programme.

The key implication of this in relation to the interpretation of the findings set out in this report regarding the effectiveness of the EAL is that any effects found are those achieved *above and beyond* the curriculum approaches that were being implemented in the settings during the control year.

Another important feature of the design was that children participated in the programme for different lengths of time each week for two reasons – because of the type of setting they were attending, Day Care or Sure Start, and whether they attended a setting full-time or part-time. The Day Care Nurseries were privately owned nurseries where children attended on a full-day or part-day basis. The Sure Start programmes were part of a new government-funded programme for 2-year olds in socially deprived areas in Northern Ireland and children attended for 2-3 hours per day. Individual children's attendance was tracked during the EAL intervention year. The average hours per week attended was 18.7 (min 5 and max 45). Sure Start children attended for 6-7 hours per week (min 5, max 8) while Day Care children attended for 23-24 hours per week (min 6, max 45).

2.3 Sample

2.3.1 Settings

As stated above, a total of 28 Early Years settings took part in the evaluation, comprising 18 Day Care Nurseries and 10 Sure Start programmes. Thirteen of the 28 settings were located in a rural area and 15 were located in an urban area. Table 1 presents the distribution of urban/rural settings by Day Care/Sure Start.

Table 1. **Number of settings per group for Day Care/Sure Start and Rural/Urban divides**

Setting Type	Location		Total
	Rural	Urban	
Day Care	8 (61.5%)	10 (66.7%)	18 (64.3%)
Sure Start	5 (38.5%)	5 (33.3%)	10 (35.7%)
Total	13 (100%)	15 (100%)	28 (100%)

2.3.2 Children

In total, 454 children participated in the evaluation: N=197 (2008/2009); N=257 (2009/2010). Table 2 presents the distribution of children in the control and intervention (EAL) groups by Day Care/Sure Start.

Table 2. **Breakdown of the sample by Sure Start and Day Care settings**

	Control Group		Intervention Group	
	n	%	n	%
Sure Start/Day Care				
Sure Start	53	26.9%	80	31.1%
Day Care	144	73.1%	177	68.9%
Total	**197**	**100%**	**257**	**100%**

Of the 197 children who participated during 2008/2009 (control cohort), 194 were assessed at pre-test, 176 were assessed at post-test and 172 children participated at both pre- and post-testing stages. Of the 257 children who participated during 2009/2010 (intervention/EAL cohort), 254 were assessed at pre-test, 234 were assessed at post-test and 231 children participated at both pre- and post-testing stages.

All target children in the evaluation were required to be between 2 years and 2 years 9 months on the 1st October 2008 (control), 2009 (intervention/EAL)[2]. At the point of pretesting, the control children ranged in age from 2 years to 3 years 1 month and the intervention children ranged in age from 2 years to 2 years 10 months. The differences in the age ranges between the control and EAL samples at point of testing is due to the more extended data collection period for the control sample

[2] For reasons related to the assessment tool being used to measure the child outcomes (the Bayley Scales), the age range of the children taking part in the evaluation was restricted to children aged between 2 years 0 months and 2 years 9 months at the time of entry into the programme (October). It was anticipated that there might be ceiling effects in the Bayley at post-test if older children were in the sample.

(Oct-Feb) compared to the intervention sample (Oct-Nov). As the Bayley scores are standardised for age, these differences do not impact on the interpretation of the data.

Overall, 241 boys (53.1%) and 213 girls (46.9%) participated. Table 3 presents the breakdown of boys and girls across Control and EAL groups.

Table 3. **Breakdown of the sample by *Gender***

	Control Group		Intervention Group	
	n	%	n	%
Gender				
Boys	107	54.3%	134	52.1%
Girls	90	45.7%	123	47.9%
Total	**197**	**100%**	**257**	**100%**

Measures of social deprivation, the Multiple Deprivation Rank (and Score), derived from the Super Output Area Statistic (NINIS, 2005) were calculated based on children's home postcode (provided through parental questionnaires). Where the child's postcode was not available, the setting postcode was used as a reasonable estimate of the geographical location of the children's homes. The overall mean rank (score) for the control cohort was 442.8 (21.0). The overall mean rank (score) for children participating during the intervention year was very similar: 429.0 (22.2). Because the Sure Start programmes were funded on the basis of area rather than on the basis of individual child need, they can enroll children from a variety of social backgrounds. Table 4 presents the breakdown of rank (score) by Day Care/Sure Start, [3] showing that, *on average*, the children who attended Sure Start settings came from more socially deprived backgrounds than those who attended the Day Care settings in the sample.

Table 4. **Breakdown of the sample by SOA deprivation rank/score**

	Control Group		Intervention Group	
	Rank	Score	Rank	Score
Setting Type				
Sure Start	231.60	33.41	237.34	33.55
Day Care	520.60	16.49	515.65	17.10

[3] 'Neighbourhood Statistics for NI' (NINIS) has recently removed Multiple Deprivation 'Scores' from their databases. NINIS now suggest that it is more meaningful to quote the 'rank score' from their database.

2.3.3 Parents

A total of 390 parents participated in the evaluation. 244 completed both the pre- and post-intervention questionnaires; the remaining 146 completed only one of either the pre- (N=72) or the post-test questionnaires (N=74).

Table 5. **Number of parents who completed pre- and post- intervention questionnaires**

Total no. of parents	Pre and post data	Pre data only	Post data only
390	244	72	74

Table 6 shows the ages of the parents/guardians in the sample. Over 50% were aged between 26-35 years, and the remainder were older rather than younger. Parents of the intervention group children were slightly older than the control group children, with 38.1% aged 36-45 years, compared to 29.7% in the control group.

Table 6. **Parent/Guardian age range**

	Control Group		Intervention Group	
	n	%	n	%
Age Group				
18-25 years	18	10.5%	21	9.6%
26-35 years	101	58.7%	112	51.4%
36-45 years	51	29.7%	83	38.1%
46-54 years	1	.6%	2	.9%
Over 55 years	0	0%	0	0%
Missing	1	.6%	0	0%
Total	172	100%	218	100%

Table 7 shows the parents' educational qualifications by type with very little difference between the control and intervention group. Over 40% of the parents had qualifications at degree level or above.

When the parents' qualifications were translated into the Census classifications Levels 1-5, it can be seen that parents of children from Rural areas had higher qualifications than those from Urban areas (M=3.93 vs M=3.41). Table 8 shows that Sure Start parents for the Intervention group held higher qualifications than Sure Start parents for the control group, although their social- economic backgrounds were almost identical, as indexed through their postcodes (see earlier table).

Table 7. **Parents' highest level of educational qualifications: Type**

	Control Group		Intervention Group	
	n	%	n	%
Qualification				
4 GCSEs (D-G) NVQ Level 1 GNVQ Foundation	9	5.2%	13	6.0%
First Cert/Dip 4-5 GCSEs (C or above) NVQ Level 2 GNVQ Intermediate	43	25.0%	45	20.6%
National Cert/Dip 2 A Levels NVQ Level 3 GNVQ Advanced	35	20.3%	49	22.5%
Higher Cert/Dip Foundation Degree NVQ Level 4	14	8.1%	7	3.2%
Honours Degree	39	22.7%	54	24.8%
Post Graduate	27	15.7%	36	16.5%
Missing	5	2.9%	14	6.4%
Total	172	100%	218	100%

Table 8. **Mean Parent Qualifications: Using the Census classification levels**

	Control Group	Intervention Group
Setting Type		
Sure Start	2.54	3.29
Day Care	3.93	3.92

2.3.4 Practitioners

A total of 180 practitioners participated in the evaluation. As detailed in Table 9, 105 completed both the pre and post intervention questionnaires; the remaining 75 completed only either the pre-(N=38) or the post-(N=37) questionnaires.

Table 9. **Number of practitioner questionnaires completed**

Total no. of practitioners	Pre and post data	Pre data only	Post data only
180	105	38	37

As outlined in Table 10, over 50% of the practitioners were under 25 years of age.

Table 10. **Practitioner age range**

	Control Group		Intervention Group	
	n	%	n	%
Age Group				
18-25 years	49	50.5%	47	56.6%
26-35 years	21	21.6%	21	25.3%
36-45 years	19	19.6%	12	14.5%
46-54 years	5	5.2%	3	3.6%
Over 55 years	0	0%	0	0%
Missing	3	3.1%	0	0%
Total	97	100%	83	100%

Table 11 shows the highest level of qualifications held by the practitioners in the sample. Very few practitioners have degree level qualifications and the practitioner qualifications in the control year and the intervention year are very similar (as might be expected given that they are the same settings separated by one year).

Table 11. **Practitioners' highest level of educational qualifications: Type**

	Control Group		Intervention Group	
	n	%	n	%
Qualification				
4 GCSEs (D-G) NVQ Level 1 GNVQ Foundation	1	1.0%	2	2.4%
First Cert/Dip 4-5 GCSEs (C or above) NVQ Level 2 GNVQ Intermediate	30	30.9%	25	30.1%
National Cert/Dip 2 A Levels NVQ Level 3 GNVQ Advanced	53	54.6%	42	50.6%
Higher Cert/Dip Foundation Degree NVQ Level 4	3	3.1%	5	6.0%
Honours Degree	6	6.2%	8	9.6%
Post Graduate	1	1.0%	0	0%
Missing	3	3.1%	1	1.2%
Total	97	100%	83	100%

Table 12 shows practitioner qualifications translated in the Census classification levels 1-5. Those practitioners who work in Sure Start settings have higher qualifications than those who work in Day Care.

Table 12. **Mean Practitioner Qualifications: Using the Census classification levels**

	Control Group	Intervention Group
Setting Type		
Sure Start	3.04	3.19
Day Care	2.78	2.77

2.4 Measures

2.4.1 Settings

Quality

The quality of the settings was measured using the The Early Childhood Environmental Rating Scale (Harms, T., Clifford, R.M., Cryer, D., 2005, Revised Edition). ECERS-R is designed to assess *process* quality in early childhood centre-based settings. It was developed in the US and is now used widely in the UK as a research tool (e.g., Sylva et al., 2004; Melhuish et al., 2004; Roberts et al., 2010), as an audit tool, and for professional development purposes (Mathers et al., 2007). The word 'environment' in the title can be taken in its broadest sense to mean not only the physical characteristics of a setting, but also the quality of social interactions, strategies to promote all-round learning, the relationships between the children as well as between the adults and children. ECERS-R has 43 items divided into 7 sub-scales, each dedicated to a different aspect of early childhood practice. The scales are: Space and Furnishings, Personal Care Routines, Language and Reasoning, Learning Activities, Interactions, Programme Structure, Parents and Staff. Each item is rated on a seven-point scale with explicit indicators for scores of 1 (inadequate), 3(minimal/adequate), 5(good), and 7 (excellent). Subscale scores, as well as total average scores, can be derived (see Appendix 1 for more details on the scale). One of the strong points of the rating scale is the rigorous criteria against which observations are made. ECERS-R has good psychometric properties and good predictive validity.

ECERS-R is normally recommended for use with children aged between 2 ½ to 5 years. As the EAL programme was designed for 2-3 year olds, the research/practitioner team debated about its suitability for these settings. In the long run it was decided to supplement the ECERS-R scale with selected subscales indicators from its sister scale, ITERS-R (The Infant-Toddler Environmental Rating Scale, Revised Edition, Cryer, D., Harms, T. & Riley, C., 2003) designed for younger children. Indicators from three ITERS subscales were included in the observations; from the language scale (1 item), the interactions scale (2 items), and the parents and staff scale (1 item).

Programme Fidelity

The programme fidelity measure was taken from the sister project, the Fidelity Implementation Study. To measure fidelity to the programme design, a series of eight key indicators of fidelity were identified. These included:

- Practitioner attendance levels at training (both initial and cluster training sessions)
- Ratio of EAL trained practitioners to children in the settings
- Number of home visits conducted by practitioners
- Number of EYS support visits conducted
- Proportion of parents attending the workshop
- Number of developmental movement experiences (MEs) completed
- Duration of implementation of the movement experiences (MEs)
- Frequency of implementation of the developmental movement experiences (MEs).

A scoring system was applied whereby settings were awarded a score of 1-5 for each indicator (one being the lowest score awarded and five the highest), the maximum potential score achievable therefore was 40 and the minimum eight. A full description of the fidelity measure can be found in the Fidelity Implementation Study (Geraghty, Molyneaux & Dunne, 2012).

2.4.2 Children

The assessment tool used for the children was the Bayley Scales for Infant and Toddler Development, 3rd Edition (2006a, 2006b), commonly known as Bayley III, developed from the long established Bayley Scales. In this revision, five domains of children's development are separately assessed – cognitive development, receptive and expressive communication, fine and gross motor movement. These domains are assessed through play-based tasks, where the children directly engage with specific test activities and their performance is rigorously assessed according to a strict marking protocol

In addition, there are two new domains based on ratings from a person who observes the children in everyday settings (normally the parent). The Social-Emotional Scale is based on the Greenspan Social-Emotional Growth Chart (Greenspan, 2004) and measures how well children meet certain social-emotional milestones. The Adaptive Behaviour Scale is designed to measure the attainment of functional skills necessary for increased independence. It is based on the Adaptive Behaviour Assessment System – 2nd Edition (ABAS-II; Harrison and Oakland, 2003) and is divided into ten sub-scales. Only 6/10 of these scales were used, due to time considerations, specifically: Communication, Functional Academics, Self-Direction, Leisure and Social Interaction (see Appendix for more details of the scales). All Bayley-III scales have high reliability and validity.

For this evaluation, the practitioners in the early years settings rated the children on the Social-Emotional Scales and the Adaptive Behaviour Scales rather than the parents. This decision was largely for practical reasons (e.g., a desire not to put too great a burden on parents who were also required to complete parental questionnaires). Nevertheless, there were distinct advantages from using the practitioners as raters. They may not have had as extensive knowledge of any single child as a parent has, but they were in a position to draw on their experience of a larger number of children, and thus were well positioned to make comparative judgements about children's

development. However, the norms for the scales are based on parents' ratings of their own children and, thus, cautious interpretations must be made based on the norms.

Using the Bayley age-based norms (from US samples), the average performance for each of the developmental domains has a scaled score of 10 and a standard deviation of 3. Because of the recency of the new Bayley III test, only the US norms were available for comparison at the time of testing in 2008. Subsequently, UK norms for Bayley III for a sample of 221 children aged from 10 months to 2 years 3 mths became available (Bayley, 2010 UK and Ireland Supplement Manual). Comparisons were made with both US and UK children – and slightly different findings emerged. These will be commented on when interpreting the findings of the evaluation. Table 13 provides a summary of the Bayley domains assessed for the trial.

***Table 13*. Child outcomes assessed for the trial**

Form of Assessment	Domain	Description
Domains Assessed through play-based tasks	*Cognitive*	This scale specifically examines the child's ability to solve problems and complete simple puzzles, to match patterns, to group objects, to engage in representational and imaginary play and to understand one-to-one correspondences.
	Receptive Communication	This scale focuses on the child's ability to comprehend and respond appropriately to words and requests, and to the number and type of words that must be recognised.
	Expressive Communication	This scale assesses the child's ability to name pictures of objects, to ask and answer more complex questions, combine words and gestures to communicate wants and needs, and use multiple-word sentences.
	Fine Motor	This scale examines the child's ability to manipulate objects through finely co-ordinated movements, to control hand-eye co-ordination, to grasp, to imitate precise strokes and to control speed of movement.
	Gross Motor	This scale examines the child's ability to demonstrate full body control in space, to plan and control movements, to maintain balance, to walk sideways or backwards, to jump, kick a ball, to stop from a full run.
Domains assessed through practitioner ratings	*Socio-emotional*	This scale assesses the child's ability to take actions to get their needs met, to use their imagination in play, to explain what they need and why, to describe how they feel and to use emotions in a purposeful manner.
	Communication	This scale assesses the child's ability to effectively communicate, to follow instructions, ask questions, describe activities and have more sustained conversations.
	Functional Pre-Academics	This scale assesses the child's emergent literacy skills, their ability to point to pictures in a book, hold a marker, imitate simple drawings, recognize and name shapes, name colours, recite rhymes, count objects using fingers.
	Leisure/Play	This scale assesses the child's ability to choose toys/games for play, to sustain play for a period, to join in and play with peers without adult supervision, to invite others to join in, to follow rules, to wait their turn.
	Self Direction	This scale assesses the child's ability to try out routine things without adult help, to persist with hard tasks, to follow routines without being reminded, to control their temper in the face of disagreements.
	Social Interaction	This scale assesses the child's ability to interact positively with other children and with adults, to share toys willingly, to show some degree of empathy with other children, to seek friendship with peers and show helping behaviour.

2.4.3 Parents and Practitioners

Parents and practitioner self-report questionnaires were designed for the study. Their purpose was to survey the perceptions of the children's parents and the practitioners in the early years setting about the developmental needs of 2-3 year olds, and their associated actions and interactions with the children. The survey was not intended to be fully comprehensive as it concentrated on the

areas that were the focus of the new Eager and Able to Learn programme – namely, play, movement and learning, and the adult-child interactions related to the children's social-emotional, physical and cognitive growth. In addition, as parental workshops were planned as part of Eager and Able to Learn programme, questions were included in both the parents' and the practitioners' survey about their current experiences and satisfaction with the level of communication and working partnerships between parents and early years settings.

Table 14 gives an overview of the topics that were covered in the survey, the numbers of questions per topic and the sources of the items. There were 89 items in the practitioner questionnaire and 79 in the parent version. Additional questions asked for demographic information, and for details of educational qualifications.

Table 14. **Practitioner and parent outcomes and related survey questions**

Outcome	Practitioner Survey	Parent Survey
Increased recognition of the importance and the different purposes of play in the development of two-year-old children; *and increased frequency in providing different types of play opportunities, both indoors and outdoors.*	9 questions eliciting viewpoints about play and the role of adults in children's play (Section 2) 10 questions asking about the frequency of different types of play opportunities and use of play materials in the early years setting during a typical week (Section 3) 13 questions about the frequency of different types of interactions during play in the early years setting during a typical day (Section 4)	8 questions eliciting viewpoints about play and the role of parents in children's play (Section 2) 17 questions about the frequency of providing play opportunities and play material for their child during a typical week (Section 3)

/Continued Overleaf

Table 14 (Continued). **Practitioner and parent outcomes and related survey questions**

Outcome	Practitioner Survey	Parent Survey
Increased responsiveness in practitioners'/parents' interactions and engagement with two-year-old children in order to support their communication, social, emotional, physical and cognitive development needs.	21 questions derived from the Child Caregiver Interaction Scale (Carl, 2007).[4] Items relate both to children's social and emotional development as well as to cognitive stimulation for learning (Section 6)	Four subscales (24 items) of the Tool for Parental Self-Efficacy, TOPSE (Kendall & Bloomfield, 2005; Bloomfield & Kendall, 2007)[5] were used – Emotion and Affection, Play and Enjoyment, Empathy and Understanding, Learning and Knowledge (Section 5)
Increased recognition of the importance of movement for two-year-old development and how it can be related to wider developmental goals (e.g. language, cognitive, social-emotional, as well as motor development)	19 questions about the importance of different types of movement and physical activities, and whether the activities were planned for in the early years setting (Section 5)	11 questions about parents' views on the role of movement and physical activities for children's learning (Section 4)
Increased recognition of the importance of working in partnership with practitioners/ parents around the developmental needs of two-year-old children, increased opportunities to communicate with parents, and increased satisfaction with the communication.	19 questions on current practices and levels of satisfaction in the early years settings on working with partners (Section 7)	19 questions on current practices and levels of satisfaction experienced by parents working with early years settings (Section 6)

[4] For the purposes of this study, practitioners were asked to rate themselves on 21 statements derived from the Child Caregiver Interaction Scale (Carl, 2007). This scale is based on developmentally appropriate principles as outlined by the National Association for the Education of Young Children (Bredekamp, 1997), now updated (NAEYC, 2009). Statements selected for self-rating cover adult-child interactions primarily related to the child's social and emotional development (11 statements) and interactions related to providing cognitive stimulation and support for the children's learning (10 statements).

[5] Four subscales of the Tool for Parental Self-Efficacy, TOPSE (Kendall & Bloomfield, 2005; Bloomfield & Kendall, 2007) were used to measure this outcome in parents. The TOPSE was developed specifically to evaluate the impact of parenting programmes in the UK and it is a very useful tool to evaluate parents' confidence about parenting, their beliefs about discipline and setting boundaries, as well as their sense of enjoyment about playing with their children and being sensitive and responsive to their children's needs.

2.5 Procedure

All data collection was undertaken by a team of fieldworkers who were fully trained and coordinated by the research team.

2.5.1 ECERS-R

Eight trained fieldworkers observed each Day Care setting for one full morning, or a full session in the case of Sure Start settings that were open for only a few hours each day. This was repeated twice, in March 2009 and in March 2010. Fieldworkers were trained by the UK national ECERS-R trainers, and their inter-rater reliability was checked before they began the first round of observations in March 2009. They received additional training to reliability, in preparation for the second round of observations in March 2010.

2.5.2 Bayley III

15-20 trained fieldworkers were trained as Bayley III assessors by the UK national Bayley trainer. They received refresher training in August 2009 before the second cohort of testing began. Each child for whom parental written consent had been provided was individually tested in the early years setting on the assessment tasks. Practitioners completed the individual child ratings during the period that fieldworkers were visiting the settings. In the control year, pre-testing was completed between October 2008 and Jan 2009, and in the intervention year, testing was completed by Nov 2009. For both cohorts, post testing was completed in May/June 2009 or 2010.

2.5.3 Practitioner and Parent Questionnaires

For the intervention cohort, pre-intervention questionnaires were distributed to practitioners during their initial training sessions and to parents at their first workshop. For the control settings, fieldworkers distributed questionnaires during their first visits. Post-intervention questionnaires were collected by the field workers during the final visit to the setting. Parents and practitioners could either return the questionnaire in a sealed envelope to the fieldworker or post it (freepost) back to the research team at Queen's. Special efforts were made by the early years staff to maximize the return of questionnaires, especially from parents in the control year, who were not attending workshops at that point in time.

2.6 Analysis

2.6.1 Setting Outcomes – Quality

For the ECERS-R measure, simple comparisons using repeated measures t-tests were made between the total scores and the subscale scores during the control year and the intervention year to evaluate the effect of the EAL intervention on the *general* quality of the setting. ECERS-R was not

intended to evaluate the specific implementation of the EAL programme or programme fidelity, which was examined in a separate study.

2.6.2 Child Outcomes

Because of the clustered nature of the data, the statistical analysis involved the use of multilevel models with children (level 1) clustered within settings (level 2). For each outcome, a linear multilevel model was estimated with the relevant post-test score being set as the dependent variable and the related pre-test score together with a dummy variable for whether the child was a member of the control or intervention group added as independent variables. Such models were used on the entire sample to estimate the main effects of the programme. The statistical significance of the coefficient for the dummy variable was used to test whether there was evidence of the programme having an effect.

For each outcome variable, the main effects models were then extended to consider whether the programme was having a differential effect in relation to:

- The level of the child's pre-test scores – high or low
- Gender
- Whether the setting was Sure Start or Day Care
- Whether the setting was Urban or Rural
- Quality of the setting as measured by ECERS-R (intervention year only)
- Programme fidelity as measured by the FIS study (intervention year only)
- Number of hours per week that the child attended the setting (intervention year only)

Full details relating to all of the multilevel models estimated are provided in the Appendices.

2.6.3 Practitioner and Parental Outcomes

A similar analysis, involving multilevel modelling, was used to analyse the effects of the programme on parents and practitioners. In this case, because of the smaller sub-samples, the analysis was restricted to a focus on the main effects of the programme.

2.6.4 Calculation of Effect Sizes

In relation to all of the outcome measures, where an effect was found to be significant the statistical models were used to calculate the post-test mean scores (in the case of continuous measures) once pre-test scores were controlled for.

The associated effect size measure used was the standardised mean difference calculated as the difference between the mean post-test scores for the control and intervention groups, once pre-test score differences were controlled for, divided by the pooled standard deviation for the post-test scores for both groups (i.e., Cohen's D).

3. Findings

3.1 Settings: ECERS-R comparison for EAL and Control Settings

3.1.1. Overall ECERS-R ratings

ECERS-R consists of 43 items and 7 subscales relating to Space and Furnishings, Personal Care Routines, Language-Reasoning, Learning Activities, Interaction, Program Structure and Parents and Staff. Ratings were completed on a 7-point scale ranging from Excellent (7) to 1 (Inadequate).

1	2	3	4	5	6	7
Inadequate		Minimal		Good		Excellent

Importantly, the meaning of the scale shows that a score of 3 shows a minimal level of quality provision. The scale is sometimes reported in terms of three 'quality bands' – less than adequate (rating <3), adequate (rating 3<5), and good plus (rating 5+)

Table 15 shows the mean ECERS-R ratings over all 7 subscales for the two observation times – in March 2009 before the EAL programme was introduced, and in March 2010 while the settings were participating in the EAL programme. On both occasions the average scores were between 3-4; the EAL settings scored higher than the Control settings and the difference between them was statistically significant (p=.069). When the ratings for individual settings were examined, 18 settings received improved ratings under the EAL programme, 7 settings received poorer ratings, and 3 remained substantially the same.

Table 15. **Mean ECERS-R ratings for Control and EAL settings (same settings at different times)**

ECERS-R	Control *Mean (SD) March 2009*	EAL *Mean (SD) March 2010*	**Significance**
Average Rating	3.44 (.81)	3.74 (.85)	.069

Beneath these average scores, there was wide variation between settings, for both the Control settings and for the EAL settings. For example, the range of scores for the Control settings was between 2.00 and 4.73. For the EAL settings, there was a general improvement but the variation was still wide - from 2.56 to 5.82.

Figures 1a and 1b show the distributions of ratings for individual settings at both time points.

Figure 1a*. Distribution of ratings for *Control Settings (March 2009)

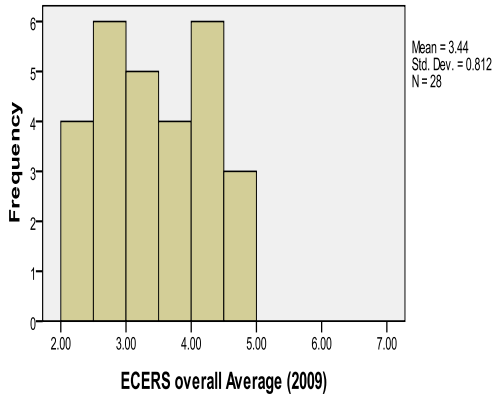
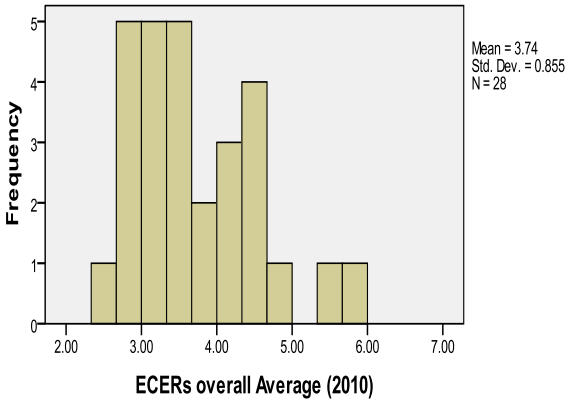


Figure 1b.* Distribution of ratings for *EAL Settings (March 2010)



3.1.2 ECERS-R and ITERS-R subscales for Control and EAL Settings

Table 16 shows the mean ratings for the 7 ECERS-R subscales and for the additional indicators from 3 ITERS-R subscales, for Control and EAL settings. Ratings on 9/10 subscales improved through participating in EAL, and 3/10 were statistically significant. There was no change in the ratings for Learning Activities.

EAL had a statistically significant impact on the ECERS-R subscale, Interactions, which included indicators relating to the quality of staff-children interactions (warmth, respect, appropriate discipline, supervision) as well as promoting positive child-child interactions (taking part, taking turns, management conflicts, including others). Also, Interactions was the highest rated ECERS-R subscale for both the Control and EAL settings (between 4 and 5) and was also rated high on the ITERS-R indicators (over 5). Provision for Parents and Staff (ECERS-R and ITERS-R) also improved significantly in EAL settings. EAL consisted of increased parent contact through workshops and home visits, and these clearly contributed to the improved quality ratings.

The scale that received the poorest rating (below the minimal quality of 3) was Learning Activities and these did not improve in the EAL settings. A substantial minority of settings had very low scores on this subscale.

ECERS-R subscales showed high internal reliability and the ITERS-R indicators correlated highly with the relevant ECERS-R subscales. The ECERS-R subscale, Personal Routines, showed slightly lower internal reliability, mainly due to the item about Naps and Rests. Not surprisingly, those settings

that were open for only a few hours per day scored poorly on this item, even if they were rated more highly on other items on that scale.

The focus of the EAL programme was on movement development, and settings received additional equipment to help with creating the movement activities. Nevertheless, the space is some settings was restricted and EAL settings continued to score poorly (mean=2.82) on the indicator, Space for Gross Motor Play (mean=2.82), with almost half of the EAL settings being rated below 3.

***Table* 16. Mean ratings for ECERS-R and ITERS-R subscales: Control vs. EAL settings**

Subscales	Control *Mean (SD) March 2009*	EAL *Mean (SD) March 2010*	**Significance**
ECERS-R			
Space and Furnishings	3.52 (1.12)`	3.82 (1.29)	.21
Personal Care Routines	3.00 (1.20)	3.45 (1.19)	.15
Language Reasoning	3.08 (1.01)	3.25 (1.01)	.37
Learning Activities	2.45 (.74)	2.45(.72)	.98
Interactions	4.60 (.93)	5.11 (.88)	**.02**
Program Structure	3.90 (1.37)	4.28 (1.26)	.20
Parents and Staff	4.32 (1.11)	4.83 (1.10)	**.04**
ITERS (selected items)			
Language	4.23 (1.24)	4.64 (1.20)	.17
Interactions	5.04 (1.35)	5.18 (1.05)	.63
Parents and Staff	4.72 (1.86)	5.68 (1.02)	**.02**

3.1.3 Key Findings

Participating in EAL improved the *average* quality for settings and this finding was marginally statistically significant. Not all settings improved; 18 improved, 7 got poorer, and 3 remained the same.

The most positive and statistically significant effects of EAL were on the subscales related to *interactions* between children and staff, *interactions* between the children, *interactions* between parents and staff and between the staff themselves. Settings were already scoring relatively high on these aspects of practice, receiving ratings between 4 and 5. Nevertheless, EAL provided an

additional boost, resulting in the average ratings moving beyond 5, and 20% of settings moved into the excellence range. The EAL programme consisted of increased contact with parents through workshops and home visits, and these clearly contributed to the improved quality ratings with regard to parents and staff.

Despite the emphasis on movement and gross motor activities in EAL, ECERS-R ratings for Space and Furnishings were not noticeably improved. Many settings had restricted space and this was reflected in the ratings for the indicator, Space for Gross Motor Play, which remained low, with almost half of the EAL settings being rated below 3. It should be remembered that ECERS was not intended to evaluate either the implementation or the impact of the specific EAL movement activities.

The lowest average quality ratings were for the ECERS-R subscale, Learning Activities, and these remained unchanged by EAL participation. This scale examines the extent to which settings engage the children with a range of stimulating materials and experiences, such as manipulatives, art materials, music, blocks, sand, water, drama, and so on. 75% of the ECERS-R ratings (Control and EAL combined) were below 3 on this subscale reflecting, perhaps, the absence of an orientation towards *using a broad range* of cognitive stimulation in the settings.

Based on the average ECERS-R ratings across all scales, and irrespective of whether the setting was a control or an EAL setting, the overall quality rating of the settings was between 3 and 4 on the ECERS-R 7-point scale. '3' indicates minimal/adequate quality and '5' indicates good. The average setting quality was more than adequate, but only approaching, good quality. There was wide variation between settings in terms of their quality. Of the 56 ECERS observations conducted (28 x 2 time points), 29% (16/56) were rated less than 3; 34% (19/56) were rated more than 4.

3.2 Children's Outcomes: Multilevel Modelling for Bayley III, Main Effects and Interaction Effects

3.2.1 Difference between EAL and Control children's Bayley scores at pre-test

Table 17 presents the pre-test mean scores for the eleven Bayley scales and the differences between them for children in the Control cohort (2008/2009) and in the EAL cohort (2009/2010). For 8 of the 11 child outcomes, there were no statistically significant differences between the mean scores pre-test scores for EAL and Control children. For the three remaining outcomes (Expressive Communication, Fine Motor, Gross Motor), the mean pre-test score for children in the Control group was significantly higher than for the EAL children. Significant differences are highlighted in bold.

These significant differences were surprising, given that the settings were identical for the EAL and Control samples and were drawing on the same neighbourhood/catchment areas over the two years. Also, and as shown above, the composition of the two cohorts on demographic variables was very similar (see Tables 2, 3 and 4). An important point to note is that the statistically significant differences were confined to the task-related Bayley scales that were fieldworker assessed, and therefore may partly relate to the level of experience of the fieldworkers at the first point of testing.

Whatever differences do exist will be controlled for in the statistical analysis and will not affect the interpretation of the findings.

Table 17. **Mean pretest scores for EAL and control group**

Outcome	Control *Mean (SD)*	EAL *Mean (SD)*	Significance *p*
Cognitive	10.46 (2.44)	10.07 (2.06)	0.252
Receptive communication	11.50 (2.43)	11.01 (2.78)	0.146
Expressive communication	11.34 (2.85)	10.50 (3.04)	***0.037***
Fine motor	11.62 (2.59)	10.89 (2.35)	***0.020***
Gross motor	9.73 (2.97)	8.82 (2.66)	***0.035***
Social-emotional	10.85 (3.52)	10.23 (3.39)	0.317
Communication	9.81 (3.22)	9.74 (3.25)	0.756
Functional Pre-Academics	9.53 (3.01)	9.58 (2.97)	0.903
Leisure	9.50 (3.18)	9.63 (3.29)	0.920
Self direction	10.30 (3.74)	10.34 (3.67)	0.608
Social	9.43 (3.45)	9.61 (3.58)	0.905

*Significance values (*p*) have been calculated using multilevel regression models in order to take into account the clustering of the data by setting

3.2.2 The impact of the EAL programme on Bayley scores: EAL vs Control

The main effects analysis was conducted using multilevel linear regressions for each outcome with children (level one) nested within settings (level two). For each model, the response variable consisted of the respective post-test score and two predictor variables were included: the relevant pre-test score and a dummy variable distinguishing between intervention/EAL and control settings.

Following on from the analyses of the main effects, a series of interaction analyses were conducted to see if the programme had different effects for subgroups of children in the sample. The largest number of statistically significant (or approaching significant) interaction effects was associated with the level of the pre-test scores of the children – that is, the developmental stage of the child at entry to the programme (Model 2, Tables 1-11, Appendix 3). Even although these interactions are

exploratory, they will be reported alongside the main effects for each of the Bayley outcomes, as they can modify the interpretation of the main effects.

Table 18 presents the key findings from the main effects analyses. There were three statistically significant effects - Functional Pre-Academics (p<.05), Cognitive (p=.058) and Social Emotional (p=.07), and these are highlighted in bold. Statistically significant interaction effects were associated with Social Emotional, Receptive Communication, Fine Motor and Leisure, and these effects are illustrated graphically in Figures below.

Table 18. **Mean posttest scores for EAL and control group, controlling for pretest differences[1]**

Outcome	Control *Mean (SD)*	EAL *Mean (SD)*	Significance *p*	**Effect Size**	**95% confidence interval**
Cognitive	10.95 (2.85)	10.27 (1.90)	0.058	***-0.29****	-0.582, +0.010
Receptive communication	11.83 (2.06)	11.67 (2.03)	0.522	-0.07	-0.301, +0.153
Expressive communication	11.47 (2.68)	11.63 (2.36)	0.574	+0.06	-0.156, +0.281
Fine motor	11.85 (2.85)	11.52 (2.12)	0.205	-0.13	-0.341, +0.073
Gross motor	10.46 (3.02)	10.48 (2.39)	0.957	+0.01	-0.283, +0.299
Social-emotional	12.32 (3.64)	13.46 (3.96)	0.071	***+0.30****	-0.025, +0.619
Communication	10.25 (3.29)	10.80 (3.19)	0.212	+0.17	-0.096, +0.433
Functional Pre-Academics	10.54 (3.06)	9.72 (2.56)	0.043	**-0.29****	-0.580, -0.009
Leisure	10.65 (3.63)	10.92 (3.75)	0.662	+0.07	-0.254, +0.400
Self direction	11.95 (4.09)	12.50 (4.10)	0.401	+0.13	-0.179, +0.448
Social	10.60 (3.92)	11.26 (3.96)	0.276	+0.17	-0.134, +0.468

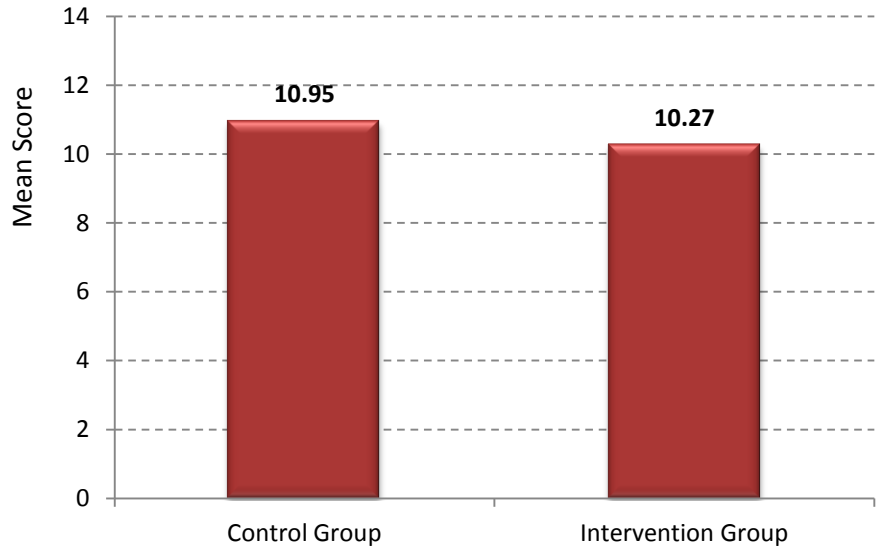
[1]Post-test mean scores and significance of differences estimated using multilevel regression models to take into account the clustered nature of the data. Effect sizes were calculated by dividing the difference in adjusted 27/mean scores by the pooled standard deviation of the relevant post-test score for the sample as a whole.

Signficance levels *p<10; **p<.05; ***p<.01

Cognitive: For the Cognitive scale (field-worker assessed), the EAL programme had a negative effect (effect size- -0.29, p=.058). The Control group scored higher than the EAL group, M=10.95 vs. 10.27. This effect is shown graphically in Figure 2. There was no interaction effect associated with pre-test

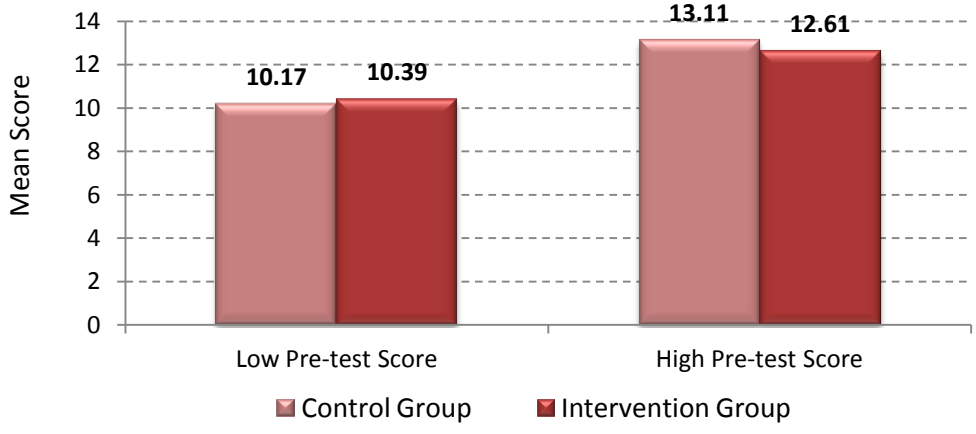
scores, indicating that the programme had similar effects irrespective of the children's pre-test scores.

***Figure 2*. Mean Cognitive post-test scores for children in the control and intervention groups, controlling for pre-test differences**



Receptive Communication: For the Receptive Communication scale (field-worker assessed), there was no statistically significant main effect of the EAL programme. There was a statistically significant interaction between the effect of EAL and the children's pre-test score, p=.027 (Model 2, Table 2, Appendix 3). Figure 3 shows that the programme had a positive effect on children with lower pre-test scores (effect size= +.11) and a negative effect on children with higher pre-test scores (effect size = -.24).

Figure 3*. Mean post-test Receptive Communication scores for children with initially low and high pre-test scores

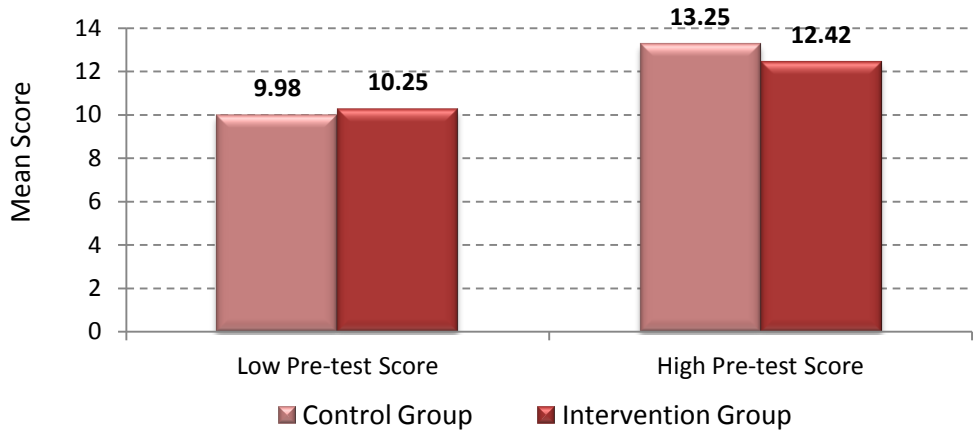


*Low and high scores defined as those one standard deviation below and above the mean score at pretest respectively.
Effect sizes: +.11 (low pre-test scores) and -.24 (high pre-test scores)

Expressive Communication: As can be seen from Table 18, there was no statistically significant main effect for the Expressive Communication scale (field-worker assessed), and no interactions associated with the pre-test scores.

Fine Motor: There was no statistically significant mean effect of the EAL programme on the Fine Motor scale (field-worker assessed). There was a statistically significant interaction between the effect of EAL and the children's pre-score, p=.01 (Model 2, Table 4, Appendix 3). Figure 4 shows that the programme had a positive effect on children with lower pre-test scores (effect size= +.11) and a negative effect on children with higher pre-test scores (effect size = -.34)

Figure 4*. Mean post-test Fine Motor scores for children with initially low and high pre-test scores



*Low and high scores defined as those one standard deviation below and above the mean score at pretest respectively.
Effect sizes: +.11 (low pre-test scores) and -.34 (high pre-test scores)

Gross Motor: As can be seen from Table 18, the EAL programme had no statistically significant main effect on the Gross Motor (field-worker assessed) scale, and there were no interactions associated with pre-test scores.

Social-Emotional: For the Social-Emotional scale (practitioner-rated), there was a positive main effect of the EAL programme (effect size = +.30, p=.07). Figure 5 shows the main effect. There was also a significant interaction effect, p=.07 (Model 2, Table 6, Appendix 3), showing that the positive effect was greater for children with higher pre-test scores (effect size=+.47) than for those with lower pre-test scores (effect size +.11), see Figure 6.

Figure 5. **Mean Socio-Emotional post-test scores for children in the control and intervention groups, controlling for pre-test differences**

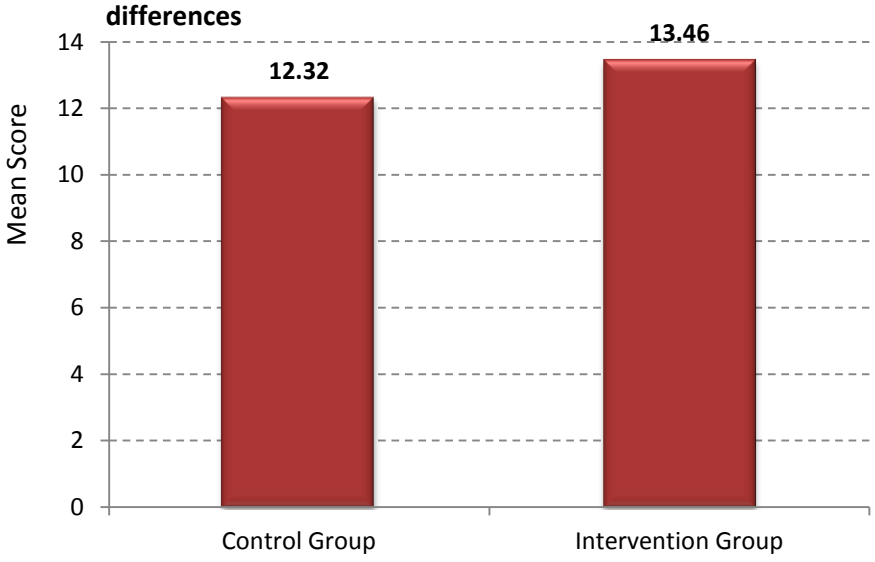
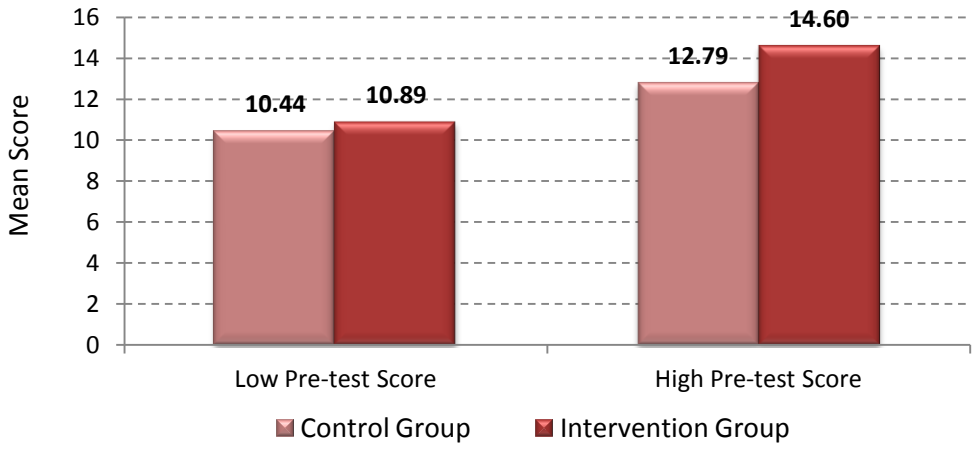


Figure 6. **Mean post-test Socio-Emotional scores for children with initially low and high pre-test scores***



*Low and high scores defined as those one standard deviation below and above the mean score at pretest respectively.
Effect sizes: +.12 (low pre-test scores) and +.47 (high pre-test scores)

Communication: For the Communication scale (practitioner-rated), there was a positive main effect (effect size=+.17) which did not approach significance. There was no effect associated with pre-test scores.

Functional Pre-Academics: For the Functional Pre-Academics scale (practitioner-rated), there was a statistically significant main effect, p=.043, which shows that, compared to the control group, the EAL programme had a negative effect (effect size=-0.29). The Control children (M=10.54) outperformed the EAL children (M=9.72) in this domain. Figure 7 shows this graphically. Figure 8 shows the associated interaction with pre-test scores, p=.160, showing a greater negative effect on children with higher pre-test scores (effect size = -.42) compared to children with lower pre-test scores (effect size= -.17), although this difference was not statistically significant.

Figure 7. **Mean Functional Pre-Academics post-test scores for children in the control and intervention groups, controlling for pre-test differences**

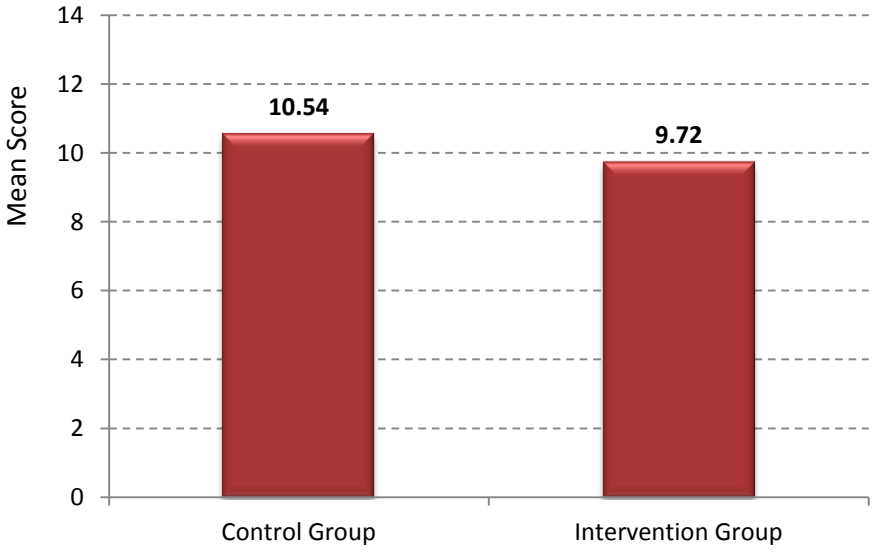
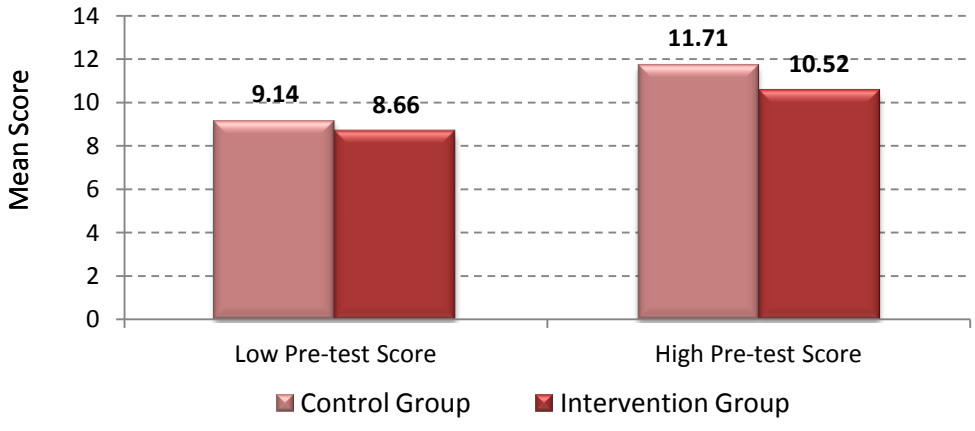


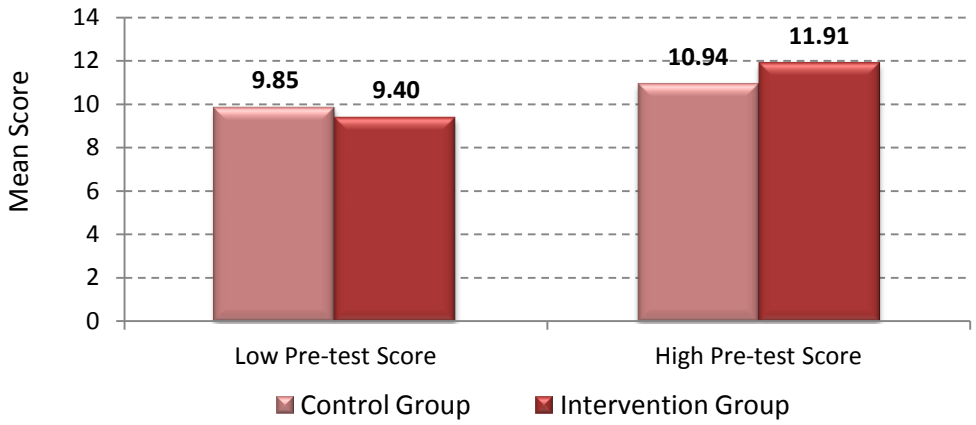
Figure 8. **Mean post-test Functional Pre-Academics scores for children with initially low and high pre-test scores***



*Low and high scores defined as those one standard deviation below and above the mean score at pretest respectively.
Effect sizes: -.17 (low pre-test scores) and -.42 (high pre-test scores)

Leisure: For the Leisure scale (practitioner-rated), there was no main effect of the programme, but there was a statistically significant interaction, p=.045 (Model 2, Table 9, Appendix 3). Figure 9 shows that, on this occasion, the programme had a positive effect (effect size +.26) on children with higher pre-test scores and a negative effect on those with lower pre-test scores (effect size= -.12).

Figure 9*. Mean post-test Leisure scores for children with initially low and high pre-test scores



*Low and high scores defined as those one standard deviation below and above the mean score at pretest respectively.
Effect sizes: -.12 (low pre-test scores) and +.26 (high pre-test scores)

Self Direction: There was a positive effect (effect size = +.13) for the Self-Direction scale which did not approach significance and no interaction effects.

Social Interaction: There was a positive effect (effect size = +.17) for the Social Interaction scale which did not approach significance and no interaction effects.

Overall, there were no statistically significant interaction effects associated with Gender (Model 3, Tables 1-11, Appendix 3), that is, the programme had similar effect for both boys and girls. There were no statistically significant interaction effects associated with type or location of setting, that is, the programme had similar effects in Sure Start and Day Care settings (Model 4, Tables 1-11, Appendix 3), and in Urban and Rural settings (Model 5, Tables 1-11, Appendix 3).

3.2.2 The effects of selected setting-related variables in the Intervention Group Only

Several interaction analyses were conducted to examine the effects of setting-related variables on children's post-test scores, controlling for pre-test scores. These analyses were completed for the

EAL group only, as the variables were relevant particularly to the EAL cohort of children. The setting variables were: (1) Programme Fidelity as measured by the Fidelity Implementation Study; (2) Quality of Setting as measured by total score on ECERS-R; and (3) Number of hours per week that the child attended the setting,

Programme Fidelity

There were statistically significant interaction effects of programme fidelity (Model 7) on two Bayley scales – Receptive Communication (p=.039, Model 7, Table 2, Appendix 3) and Expressive Communication (p=.032, Model 7, Table 3, Appendix 3), see Figures 10 and 11. Both showed that greater programme fidelity (higher percentiles) had a positive effect on children's communication. However, it should be noted that for Receptive Communication, even the post-test scores associated with the higher levels of fidelity only equaled the average scores for the Control group (M=11.83).

***Figure 10*. The relationship between fidelity scores of settings and mean post-test receptive communication scores of children, controlling for pre-test scores**

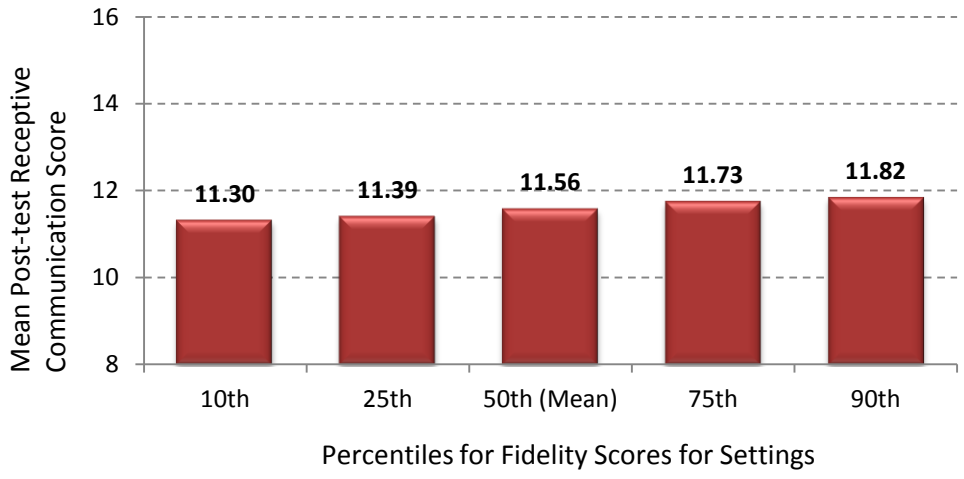
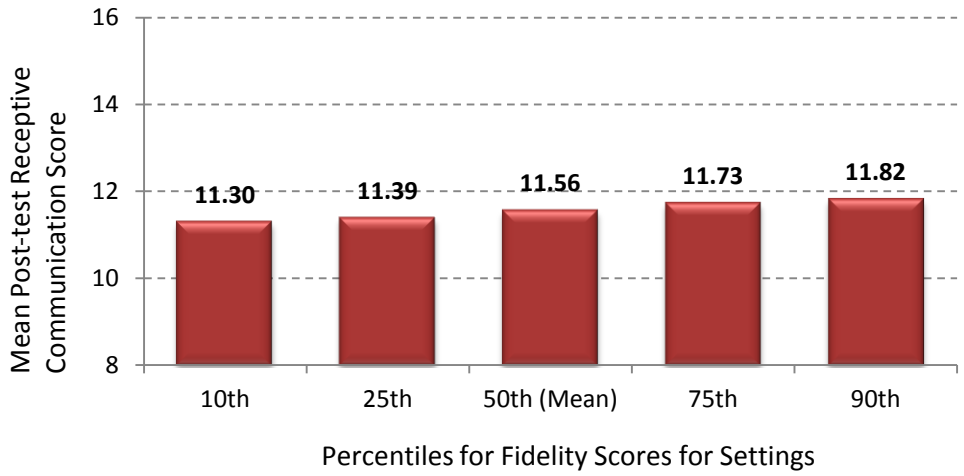


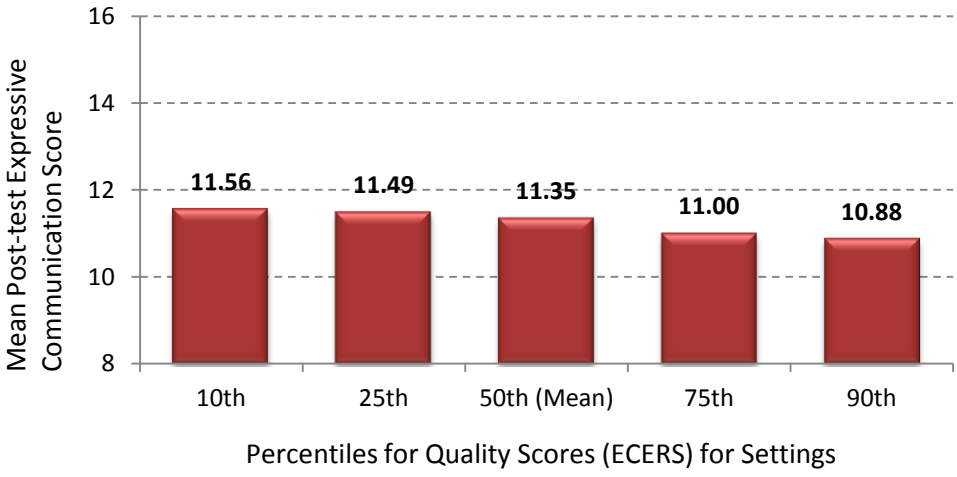
***Figure 11*. The relationship between fidelity scores of settings and mean post-test receptive communication scores of children, controlling for pre-test scores**



Quality of Setting

There was a statistically significant interaction effect of Quality of Setting as measured by the total average ECERS-R rating on only one Bayley scale –Expressive Communication (p=.039, Model 6, Table 3, Appendix 3), see Figure 12 below. Surprisingly, the effect of quality of setting had a negative effect on children's post-test scores. Higher quality settings were associated with poorer children's post-test scores. This effect requires further investigation.

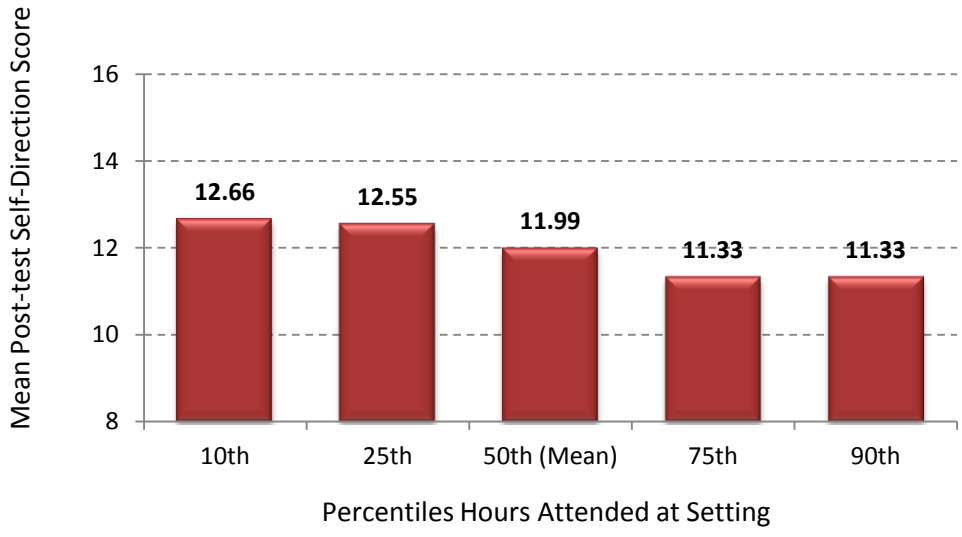
***Figure 12*. The relationship between quality scores of settings and mean post-test expressive communication scores of children, controlling for pre-test scores (intervention settings only)**



Hours in Setting

Because some settings were open for full days and some for 2-3 hours per day, and because some children attended part-time and others full-time, the number of hours attended per week for each child was tracked during the Intervention year, on the expectation that the number of hours attended would indicate the degree of exposure to the programme experienced by the child. The hours spent in the setting had a statistically significant effect on only one variable – Self-Direction (p-.038, Model 8, Table 10, Appendix 3), where the effect was negative – children who spent very high number of hours in the setting had poorer self-direction scores, indicating perhaps that such children were over-supervised and failed to develop appropriate levels of independence and persistence with tasks without help, see Figure 13. This finding is probably unrelated to the EAL programme. As the hours in setting were not tracked for the children in the control cohort, it is not possible to draw firm conclusions about this.

***Figure 13*. The relationship between hours attended at setting by children and their mean self-direction scores, controlling for pre-test scores**



3.2.3 Key Findings on the effects of EAL on Children's Outcomes

The impact of the EAL programme on the developmental outcomes for children as assessed by Bayley was not consistent and presented a rather surprising pattern. There were both statistically significant positive and negative *overall* effects as well and positive and negative effects for subgroups of children. The main subgroup variation was due to the developmental stage of the children at entry into the programme – pre-test scores – rather than to factors such as gender, type or location of early years setting attended.

The main positive effects of the EAL programme were on social emotional development for all children (effect size=+.30), with some indication of positive effects on their social skills (effect size=+.17) and self-direction (effects size=+.13). Those who were more developmentally advanced when they joined the programme seemed to benefit most, with a larger effect size (effect size=+.47) in the social emotional domain and additional positive effects noted on play-related behaviours from the Leisure scale (effect size=+.26)

The main negative effects were on Functional Pre-Academics (effect size=-.29) and on the Cognitive scale (effect size=-.29) for all children. Functional Pre-Academics was particularly negatively affected for those who were more developmentally advanced when entering the programme (effect size=-.42).

Thus, high pre-test scoring children experienced both positive and negative effects of the programme. The programme positively affected their social emotional development (effect size=-.47) and their play=related/ behaviour (effect size=+.26, Leisure scale), with some positive effects on social skills and self-direction. In contrast, the programme negatively affected their emergent literacy skills (Functional Pre-Academics, effect size=-.42), their cognitive development (Cognitive Scale, effect size =-.29), with specific additional negative effects associated with Receptive Communication (effect size=-.24), Fine Motor (effect size=-.34)

Lower pre-test scoring children also experienced positive and negative effects of the programme. In general, the effect sizes were smaller for this group of children and while they mirrored some of the effects for the higher pre-test scoring children, the pattern was distinctly different. For example, the EAL programme had a positive effect on social emotional development (effect size= +.12) and on social skills and self-direction to a lesser extent. But it also positively affected Receptive Communication (effect size=+.11) and Fine Motor development (effect size = +.11) Similar to the more developmentally advanced children, there were negative effects for Functional Pre-Academics (effect size= -.17) and for the Cognitive scale, but unlike the more advanced children, play-related behaviours were also negatively affected (effect size=-.12).

The main focus of the programme on developmental movement activities had no effects on the children gross motor development.

For the EAL cohort, setting variables that were expected to influence the children's outcomes had limited effects on a restricted range of Bayley scales. For example, programme fidelity measures had a positive influence on Receptive and Expressive Communication, but there were no significant overall differences between the Control and EAL cohorts on these scales. Quality as measured by ECERS-R while the children were experiencing the EAL programme had a negative influence on Expressive Communication- which needs further investigation and explanation. Higher number of

hours spent in the setting had a negative effect on the children's capacity to self-direct their behaviour and to develop independence and work without adult supervision.

3.3 Outcomes for Practitioners: Multilevel Modelling of Practitioner Questionnaire Responses

Multilevel modelling was used to analyse the effects of the programme on the practitioners' self-reported responses to pre and post intervention questionnaires. In this case, because of the smaller sub-samples, the analysis was restricted to a focus on the main effects of the programme. 89 multilevel modelling analyses were conducted on the individual questionnaire items, and 18 analyses produced statistically significant effects.

Table 19 shows the post-test means for ratings by the Control and EAL practitioners, controlling for differences in pre-test responses. The questionnaire items are displayed under the general domains that were related to the practitioner outcomes for the study. The evidence suggests that the training and experience of participating in the EAL programme had positive effects on the practitioners' beliefs, attitudes and behaviours that were consistent with the training they had encountered. For example, with regard to providing new and different opportunities and materials for play (Section 3), EAL practitioners reported that they were using more 'everyday' materials' (e.g., pots, pans, crumpled papers) and props to help with movement games (e.g., scarves, balls, hoops). They were less likely to use number games. Also, it was clear that practitioners began to plan movement activities more deliberately (Section 5) and with more specific purposes.

Questions in Section 6 had focused specifically on the practitioners' general style of interacting with the children. The responses to a series of questions in this section showed that, compared to the control year, the EAL practitioners were much less likely to adopt 'harsh and controlling' interaction styles with the children and more likely to explain the reasons for things in order to encourage the children to think for themselves. Two items in this section showed slightly anomalous results. For example, control practitioners continued to confirm that they would 'put toys and objects out of sight when children lose interest in them' (-.47) and that they 'would use books and pictures for story-telling so that the children can understand what books are for' (-.45) more frequently than EAL practitioners. The first item was designed to see if practitioners had strategies to re-motivate children's interest but the item could also be interpreted as overly controlling. The second item re books was designed as a measure of cognitive stimulation, but may have been displaced with the new emphasis on the EAL developmental movement experiences.

Also, consistent with the new partnerships arrangements with parents during the EAL year, practitioners were more positive about how their setting worked with parents, and less doubtful than they were during the control year about the contribution that parents can make to support their children's learning in the setting , see Section 7.

Table 19. **Practitioners' Questionnaire: Mean post-test ratings for Control and EAL practitioners[1]**

Practitioner Questionnaire Items	Control Mean (SD)	EAL Mean (SD)	Effect Size (Significance)
Section 2 Viewpoints on Play			
Q2.3 Children's play needs to be stimulated and extended by adults	4.54 (.70)	4.29 (.71)	-.35*
Section 3 Different Opportunities and Materials for Play			
Q3.2 Provide play materials –crumpled papers, pots, pans, cardboard boxes, etc	3.69 (.63)	3.89 (.33)	+.38**
Q 3.6 Use simple props to play movement games – scarves, balls, hoops	3.33 (.78)	3.79 (.50)	+.68***
Q 3.10 Use number games	3.13 (.96)	2.61 (1.13)	-.48**
Section 4 Interacting with Children during Play			
Q 4.6 Accept invitations to join in children's play	5.00 (.32)	4.86 (.39)	-.38*
Section 5 Movement and Learning			
Q 5.3a Develop balance – importance	4.41 (.83)	4.83 (.59)	+.56***
Q 5.3b Develop balance – plan activities	4.44 (.71)	4.78 (.58)	+.51**
Q 5.4b Develop body sense –plan activities	4.46 (.86)	4.73 (.62)	+.35*
Q.5.6b Use movement as a form of play – plan activities	4.65 (.60)	4.85 (.54)	+.34**
Q 5.8 Use movement as a way to help children express and communicate feelings – plan activities	4.33 (.88)	4.62 (.79)	+.34*

/Continued Overleaf

Table 19 (Continued). **Practitioners' Questionnaire: Mean post-test ratings for Control and EAL practitioners[1]**

Practitioner Questionnaire Items	Control Mean (SD)	EAL Mean (SD)	Effect Size (Significance)
Section 6 Your own Style as a Practitioner			
Q 6.3 I expect the children to be obedient and to follow a strict routine	2.67 (.97)	2.08 (1.09)	*-.57***
Q 6.7 Despite my best intentions, I can get irritated and be impatient with children	1.90 (.96)	1.63 (.88)	*-.29**
Q 6.11 I put toys and objects out of sight when children lose interest, and bring them out again at a later point	4.38 (.87)	3.94 (.96)	*-.47***
Q 6.13 I try to explain the reasons for things in order to encourage the children to think for themselves	4.33 (.83)	4.59 (.63)	*+.35***
Q 6.14 I can be abrupt with the children when they misbehave	2.00 (1.01)	1.57 (.93)	*-.44***
Q 6.18 I use books and pictures for story-telling so that the children can understand what books are for	4.87 (.48)	4.62 (.62)	*-.45**
Section 7 Working in Partnership with Parents			
Q 7.3 To what extent do you think that the parents and your setting work in partnership to promote children's learning and development?	4.32 (.72)	4.54 (.67)	+.32*
Q 7.11 Parents are too busy to support their children's learning. That is what we as practitioners do best.	2.01 (.87)	1.65 (.79)	-.43**

[1]Post-test mean ratings and significance of differences estimated using multilevel regression models to take into account the clustered nature of the data. Effect sizes calculated by dividing the difference in adjusted mean ratings by the pooled standard deviation for sample as whole at post-test.

Signficance levels *p<10; **p<.05; ***p<.01

3.4 Outcomes for Parents: Multilevel Modelling of Parent Questionnaire Responses

Multilevel modelling was used to analyse the effects of the programme on the parents' self-reported responses to pre- and post-intervention questionnaires. In this case, because of the smaller sub-samples, the analysis was restricted to a focus on the main effects of the programme. 57 multilevel modelling analyses were conducted on the individual questionnaire items, and 15 analyses produced statistically significant or approaching significant results.

Table 20 shows the post-test means for ratings by the Control and EAL parents, controlling for differences in pre-test responses. The questionnaire items are displayed under the general domains that were related to the parent outcomes for the study. The evidence suggests that the workshops and experience of participating in the EAL programme had positive effects on the parents' knowledge, beliefs and behaviours that were consistent with aims of the parental involvement in the programme.

Table 20. **Parents' Questionnaire: Mean post-test ratings for Control and EAL parents, controlling for differences at pre-test[1]**

Parent Questionnaire Items	Control *Mean (SD)*	EAL *Mean (SD)*	**Effect Size (d)[2]** **(Significance)**
Section 2 Children, Parents and Play			
Q 2.3 Some children find it hard to know how to play with other children	3.70 (.80)	3.85 (.81)	+.19*
Q 2.6 Parents should join in and play alongside the children to make the most of their play	3.94 (.84)	4.15 (.75)	+.26**
Section 3 Types of Play			
Q 3.3 Use songs and dance in a playful way	3.27 (.77)	3.46 (.60)	+.27**
Q3.13 Encourage my child to play with materials that make different sounds, have different colours, or have a different feel to them	2.67 (.99)	2.85 (.84)	+.19*
Section 4 Physical Movement and Physical Activity			
Relevance in terms of….			
Q 4.1 Learning new words	1.55 (.88)	1.25 (.60)	+.40***
Q 4.2 Helping children's imagination	1.20 (.49)	1.07 (.27)	+.33**
Q 4.5 Helping children to better understand the world around them	1.28 (.61)	1.09 (.32)	+.40***
Q 4.9 Helping children to be better at reading in the future	2.33 (1.24)	1.86 (.99)	+.42***
Q 4.10 Helping children to be better at writing in the future	2.35 (1.25)	1.89 (.95)	+.40***
Q 4.11 Helping children to be better at problem solving	1.74 (.90)	1.44 (.69)	+.37***
Section 5 You and Your Child			
Q 5.1-5.6 Parental Self Efficacy, Emotion and Affection	8.02 (.47)	7.92 (.43)	-.22*

/Continued Overleaf

[1]Post-test mean ratings and significance of differences estimated using multilevel regression models to take into account the clustered nature of the data. Effect sizes calculated by dividing the difference in adjusted mean ratings by the pooled standard deviation of the relevant post-test rating for the sample as a whole.

[2]For clarity of presentation, the direction of the effect (plus or minus) is displayed to ease the interpretation of the data for the reader and takes into account those items where lower scores by the EAL group show positive effects (e.g., Qs 4.1 to 4.11)

Signficance levels *p<10; **p<.05; ***p<.0

Table 20 (Continued). **Parents' Questionnaire: Mean post-test ratings for Control and EAL parents, controlling for differences at pre-test[1]**

Parent Questionnaire Items	Control *Mean (SD)*	EAL *Mean (SD)*	**Effect Size (d)[2] (Significance)**
Section 6 You and Your Child's Setting			
Q 6.8 Parents can disturb the children if they are frequently in the setting	3.31 (1.14)	3.62 (.88)	+.31**
Q 6.10 The setting staff talk to me about my child's development and what we can do to help this along	3.98 (.93)	4.22 (.88)	+.27**
Q 6.12 I am given the opportunity to share my views, concerns and wishes with the staff in my child's setting	4.43 (.74)	4.58 (.62)	+.22**
Q 6.18 My child's setting provides me with materials and training to help my child's development	3.48 (1.16)	4.06 (.97)	+.51***
Q 6.19 Overall I am satisfied with the level and quality of communication I have with child's setting	4.40 (.78)	4.59 (.67)	+.26**

[1]Post-test mean ratings and significance of differences estimated using multilevel regression models to take into account the clustered nature of the data. Effect sizes calculated by dividing the difference in adjusted mean ratings by the pooled standard deviation of the relevant post-test rating for the sample as a whole.

[2]For clarity of presentation, the direction of the effect (plus or minus) is displayed to ease the interpretation of the data for the reader and takes into account those items where lower scores by the EAL group show positive effects (e.g., Qs 4.1 to 4.11).

Signficance levels *p<10; **p<.05; ***p<.01

For example, compared to the Control parents, the EAL parents showed a new sensitivity to the difficulties that some children might have when playing with other children, and that parents might need to join in to help children make the most of their play (Section 2). EAL parents also reported playing with their children in different types of ways – with song and dance, and using different materials (Section 3). EAL parents showed a sharper recognition than Control parents about the relevance of play to different forms of learning both in the present and in the future (Section 4).

Parents were also substantially more satisfied than Control parents with communicating and sharing views with staff in the early years settings and also reported more help with materials and training for promoting their child's development (Section 5).

However, in terms of overall self-efficacy as measured by the Parental Self-Efficacy Scale, there was very little difference between the two groups of parents, and Control parents scored marginally higher then EAL parents on expressing emotion and affection to their children, though the scores were high for both groups.

4. Summary and Conclusions

This section summarises the key findings from the trial and considers the implications of these for the further development of the Eager and Able to Learn programme.

4.1 Preliminary Points

EAL is a newly designed service for 2-3 year olds and was piloted by Early Years in 2008-2009 in a small group of seven settings. The current full evaluation was conducted in 2009-2010 on a new group of 28 settings. Thus, the EAL intervention settings described in this report were implementing the programme for the first time. The programme ran for 8-9 months.

The programme focussed on a series of twelve developmental movement experiences and on promoting positive interactions between the practitioners and children in the settings, and between the practitioners and parents. The movement experiences were designed to create opportunities to extend learning and development into other domains beyond the physical. Hence, accompanying each movement experience was a fan of four learning experiences – physical, cognitive, language and social-emotional. Consequently, the expected child outcomes covered the full range of developmental domains usually associated with child learning and development – cognition and language, social-emotional, physical and motor (National Research Council, 2001, Chapter 3). The new Bayley III was chosen as the assessment tool that best mapped onto the expected child outcomes.

As well as becoming more knowledgeable about movement and its role in learning, the outcomes for practitioners also expected that the practitioners would acquire new understandings of the meaning and role of play for two-year olds and become more responsive in their interactions with the children. For parents, the home-package of play activities and the home visits were expected to help parents to recognise the importance of play, to diversify the opportunities for different types of children's play, as well as to increase their general responsiveness and empathy with their children. For both parents and practitioners, it was hoped that the programme would help communication between parents and the early years settings in the interests of the children's learning, as well as create greater feelings of satisfaction about working in partnership with one another. The EAL programme was ambitious in the outcomes it expected for children and adults.

It is important to note that the current trial evaluates the immediate post-intervention effects of the programme – not any longer term effects as the children moved into pre-school settings or onto statutory schooling. One of the strengths of the evaluation is that many different types of evaluation data was collected – performance data, ratings, structured observations, self-report questionnaires, demographic information – and from many different participants, the children, the parents, and the practitioners. The challenge for interpretation is to create a composite picture if the data from different sources appears to conflict.

4.2 Summary of findings in relation to setting outcomes

The key findings with regard to the impact of the EAL programme on the quality of the settings were:

4.2.1 There is some evidence that the EAL programme improved the quality of the settings

Participating in EAL improved the *average* quality for settings. The average ECERS-R rating changed from 3.44 to 3.74, shifting the average quality of the settings to the higher end of the 'adequate' quality band. Additionally 4 settings moved from the inadequate band (<3) to the adequate band (3<5) and two settings moved from adequate to the good band (5+). 6/7 ECERS-R subscales showed some evidence of improvement. Not all settings improved; 18 improved, 7 got poorer, and 3 remained the same.

4.2.2 There is strong evidence that the EAL programme improved the social interactions between children and staff and between parents and staff

The most positive and statistically significant effects of EAL were on the subscales related to *interactions* between children and staff, *interactions* between the children, *interactions* between parents and staff and between the staff themselves. Settings were already scoring relatively high on these aspects of practice, receiving ratings between 4 and 5. Nevertheless, EAL provided an additional boost, resulting in the average ratings moving beyond 5, and 20% of settings moved into the excellent range, confirming that the EAL settings were most successful in having warm and respectful relationships with the children, helping the children get along with their peers, and providing appropriate levels of discipline. The EAL programme consisted of increased contact with parents through workshops and home visits, and these clearly contributed to the improved quality ratings with regard to parents and staff. These findings are corroborated by the findings from the practitioners' and parents' survey questionnaires.

4.2.3 There is strong evidence that the overall quality of settings in the sample is not satisfactory

Irrespective of the positive impact of the EAL programme, the quality of the settings as rated by ECERS-R was relatively low, with the average ratings remaining in the 'adequate' band (3<5). When benchmarked against other Northern Ireland/UK studies of settings catering for same-aged or slightly older children, the current ratings are notably poorer across all quality areas that are rated by both the ECERS-R and ITERS-R scales. International comparisons also confirm wide variation in the quality of settings both for infants/toddlers and for pre-school children, with generally poorer quality reported for younger children. This may be due to contextual factors such as country/state's early years policy, the development of services, accreditation/licensing arrangements, stage of expansion of services, and so on.

4.3 Summary of findings in relation to child outcomes

Table 21 provides a summary of the findings arising from the trial evaluation in relation to each of the outcomes identified for the children. The key points to note are as follows:

4.3.1 Evidence for the impact of the EAL programme across the expected broad range of developmental domains for all children produced some unexpected and puzzling findings. For some developmental domains, the programme had a positive impact on the children's development, for other domains it had a negative impact, and in some domains it had no effect.

Table 21 shows how the eleven Bayley measures mapped on to the five identified outcomes for children. There were significant statistical effects on 8/11 measures – either as main effects, as interactions, or both. Many of the effect sizes were substantial. Effect sizes for observed main effects ranged from +.30 to -.29. Effect sizes for interactions varied from -.17 for one subgroup in an interaction to +.47 in another subgroup. The expectation that the programme would have a broad positive effect across the developmental domains for all children was not confirmed.

4.3.2 The strongest evidence that the EAL is effective in improving outcomes for children is in relation to socio-emotional development

All children who participated in the EAL programme improved in this domain. The socio-emotional status of the children is one of the new scales to be included in Bayley III and is based on

Greenspan's functional emotional milestones. Functional emotional milestones are distinguished from specific emotions and social skills and are defined in the Bayley manual as incorporating milestones such as "using symbols and ideas to convey increasingly complex intentions and feelings, dealing with increasingly complex emotional themes, forming local bridges between emotions and ideas and forming logical bridges between their own emotional ideas and those of others". In other words, the children showed improved ability to take actions to get their needs met, to explain what they needed and why, to use their imagination in play, to describe how they feel and use emotions in a purposeful manner. In addition, some improvement was observed in two other domains that were separately assessed – social skills, where their ability to interact positively with other children and with adults improved, and self-direction, where they showed increased independence and ability to manage their emotions in the face of frustrations. These domains were all rated by the practitioners in the early years settings.

Although the programme was not explicitly about social-emotional development, the movement experiences clearly provided new opportunities for staff and children to interact positively (confirmed from the ECERs-R setting data), and encouraged children to participate in an extended range of activities, to experience and express a range of emotions (e.g., enjoyment, excitement, fear, reluctance) and to play with other children. The fan of experiences in the socio-emotional domain was also likely to help children develop vocabulary to label emotions and feelings. Also, those children who were more developmentally advanced in this domain at the point of entry to the programme seemed to improve most. These children were probably best positioned to take advantage of the social-emotional opportunities provided by the programme.

Table 21. **Effects of the Eager and Able to Learn programme on Children**

Outcomes	Main Effects	Differential Effects for Specific Subgroups[a]
Language and Communication Skills, Vocabulary		
Receptive Language	**No Effect**	**Pretest scores** (Low +.11; High -.24)** No gender differences No SS/DC differences No Urban/Rural differences **Difference re Fidelity**** No difference re Quality of Setting or Hours Attend
Expressive Language	**No Effect**	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences **Differences re Fidelity**** **Differences re Quality of Setting** (negative)** No difference re Hours attended
Communication Skills	**No Effect**	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Social/Emotional Skills and Behaviours, Independence, Self-Help Skills		
Social Emotional Milestones	**+.30***	**Pre-test scores* (Low +.12; High +.47)** No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Social Skills	**No Effect**	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Play Behaviours/Leisure	**No Effect**	**Pre-test scores* (Low -.12; High +.26)** No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Self-Direction	**No Effect**	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting **Difference for Hours Attended****

/Continued Overleaf

*p<0.10; **p<0.05; ***p<0.01

Table 21 (Continued). **Effects of the Eager and Able to Learn programme on Children**

Outcomes	Main Effects	Differential Effects for Specific Subgroups[a]
Thinking and Problem-Solving		
Cognitive, Pretend Play, One-to One Correspondence	**-.29***	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Functional Pre-Academics	**-.29****	**Pretest scores marginal (Low -.17; High -.42)** No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Involvement, Concentration, Persistence, Precision		
Indirectly through Fine Motor	**No Effect**	
Indirectly through Self-Direction	**No Effect**	
Physical Movement, Gross, Fine and Sensory Motor		
Fine Motor	**No Effect**	**Pretest scores* (Low=+11; High: -.34)** No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend
Gross Motor	**No Effect**	No pretest differences No gender differences No SS/DC differences No Urban/Rural differences No differences re Fidelity, Quality of Setting or Hours Attend

*p<0.10; **p<0.05; ***p<0.01

4.3.3 The strongest evidence that the EAL is having a negative effect on the children is in relation to cognitive development

All children experienced negative effects with regard to their cognitive development. There was evidence for these negative effects from the Bayley cognitive scale as well as from the functional pre-academic scale which assesses emergent literacy for this age group. These scales measure different aspects of cognitive development. The Cognitive scale evaluates the child's development with regard to concepts that are regarded as important for cognitive growth – the ability to match patterns, to sort objects according to different dimensions, to engage in representational and imaginary plays and to understand one-to-one correspondence. On the other hand, the Functional

Pre-Academics scale includes items that are more related to emergent literacy such as pointing at pictures in books, holding crayons and pencils, imitating simple drawings, counting objects, naming colours and so on. It may well be that the shift in practice from the previous control year towards the developmental movement experiences in the EAL programme reduced the amount of time available to spend on these more 'traditional' early years activities. However, it should be noted that practitioners said that they did not 'give up' any of their previous activities to implement EAL. Nevertheless, ratings from the practitioners' questionnaires showed that they reported less frequent use 'number games' and 'books for story-telling' (see Table 19) Also, all settings scored noticeably low on the ECERs-R subscale relating to Learning Activities (see Table), indicating perhaps that, even within and around the new EAL programme, additional attention needs to be given to activities related to cognitive stimulation.

4.3.4 There is no evidence that the programme had an effect on the children's gross motor development

Surprisingly, given the amount of time spent on the twelve developmental experiences, and the high levels of reported fidelity of implementation, there were no improvements in the children's gross motor development compared to the control group. Also, this domain had one of the lowest scores at pre-test and had been identified as a domain of particular developmental concern in the baseline study, as well as in the UK and Ireland norms. The items from the Bayley gross motor measure included testing children's balance while walking up stairs (with and without support), the ability to stand on one leg, to hop, to step backwards, to catch a ball, to run with co-ordination etc and seemed well matched to features of the programme.

4.3.5 There is some evidence that the EAL programme had differential effects on subgroups of children.

The main sub-group variation was due to the developmental stage of the children at the point of entry into the programme – their pre-test scores. There were interactions on five developmental domains that showed that those who had higher pre-test scores in a certain domain reacted differently to the programme than those who had lower pre-test scores. Sometimes the high pre-scorers benefitted most from the programme and sometimes it was the other way around.

It seemed that EAL programme had distinctive polarising effects on developmental domains for high and low pre-test scorers in those domains. For example, the more developmentally advanced children appeared to benefit most from, and were able to capitalise upon, the social-emotional and play-related opportunities that the EAL programme provided but they were also most disadvantaged by the programme in relation to their cognitive development, emergent literacy, receptive language and fine motor development. The effect sizes for this group – in both directions – were largest. In

contrast, the polarising effect for the less advanced children on entering the programme was less sharp and effects sizes were all smaller. They too benefited social-emotionally and were disadvantaged cognitively and with regard to emergent literacy, but they also made small gains in receptive language and fine motor movement compared to their control group peers.

4.3.6 There is strong evidence that the programme has similar effects regardless of gender or of the type and location of the setting that the child attends

One strongly consistent pattern emerged from the analysis. The programme had similar effects regardless of gender or of type and location of the setting. Boys and girls responded similarly to the programme. Also there were no differences attributed solely to attending Sure Start and Day Care programmes, although hours spent attending a setting did impact negatively on children's ability to become independent and to regulate themselves without adult supervision. Longer hours attending was associated with Day Care settings. Children in urban and rural settings responded similarly.

4.3.7 The fidelity of implementation of the programme has very limited effects on the children's outcomes.

Nearly all settings delivered the programme with high fidelity. However, the effect of programme fidelity was restricted to language, both receptive and expressive, and in the predicted direction.

4.4 Summary of findings in relation to practitioner and parent outcomes

Table 22 gives a summary of the main effects of the EAL programme on practitioner and parent outcomes, drawn from Tables 19 and 20.

4.4.1 There is strong evidence that the EAL programme has positive effects on practitioners' and parents' beliefs, attitudes and self-reported behaviours with regard to 2-3 year old children's development

Participating in the EAL pilot programme had significantly positive effects on practitioners' and parents' beliefs, attitudes and self-reported behaviours. Although significant effects emerged for only a limited number of survey questionnaire items, the effect sizes were often large, ranging from .68 to .19, and almost entirely in the direction expected by the aims and goals of the programme. For example, with regard to providing new and different opportunities and materials for play, EAL practitioners reported that they were using more 'everyday' materials' (e.g., pots, pans, crumpled

papers) and props to help with movement games (e.g., scarves, balls, hoops). They also reported less frequent use of 'books and story-telling' and 'number games'. With regard to their interaction with children, the EAL practitioners were much less likely to adopt 'harsh and controlling' interaction styles and more likely to explain the reasons for things in order to encourage the children to think for themselves. Consistent with the new partnership arrangements with parents during the EAL year, practitioners were more positive about how their setting worked with parents, and less doubtful than they were during the control year about the contribution that parents can make to support their children's learning in the setting.

***Table 22*. Effects of the Eager and Able to Learn Programme on Practitioners and Parents**

Outcomes	Main Effects	Main Effects[1]
Practitioner and Parent Outcomes	**Practitioners**	**Parents**
Increased recognition of the importance and the different purposes of play in the development of two-year-old children;	**-.35***	**+.19*** **+.26****
and increased frequency in providing different types of play opportunities, both indoors and outdoors.	**-.38**** **+.68***** **+.48****	**+.27**** **+.19***
Increased responsiveness in practitioners'/parents' interactions and engagement with two-year-old children in order to support their communication, social, emotional, physical and cognitive development needs.	**+.57**** **+.29*** **-.47**** **+.35**** **+.44**** **-.45***	**-.22***
Increased recognition of the importance of movement for two-year-old development and how it can be related to wider developmental goals (e.g. language, cognitive, social-emotional, as well as motor development)	**+.56**** **+.51**** **+.35*** **+.34**** **+.34***	**+.40**** **+.33**** **+.40***** **+.42***** **+.40***** **+.37*****
Increased recognition of the importance of working in partnership with practitioners/ parents around the developmental needs of two-year-old children, increased opportunities to communicate with parents, and increased satisfaction with the communication.	**+.32*** **+.43****	**+.31**** **+.27**** **+.22**** **+.51***** **+.26***

[1]The effects sizes in the table refer to differences on specific survey questions related to the outcome.
*p<0.10; **p<0.05; ***p<0.01

EAL parents also reported playing with their children in different types of ways – with song and dance, and using different materials. EAL parents showed a sharper recognition than Control

parents about the relevance of play to different forms of learning both in the present and in the future. EAL parents were also substantially more satisfied than Control parents with communicating and sharing views with staff in the early years settings and also reported more help with materials and training for promoting their child's development.

5. Recommendations

5.1. Specific Recommendations about the EAL programme

The impact of the EAL programme on the children's development produced a surprising and unexpected pattern of results, with positive effects on the social emotional development and negative effects on cognitive and emergent literacy outcomes. This 'polarising' effect was more noticeable in those subgroups of children who were more developmentally advanced when they joined the programme. The absence of an effect on gross motor development was also surprising given the emphasis on the movement activities and the high fidelity implementation in this component of the programme.

The positive social emotional impact (from Bayleys) is consistent with the observed positive boost on the rated quality of the EAL settings compared to the control settings, especially on interactions between staff/child and child/child (ECERs-R), and on the practitioners' reports post-EAL that they were interacting with children in a more positive way (practitioners' survey questionnaires). Parents also appeared to learn more about the role of play in children's development and to experiment with different types of play. They were also more positive about their own interactions with the early years settings (parents' questionnaires). The findings from the Fidelity Study (Geraghty et al., 2012) show that the large majority of settings implemented the programme with high fidelity and that the programme was warmly welcomed by the vast majority of practitioners, setting managers and parents.

However, in the light of the mixed findings on child outcomes, the content of the programme needs to be re-evaluated to ensure that the positive child outcomes can be maintained and the negative impact minimised.

From the EAL evaluation, with regard to the development of the programme

- Given the unusual pattern of findings for children's outcomes, Early Years should re-evaluate the content of the EAL programme to ensure that the positive impacts on children, quality of settings, practitioners and parents are maintained and the negative impacts are minimised or turned around.

- Specifically, the dominance of the movement activities in terms of time allotted should be reassessed to create a more balanced programme that focuses directly on socio-emotional development, language, movement and conceptual development.
- The focus on high quality interactions between adults and children should be maintained and enhanced in any future programmes.
- The focus on partnerships between settings and parents should be maintained and enhanced, following the advice from the Fidelity Implementation Study on involving parents and on managing home visits.
- Fidelity monitoring should be part of any future roll-out of the programme.

5.2 General Recommendations for Policy and Research

5.2.1 The importance of a specific focus on provision for 2-3 year olds

A focus on provision for 2-3 year olds has emerged only recently as a national priority, with the launching of the 2-year old Sure Start programmes in England, Wales and Northern Ireland. Previously, both policy and research had focussed on 3-4 years in the pre-school year (e.g., the EPPE and EPPNI longitudinal research studies and the expansion of free pre-school places). The research base on what we know about the impact of provision for two years in the UK is at a very early stage. For example, the National Evaluation of Neighbourhood Nurseries, 2007, in England and the evaluation of the Early Education Pilot for Two Year Old Children, 2009, in England, both focussed on disadvantaged children. The current studies contribute substantially to the research base in Northern Ireland. From a research/policy perspective, it is important that, as well as evaluating the impact of specific programmes, participating in early years provisions (of whatever kind) is included as part of current and any future longitudinal cohort tracking (e.g., the Northern Ireland Millennium Cohort and any future cohort studies in Northern Ireland).

5.2.2 The importance of the quality of early years settings

A consistent finding across many pre-school studies is the importance of the quality of the settings for early years outcomes. This point has been confirmed again in the pilot evaluation for two-year olds in England, where positive outcomes for children were reported only for those who attended the very highest quality settings. The average rated quality of the early years settings in the current study deserves immediate attention.

5.2.2. Motor development as an approach for early years intervention

The EAL trial is one evaluation of an innovative pilot programme that focussed on developmental movement experiences as a potential approach for accelerating more general development. Although the findings from the EAL evaluation on child outcomes are surprising, it is important that research continues on the relationship between different kinds of movement development as a potential approach for early years intervention.

5.2.3 Evidence-based policy in the early years

It is important to appreciate the scale and scope of these early years studies for Northern Ireland and to understand the logistical demands of running research studies on this scale with 2-3 year old children. They need careful consideration so that policy decisions are research informed and are appropriately benchmarked with international developments.

6. References

Bayley, N. (2006a). *Bayley scale of infant and toddler development (3rd Edition). Technical Manual.* San Antonio: Harcourt Assessment, Inc.

Bayley, N. (2006b). *Bayley scale of infant and toddler development (3rd Edition). Administration Manual.* San Antonio: Harcourt Assessment, Inc.

Bayley, N. (2010). *Bayley scale of infant and toddler development (3rd Edition). UK and Ireland Supplement Manual. London: Pearson Assessment.*

Bloomfield, L. & Kendall, S. (2007). Testing a parenting programme evaluation tool as a pre- and post-course measure of parenting self-efficacy. *Journal of Advanced Nursing,* 60(5), 487-493.

Bredekamp, S. (1997). (Eds.).*Developmentally Appropriate Practice for Early Childhood Programs*. Revised Edition. Washington DC: NAEYC.

Carl, B. (2007). *The Child Caregiver Interaction Scale*. Unpublishd dissertation, Indiana University of Pennsylvania.

Cryer, D., Harms, T. & Riley, C. (2003). *All about the ECERS-R: A detailed guide in words and pictures to be used with the ECERS-R.* New York: Teachers College Press.

Cryer, D., Harms, T. & Riley, C. (2003). *All about the ITERS-R: A detailed guide in words and pictures to be used with the ITERS-R.* New York: Teachers College Press.

Geraghty, T.,Molyneaux, F. & Dunne, C. (2012). *A fidelity and implementation study of the Eager and Able to Learn programme.* Belfast: National Children's Bureau. Available at http://www.ncb.or.uk/resources

Greenspan, S. (2004). *Greenspan social-emotional growth chart: a screening questionnaire for infants and young children*. San Antonio, TX: Harcourt Assessment.

Harms, T., Clifford, R.M., Cryer, D. (2005). *Early Childhood Environment Rating Scale, Revised Edition.* New York: Teachers College Press.

Harrison, P. & Oakland, T. (2003). *Adaptive behavior assessment system 2nd edition*. San Antonio, TX: The Psychological Corporation

Kendall, S. & Bloomfield, L. (2005). Developing and validating a tool to measure parenting self-efficacy. *Journal of Advanced Nursing,* 51 (2) 174-181.

Mathers, S., Linskey, F., Seddon, J., & Sylva, K. (2007). Using quality rating scales for professional development: experiences fro the UK. *International Journal of Early Years Education*, 15(3), 261-274.

Melhuish, E., Quinn, L., Hanna, K., Sylva, K., Sammon, P., Siraj-Blatchford, I. & Taggart, B. (2004). *Effective Pre-School Provision in Northern Ireland (EPPNI) Summary Report.* Department of Education, Northern Ireland.

McGuinness, C., Connolly, P., Eakin, A. and Miller, S. (2012) *The developmental status of 2-3 year old children entering group-based settings in Northern Ireland: Survey findings*. Belfast: Centre for Effective Education, Queen's University Belfast. Available at http://www.qub.ac.uk

Molyneaux, R., Walsh, G., McConnell, B. & McGuinness, C. (2012). *The Eager and Able to Learn Pilot Year: Lessons learned – implementation and evaluation of the pilot programme*. Belfast: National Children's Bureau. Available on request from the authors

Neighbourhood Nurseries Initiative (NNI) Research Team (2007). *National Evaluation of the Neighbourhood Nurseries Initiative: Integrated Report.* Nottingham: DfES Publications.

Smith, R. et al., (2009). *Early education pilot for two year old children – evaluation*. London: Department of Children, Schools and Families, Research Report No DCSF-RR134.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2004). *Effective Provision of Pre-School Education (EPPE) project: Final report.* London: DfES.

7. Appendices

Appendix 1 Research Instrument Setting ECERS-R

Appendix 2 Research Instrument Children Bayleys III

Appendix 3 Statistical Models

Practitioner and Parent Survey Questionnaire available on request

Appendix 1

EARLY CHILDHOOD ENVIRONMENT RATING SCALE, REVISED EDITION (ECERS-R)

http://www.fpg.unc.edu/~ECERS/

The Early Childhood Environmental Rating Scale (Revised Edition) is designed to assess process quality in an early childhood centre-based setting. Process quality consists of the various interactions that go on in a setting between staff and children, staff, parents, and other adults, among the children themselves, and the interactions children have with the many materials and activities in the environment, as well as those features, such as space, schedule and materials that support these interactions. Process quality is assessed primarily through observation and has been found to be more predictive of child outcomes than structural indicators such as staff to child ratio, group size, cost of care, and even type of care, for example, child care center or family child care home (Whitebook, Howes & Phillips, 1995).

In order to provide care and education that will permit children to experience a high quality of life while helping them develop their abilities, a quality program must provide for the three basic needs all children have:

- Protection of their health and safety
- Building positive relationships
- Opportunities for stimulation and learning from experience

No one component is more or less important than the others, nor can one substitute for another. It takes all three to create quality care. Each of the three basic components of quality care manifests itself in tangible forms in the program's environment, curriculum, schedule, supervision and interaction, and can be observed. These are the key aspects of process quality that are included in the environmental rating scales.

The scales define environment in a broad sense and guides the observer to assess the arrangement of space both indoors and outdoors, the materials and activities offered to the children, the supervision and interactions (including language) that occur in the classroom, and the schedule of the day, including routines and activities. The support offered to parents and staff is also included.

The scale was developed in close collaboration with realistic field-based sites. It has good inter-rater reliability and validity, thus making it suitable for research and program evaluation, as well as program improvement efforts. ECERS-R was designed to assess group programs for children of preschool through kindergarten age, 2½ years through 5.

Overview of the Subscales and Items of the ECERS-R, 43 items into 7 subscales

Space and Furnishings

1. Indoor space

2. Furniture for routine care, play and learning

3. Furnishings for relaxation and comfort

4. Room arrangement for play

5. Space for privacy

6. Child-related display

7. Space for gross motor play

8. Gross motor equipment

Personal Care Routines

9. Greeting/departing

10. Meals/snacks

11. Nap/rest

12. Toileting/diapering

13. Health practices
14. Safety practices

Language-Reasoning

15. Books and pictures
16. Encouraging children to communicate
17. Using language to develop reasoning skills
18. Informal use of language

Activities

19. Fine motor
20. Art
21. Music/movement
22. Blocks
23. Sand/water
24. Dramatic play
25. Nature/science
26. Math/number
27. Use of TV, video, and/or computers
28. Promoting acceptance of diversity

Interaction

29. Supervision of gross motor activities
30. General supervision of children (other than gross motor)
31. Discipline
32. Staff-child interactions
33. Interactions among children

Program Structure

34. Schedule
35. Free play
36. Group time
37. Provisions for children with disabilities

Parents and Staff

38. Provisions for parents
39. Provisions for personal needs of staff
40. Provisions for professional needs of staff
41. Staff interaction and cooperation
42. Supervision and evaluation of staff
43. Opportunities for professional growth

Appendix 2

The Bayley Scale of Infant and Toddler Development III

The following descriptions of the various components of the Bayley Scales can be found (in greater detail) in the Bayley III Technical Manual.

The Cognitive Scale

The cognitive scale specifically examines the roles of play, information processing and number concepts and counting in cognitive development.

Play

Play in early childhood is believed to promote cognitive growth. Among pre-school age children, social make-believe play has been correlated with indexes of cognitive ability. Items have been developed for the Bayley III Cognitive Scale that assesses children's play skills, ranging from solitary non-relational play to social fantasy play.

Information processing

Cognitive functions in early life are correlated with cognitive functioning later in life. Information processing types of tasks (including novelty preference, habituation and paired comparisons, memory, reaction time and anticipation of patterns) have been found to correlate with both later cognitive functioning and intelligence tests. The Bayley III Cognitive Scale includes items that assess attention to novelty, habituation, memory and problem solving.

Number concepts and counting

Items in the Bayley III Cognitive Scale measure skills in one-to-one correspondence, counting and cardinality. Cardinality (the ability to assign a number accurately based on the numerosity of a set of items) is more difficult to test than earlier counting skills and the Bayley III includes further queries in the administration directions of one of the more difficult counting items to test this concept.

The Language Scale

Receptive and expressive language requires different abilities and can develop independently. As such the Bayley III assesses both.

Items on the Receptive Communication subtest focus on the child's ability to comprehend and respond appropriately to words and requests. Item difficulty is reflected in the number of words and the type of words that must be recognised. Because non-linguistic behaviours and cues can make it appear that the child understands more words than they truly know, items assess the child's comprehension in the absence of contextual cues.

The Expressive Communication subtest includes items that assess the child's ability to vocalise, babble and speak. Although there is great variation in the age at which children acquire language, the general sequence of the phases of language development is preserved. The rate of new words acquisition is influenced by genetic factors and by the child's opportunity to engage in reciprocal vocal interchange with caregivers. The Expressive Communication subtest includes items that measure the use of one word approximations, the ability to name pictures of objects and actions, the ability to communicate wants and needs, the ability to respond to questions and the ability to use multiple-word sentences. Items in the Expressive Communication subtest also measure the child's ability to combine words and gestures.

The Motor Scale

The Fine Motor and Gross Motor subtests include items that measure quality of movement, sensory integration, and perceptual-motor integration as well as basic milestones of prehension (grasping)

and locomotion. Items in the Fine Motor scale include outing coins in a slot, stacking blocks in increasingly complex ways and patterns, imitating and drawing lines and circles with a crayon and using a scissors in increasingly complex ways. The Gross Motor scale includes items such as stepping backwards, climbing stairs (there are various ways children do this depending on their developmental stage), running with coordination, balancing on each foot (with and without support), jumping forward, kicking a ball, imitating increasingly complex postures and hopping.

Social-Emotional and Development Scale

This scale is completed by a practitioner who knows the child well. It is designed to measure whether the child has reached certain social-emotional milestones for their age. While our understanding of social and emotional functioning in young children has increased over the last number of decades, challenges remain in terms of assessing and measuring the in-depth aspects of emotional functioning. Most assessments do not systematically assess the vital structure-building aspects of emotional interactions, such as the ability to relate to others, to symbolise wishes and affects (emotions) and to test reality.

It is important to distinguish between specific emotions or social skills and the acquisition of functional emotional milestones. Functional emotional milestones focus on the larger emotional patterns that define healthy emotional functioning and provide purpose to many mental processes. These milestones include the capacity to engage with a range of emotions, to experience, express and comprehend a variety of emotional signals, and to elaborate a range of feelings with words and symbols (e.g. pretend play).

The Bayley social-emotional scale is based on the functional emotional milestones identified by Stanley Greenspan. It incorporates milestones such as using symbols or ideas to convey increasingly complex intentions or feelings, dealing with increasingly complex emotional themes, forming logical bridges between emotions and ideas and forming logical bridges between their own emotional ideas and those of others.

The Adaptive Behaviour Scale

The Adaptive Behaviour Scale is designed to evaluate the attainment of functional skills necessary for the increasing independence of the child and is completed by an adult who knows the child well. As such, the scale focuses on behaviours, and measures what a child actually does in addition to measuring what the child may be able to do. Adaptive skills are divided into three clusters, which have been described as: 'the collection of conceptual, social and practical skills that have been learned by people in order to function in their everyday lives'.

The Adaptive Behaviour Scale contains 9 subscales that group into these three clusters: conceptual, social and practical skills. For the purpose of this research we are only using the subscales related to two of these groups: conceptual skills and social skills. (The practical skills cluster was omitted because it required knowledge of how children behaved at home which the practitioner would not have known about).

Social Skills

- Social behaviours including responding differently to familiar and unfamiliar people, imitating actions of adults, sharing toys willingly, saying thank you, seeking friendship with other children of the same age, saying when they feel happy, sad, scared or angry etc.
- Play related behaviours (called Leisure) including playing with a toy for more than five minutes, playing games with other people, playing simple games without adult supervision etc
-

Conceptual skills

- Communication: the ability to effectively communicate through vocalisations, gestures and the use of words and sentences

- Functional pre-academics: this includes behaviours such as imitating simple drawings, reciting nursery rhymes, counting, naming shapes, reading their own name etc.
- Self direction: this includes behaviours such as ceasing to cry or fuss when picked up, sitting quietly for at least one minute without demanding attention, showing interest in an object by pointing to it, obeying an adult's request to 'quiet down' or 'behave', following simple rules, not hitting or pushing another child when angry and upset, etc.

Table 1 . **Multilevel Linear Regression Models with the Children's Post-Test standardised Cognitive score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.473 (.043)	.501 (.060)	.454 (.042)	.474 (.043)	.461 (0.43)	.493 (.051)	.497 (.051)	.513 (.051)
Intervention$_j$	**-.688**[a] **(.363)**	-.090 (.957)	-.478 (.394)	-.657 (.442)	-.615 (.482)			
Gender$_{ij}$			.860 (.260)					
SS/DC$_{ij}$				.129 (.557)				
Urban/Rural$_{ij}$					.762 (.516)			
Quality (ECERS)$_j$						**-.162**[f] **(.156)**		
Fidelity$_j$							**.027**[g] **(.038)**	
Hours of Attendance$_j$								**-.001**[h] **(.011)**
Intervention*Pretest$_{ij}$		**-.058**[b] **(.086)**						
Intervention*Gender$_{ij}$			**-.496**[c] **(.339)**					
Intervention*SS/DC$_{ij}$				**-.092**[d] **(.757)**				
Intervention*Urban/ Rural$_{ij}$					**-.186**[e] **(.709)**			
Intervention*Quality$_j$								
Constant	5.992 (.522)	5.696 (.682)	5.800 (.520)	5.937 (.572)	5.77 (.548)	5.678 (.814)	4.125 (1.413)	4.920 (.567)
Sample Size (n)	406	406	406	406	406	231	231	229
Ω_u	1.387 (0338)	1.393 (.339)	1.398 (.336)	1.386 (.338)	1.300 (.318)	.164 (.127)	.163 (.129)	.155 (.129)
Ω_e	2.637 (.198)	2.632 (.198)	2.538 (.191)	2.636 (.198)	2.631 (.198)	2.179 (.218)	2.186 (.219)	2.176 (.219)
-2*loglikelihood	1629.508	1629.052	1616.070	1629.450	1626.078	848.769	849.363	843.267

[a]p=0.058, [b]p=0.499, [c]p=0.143, [d]p=0.903, [e]p=0.793, [f]p=0.299, [g]p=0.487, [h]p=0.907

***Table 2*. Multilevel Linear Regression Models with the Children's Post-Test standardised Receptive Communication score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.476 (.030)	.556 (.047)	.466 (.030)	.477 (.030)	.467 (.030)	.424 (.040)	.431 (.039)	.431 (.040)
Intervention$_j$	**-.151**[a] **(.237)**	1.348 (.738)	-.006 (.270)	-.088 (.285)	-.152 (.303)			
Gender$_{ij}$			.523 (.230)					
SS/DC$_{ij}$				.172 (.372)				
Urban/Rural$_{ij}$					.590 (.326)			
Quality (ECERS)$_j$						**-.113**[f] **(.179)**		
Fidelity$_j$							**.085**[g] **(.041)**	
Hours of Attendance$_j$								**-.009**[h] **(.012)**
Intervention*Pretest$_{ij}$		**-.136**[b] **(.061)**						
Intervention*Gender$_{ij}$			**-.331**[c] **(.301)**					
Intervention*SS/DC$_{ij}$				**-.197**[d] **(.503)**				
Intervention*Urban/ Rural$_{ij}$					**-.000**[e] **(.444)**			
Intervention*Quality$_j$								
Constant	6.327 (.391)	5.398 (.574)	6.199 (.393)	6.261 (.418)	6.148 (.409)	7.169 (.830)	3.741 (1.501)	6.858 (.495)
Sample Size (n)	403	403	403	403	403	229	229	227
Ω_u	.442 (.145)	.480 (154)	.418 (.140)	.440 (.145)	.352 (.129)	.314 (.178)	.220 (.151)	.312 (.178)
Ω_e	2.079 (.157)	2.037 (.154)	2.055 (.155)	2.077 (.157)	2.081 (.158)	2.186 (.220)	2.199 (.222)	2.192 (.222)
-2*loglikelihood	1488.123	1483.363	1482.074	1487.904	1481.551	850.102	846.637	843.267

[a]p=0.522, p[b]=0.027, [c]p=0.272, [d]p=0.695, [e]p=0.999, [f]p=0.530, [g]p=0.039, [h]p=0.430

Table 3. **Multilevel Linear Regression Models with the Children's Post-Test standardised Expressive Communication score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.497 (.034)	.545 (.052)	.497 (.034)	.499 (.034)	.492 (.034)	.469 (.041)	.470 (.041)	.460 (.042)
Intervention$_j$	**.157** [a] **(.279)**	1.066 (.806)	-.062 (.333)	.431 (.330)	.210 (.357)			
Gender$_{ij}$			-.137 (.299)					
SS/DC$_{ij}$				.531 (.436)				
Urban/Rural$_{ij}$					.742 (.385)			
Quality (ECERS)$_j$						**-.392**[f] **(.189)**		
Fidelity$_j$							**.100**[g] **(.047)**	
Hours of Attendance$_j$								**.007**[h] **(.014)**
Intervention*Pretest$_{ij}$		**-.083**[b] **(.069)**						
Intervention*Gender$_{ij}$			**.464**[c] **(.391)**					
Intervention*SS/DC$_{ij}$				**.865**[d] **(.585)**				
Intervention*Urban/ Rural$_{ij}$					**-.101**[e] **(.521)**			
Intervention*Quality$_j$								
Constant	5.853 (.436)	5.315 (.626)	5.916 (.446)	5.666 (.469)	5.551 (.464)	7.746 (.829)	2.838 (1.683)	6.259 (.491)
Sample Size (n)	401	401	401	401	401	229	229	227
Ω_u	.515 (.207)	.550 (.216)	.528 (.209)	.482 (.198)	.393 (.183)	.280 (.187)	.265 (.183)	.366 (.206)
Ω_e	3.489 (.266)	3.459 (.264)	3.465 (.264)	3.485 (.265)	3.491 (.266)	2.978 (.298)	2.984 (.299)	2.965 (.297)
-2*loglikelihood	1677.738	1676.311	1675.848	1675.546	1671.132	915.395	915.202	909.809

[a]p=0.574, p=0.227, [c]p=0.234, [d]p=0.139, [e]p=0.847, [f]p=0.039, [g]p=0.032, [h]p=0.613

Table 4. **Multilevel Linear Regression Models with the Children's Post-Test standardised Fine Motor score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.538 (.043)	.659 (.063)	.510 (.042)	.537 (.043)	.543 (.043)	.436 (.051)	.436 (.051)	.432 (.051)
Intervention$_j$	**-.332**[a] **(.262)**	2.203 (1.016)	-.344 (.317)	-.184 (.312)	-.256 (.355)			
Gender$_{ij}$			1.120 (.308)					
SS/DC$_{ij}$				.253 (.419)				
Urban/Rural$_{ij}$					-.141 (.388)			
Quality (ECERS)$_j$						**.019**[f] **(.204)**		
Fidelity$_j$							**.013**[g] **(.051)**	
Hours of Attendance$_j$								**.021**[h] **(.013)**
Intervention*Pretest$_{ij}$		**-.222**[b] **(.086)**						
Intervention*Gender$_{ij}$			**-.057**[c] **(.403)**					
Intervention*SS/DC$_{ij}$				**-.486**[d] **(.561)**				
Intervention*Urban/ Rural$_{ij}$					**-.155**[e] **(.518)**			
Intervention*Quality$_j$								
Constant	5.658 (.542)	4.229 (.771)	5.470 (.531)	5.587 (.560)	5.657 (.558)	6.378 (.992)	5.983 (1.851)	6.097 (.626)
Sample Size (n)	401	401	401	401	401	229	229	227
Ω_u	.322 (.178)	.384 (-.190)	.334 (.176)	.315 (.175)	.312 (.174)	.381 (.201)	.375 (.201)	.366 (.196)
Ω_e	4.042 (.307)	3.936 (.299)	3.743 (.285)	4.039 (.306)	4.040 (.306)	2.980 (.295)	2.982 (.296)	2.976 (.296)
-2*loglikelihood	1722.525	1716.024	1693.973	1721.770	1721.679	919.411	919.350	910.622

[a]p=0.205, [b]p=0.010, [c]p=0.888, [d]p=0.386, [e]p=0.765, [f]p=0.928, [g]p=0.791, [h]p=0.115

Table 5. **Multilevel Linear Regression Models with the Children's Post-Test standardised Gross Motor score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.269 (.043)	.305 (.062)	.269 (.043)	.269 (.043)	.252 (.043)	.237 (.056)	.238 (.056)	.231 (.057)
Intervention$_j$	**.022**[a] **(.398)**	.684 (.903)	.049 (.454)	.163 (.482)	.151 (.465)			
Gender$_{ij}$			.421 (.360)					
SS/DC$_{ij}$				.136 (.618)				
Urban/Rural$_{ij}$					1.658 (.504)			
Quality (ECERS)$_j$						**-.167**[f] **(.272)**		
Fidelity$_j$							**.054**[g] **(.068)**	
Hours of Attendance$_j$								**.006**[h] **(.017)**
Intervention*Pretest$_{ij}$		**-.071**[b] **(.087)**						
Intervention*Gender$_{ij}$			**-.068**[c] **(.471)**					
Intervention*SS/DC$_{ij}$				**-.436**[d] **(.839)**				
Intervention*Urban/ Rural$_{ij}$					**-.361**[e] **(.680)**			
Intervention*Quality$_j$								
Constant	7.731 (.515)	7.380 (.670)	7.534 (.542)	7.684 (.549)	7.131 (.521)	8.647 (1.145)	6.165 (2.416)	7.977 (.621)
Sample Size (n)	400	400	400	400	400	229	229	227
Ω_u	1.363 (.398)	1.346 (.394)	1.366 (.397)	1.364 (.400)	.799 (.297)	.826 (.345)	.817 (.342)	.879 (.366)
Ω_e	4.972 (.376)	4.969 (.376)	4.934 (.374)	4.967 (.376)	4.991 (.379)	4.284 (.423)	4.283 (.422)	4.293 (.426)
-2*loglikelihood	1834.451	1833.786	1831.742	1834.125	1818.982	1008.764	1008.523	1001.472

[a]p=0.957, p=0.414, [c]p=0.886, [d]p=0.604, [e]p=0.596, [f]p=0.538, [g]p=0.429, [h]p=0.718

Table 6. **Multilevel Linear Regression Models with the Children's Post-Test standardised Social-Emotional score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.443 (.056)	.341 (.080)	.430 (.056)	.435 (.057)	.443 (.056)	.538 (.077)	.538 (.077)	.536 (.081)
Intervention$_j$	**1.139**[a] **(.631)**	-.933 (1.326)	.950 (.703)	1.300 (.746)	.491 (.846)			
Gender$_{ij}$			.512 (.530)					
SS/DC$_{ij}$				-.059 (1.014)				
Urban/Rural$_{ij}$					-1.049 (1.252)			
Quality (ECERS)$_j$						**-.034**[f] **(.491)**		
Fidelity$_j$							**-.096**[g] **(.122)**	
Hours of Attendance$_j$								**.000**[h] **(.028)**
Intervention*Pretest$_{ij}$		**.196**[b] **(.111)**						
Intervention*Gender$_{ij}$			**.349**[c] **(.669)**					
Intervention*SS/DC$_{ij}$				**-.556**[d] **(1.345)**				
Intervention*Urban/ Rural$_{ij}$					**1.404**[e] **(.876)**			
Intervention*Quality$_j$								
Constant	6.938 (.761)	8.037 (.979)	6.846 (.777)	7.034 (.855)	7.422 (.876)	7.224 (2.053)	10.409 (4.317)	7.091 (.933)
Sample Size (n)	370	370	370	370	370	223	223	222
Ω_u	3.598 (1.023)	3.531 (1.005)	3.663 (1.033)	3.524 (1.013)	3.493 (.999)	3.331 (1.271)	3.132 (1.246)	3.283 (1.262)
Ω_e	8.945 (.711)	8.879 (.705)	8.788 (.698)	8.953 (.712)	8.934 (.710)	9.071 (.917)	9.097 (.923)	9.008 (.914)
-2*loglikelihood	1927.405	1924.284	1922.173	1926.956	1925.952	1160.460)	1160.333	1154.118

[a]p=0.071, p[b]=0.077, [c]p=0.602, [d]p=0.679, [e]p=0.262, [f]p=0.944, [g]p=0.432, [h]p=0.996

Table 7. **Multilevel Linear Regression Models with the Children's Post-Test standardised Communication score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.460 (.044)	.414 (.065)	.450 (.044)	.456 (.045)	.461 (.044)	.495 (.055)	.495 (.056)	.498 (.058)
Intervention$_j$	**.546**[a] **(.438)**	-.293 (.973)	.780 (.504)	.446 (.522)	.489 (.591)			
Gender$_{ij}$								
SS/DC$_{ij}$				-.515 (.702)				
Urban/Rural$_{ij}$					.238 (.645)			
Quality (ECERS)$_j$						**-.010**[f] **(.356)**		
Fidelity$_j$							**.049**[g] **(.091)**	
Hours of Attendance$_j$								**-.003**[h] **(.020)**
Intervention*Pretest$_{ij}$		**.085**[b] **(.089)**						
Intervention*Gender$_{ij}$			**-.552**[c] **(.533)**					
Intervention*SS/DC$_{ij}$				**.344**[d] **(.934)**				
Intervention*Urban/ Rural$_{ij}$					**.112**[e] **(.616)**			
Intervention*Quality$_j$								
Constant	5.521 (.542)	5.969 (.713)	5.227 (.567)	5.719 (.600)	5.403 (.616)	5.767 (1.463)	4.038 (3.210)	5.750 (.674)
Sample Size (n)	397	397	397	397	397	231	231	229
Ω_u	1.585 (.500)	1.580 (.450)	1.659 (.513)	1.551 (.493)	1.557 (.496)	1.761 (.652)	1.759 (.648)	1.765 (.658)
Ω_e	6.321 (.483)	6.307 (.482)	6.200 (.474)	6.324 (.483)	6.323 (.483)	5.179 (.513)	5.173 (.512)	5.223 (.520)
-2*loglikelihood	1912.765	1911.835	1907.262	1912.162	1912.280	1071.609	1071.322	1064.130

[a]p=0.212, [b]p=0.335, [c]p=0.300, [d]p=0.712, [e]p=0.898, [f]p=0.977, [g]p=0.591, [h]p=0.889

Table 8. **Multilevel Linear Regression Models with the Children's Post-Test standardised Functional Pre-Academic score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.370	.431	.350	.359	.374	.299	.306	.303
	(.042)	(.061)	(.042)	(.042)	(.042)	(.052)	(.052)	(.052)
Intervention$_{j}$	**-.824**[a]	.297	-.667	-.844	-1.083			
	(.407)	(.896)	(.464)	(.461)	(.531)			
Gender$_{ij}$			.953					
			(.346)					
SS/DC$_{ij}$				-1.200				
				(.615)				
Urban/Rural$_{ij}$					.463			
					(.577)			
Quality (ECERS)$_{j}$						**-.275**[f]		
						(.324)		
Fidelity$_{j}$							**-.064**[g]	
							(.081)	
Hours of Attendance$_{j}$								**.016**[h]
								(.017)
Intervention*Pretest$_{ij}$		**-.118**[b]						
		(.084)						
Intervention*Gender$_{ij}$			**-.394**[c]					
			(.446)					
Intervention*SS/DC$_{ij}$				**.143**[d]				
				(.820)				
Intervention*Urban/ Rural$_{ij}$					**.537**[e]			
					(.785)			
Intervention*Quality$_{j}$								
Constant	6.877	6.301	6.647	7.351	6.638	7.744	8.872	6.403
	(.498)	(.647)	(.512)	(.535)	(.558)	(1.385)	(2.862)	(.618)
Sample Size (n)	397	397	397	397	397	231	231	229
Ω_u	1.530	1.550	1.612	1.277	1.354	1.548	1.520	1.463
	(.416)	(.418)	(.443)	(.362)	(.385)	(.517)	(.519)	(.511)
Ω_e	4.460	4.430	4.285	4.454	4.466	3.388	3.396	3.428
	(.339)	(.336)	(.326)	(.338)	(.339)	(.334)	(.336)	(.341)
-2*loglikelihood	1785.368	1783.403	1774.332	1778.354	1781.390	979.634	979.744	972.608

[a]p=0.043, p=0.160, [c]p=0.377, [d]p=0.861, [e]p=0.494, [f]p=0.397, [g]p=0.432, [h]p=0.352

Table 9. **Multilevel Linear Regression Models with the Children's Post-Test standardised Leisure score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.287 (.055)	.169 (.081)	.282 (.055)	.279 (.056)	.290 (.055)	.388 (.071)	.388 (.072)	.393 (.073)
Intervention$_j$	**.270**[a] **(.618)**	-1.847 (1.223)	.447 (.673)	-.024 (.732)	-.228 (.832)			
Gender$_{ij}$			.869 (.465)					
SS/DC$_{ij}$				-1.444 (.952)				
Urban/Rural$_{ij}$					-.441 (.895)			
Quality (ECERS)$_j$						**.060**[f] **(.482)**		
Fidelity$_j$							**.004**[g] **(.123)**	
Hours of Attendance$_j$								**-.027**[h] **(.025)**
Intervention*Pretest$_{ij}$		**.220**[b] **(.110)**						
Intervention*Gender$_{ij}$			**-.431**[c] **(.602)**					
Intervention*SS/DC$_{ij}$				**.964**[d] **(1.279)**				
Intervention*Urban/ Rural$_{ij}$					**1.080**[e] **(1.227)**			
Intervention*Quality$_j$								
Constant	7.652 (.693)	8.779 (.891)	7.321 (.714)	8.180 (.769)	7.820 (.802)	6.715 (1.957)	6.812 (4.327)	7.400 (.896)
Sample Size (n)	397	397	397	397	397	231	231	229
Ω_u	3.906 (1.022)	3.823 (1.008)	3.892 (1.015)	3.709 (.975)	3.799 (1.007)	3.563 (1.251)	3.568 (1.256)	3.760 (1.322)
Ω_e	7.849 (.602)	7.781 (.597)	7.748 (.594)	7.840 (.601)	7.856 (.603)	7.173 (.714)	7.173 (.714)	7.140 (.715)
-2*loglikelihood	2024.298	2020.317	2019.535	2021.748	2023.488	1154.735	1154.750	1145.128

[a]p=0.662, [b]p=0.045, [c]p=0.473, [d]p=0.451, [e]p=0.379, [f]p=0.900, [g]p=0.976, [h]p=0.277

Table 10. **Multilevel Linear Regression Models with the Children's Post-Test standardised Self-Direction score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.392 (.049)	.342 (.074)	.380 (.049)	.388 (.049)	.393 (.049)	.431 (.063)	.430 (.060)	.435 (.063)
Intervention$_j$	**.554**[a] **(.659)**	-.383 (1.219)	.433 (.726)	.496 (.776)	.499 (.895)			
Gender$_{ij}$			.512 (.501)					
SS/DC$_{ij}$				-1.009 (1.010)				
Urban/Rural$_{ij}$					.050 (.963)			
Quality (ECERS)$_j$						**.203**[f] **(.530)**		
Fidelity$_j$							**.012**[g] **(.135)**	
Hours of Attendance$_j$								**-.056**[h] **(.027)**
Intervention*Pretest$_{ij}$		**.090**[b] **(.098)**						
Intervention*Gender$_{ij}$			**.217**[c] **(.648)**					
Intervention*SS/DC$_{ij}$				**.222**[d] **(1.358)**				
Intervention*Urban/ Rural$_{ij}$					**.114**[e] **(1,321)**			
Intervention*Quality$_j$								
Constant	7.369 (.703)	7.901 (.913)	7.277 (.732)	7.728 (.775)	7.340 (.833)	6.766 (2.156)	7.131 (4.771)	8.491 (.914)
Sample Size (n)	397	397	397	397	397	231	231	229
Ω_u	4.408 (1.158)	4.387 (1.155)	4.532 (1.184)	4.118 (1.110)	4.397 (1.158)	4.300 (1.521)	4.338 (1.533)	5.198 (1.844)
Ω_e	9.114 (.699)	9.100 (.698)	8.981 (.689)	9.142 (.702)	9.116 (.699)	8.525 (.849)	8.523 (.849)	8.205 (.827)
-2*loglikelihood	2082.449	2081.617	2078.350	2080.783	2082.414	1194.957	1195.096	1181.112

[a]p=0.401, [b]p=0.361, [c]p=0.737, [d]p=0.870, [e]p=0.931, [f]p=, [g]p=0.931, [h]p=0.038

Table 11 . **Multilevel Linear Regression Models with the Children's Post-Test standardised Social score as the Dependent Variable**

Independent Variables	Statistical Models (Standard Errors in Parentheses)							
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Main	Pretest	Gender	SS/DC	Urban/ Rural	Quality (ECERS)	Fidelity	Hours
Pre-test Score$_{ij}$	.458 (.049)	.501 (.075)	.444 (.049)	.441 (.050)	.463 (.049)	.409 (.062)	.408 (.062)	.414 (.063)
Intervention$_j$	**.661**[a] **(.607)**	1.365 (1.121)	.706 (.669)	.470 (.707)	.461 (.808)			
Gender$_{ij}$			.645 (.444)					
SS/DC$_{ij}$				-.1.506 (.926)				
Urban/Rural$_{ij}$					.595 (.869)			
Quality (ECERS)$_j$						**.174**[f] **(.548)**		
Fidelity$_j$							**-.093**[g] **(.138)**	
Hours of Attendance$_j$								**.004**[h] **(.024)**
Intervention*Pretest$_{ij}$		**-.074**[b] **(.099)**						
Intervention*Gender$_{ij}$			**-.125**[c] **(.573)**					
Intervention*SS/DC$_{ij}$				**.648**[d] **(1.232)**				
Intervention*Urban/ Rural$_{ij}$					**.397**[e] **(1.193)**			
Intervention*Quality$_j$								
Constant	5.783 (.638)	5.378 (.834)	5.632 (.661)	6.417 (.729)	5.479 (.748)	6.267 (2.164)	10.144 (4.870)	6.780 (.833)
Sample Size (n)	397	397	397	397	397	231	231	229
Ω_u	3.857 (.982)	3.835 (.977)	3.964 (1.003)	3.482 (.908)	3.640 (.949)	5.012 (1.609)	4.887 (1.591)	4.982 (1.631)
Ω_e	7.101 (.544)	7.095 (.544)	6.998 (.537)	7.118 (.545)	7.112 (.546)	6.282 (.625)	6.290 (.626)	6.284 (.629)
-2*loglikelihood	1988.188	1987.630	1984.127	1984.763	1986.296	1134.887	1134.543	1125.226

[a]p=0.276, p[b]=0.455, [c]p=0.827, [d]p=0.599, [e]p=0.739, [f]p=0.750, [g]p=0.501, [h]p=0.884