preparing for life

Early Childhood Intervention

Assessing the Impact of *Preparing For Life* at 48 Months

By the *PFL* Evaluation Team, UCD Geary Institute for Public Policy

UCD Geary Institute for Public Policy

preparing for life
planning together for our children

Preparing For Life:
Early Childhood Intervention

Assessing the Impact of *Preparing For Life* at Forty-Eight Months

EVALUATION OF THE *'Preparing For Life'*
EARLY CHILDHOOD INTERVENTION PROGRAMME

By
PFL EVALUATION TEAM AT THE UCD GEARY INSTITUTE FOR PUBLIC POLICY
November, 2015

UCD Geary Institute for Public Policy

preparing for life
planning together for our children

Table of Contents

List of Tables

List of Figures

Acknowledgements

The *Preparing For Life* Team and the UCD Geary Institute Evaluation Team would like to thank all those who participated in and supported this research, especially the families participating in the *PFL* programme and evaluation.

Preparing For Life would particularly like to thank all members of UCD Geary Institute Evaluation Team over the last eight years (Ailbhe Booth, Carly Cheevers, Maria Cosgrove, Deirdre Coy , Sarah Finnegan, Nick Fitzpatrick, Olivia Joyner, Judy Lovett, Louise McEntee, Edel McGlanaghy, Kelly McNamara, Eylin Palamaro Munsell, Christine O'Farrelly, Catherine O'Melia, Caroline Rawdon, Sarah Thompson, Gerard Victory, and all of the interns who have worked on the evaluation to date), under the direction of Dr. Orla Doyle and the scientific advisory committee (Professor Colm Harmon, Professor James Heckman, Professor Cecily Kelleher, Professor Sharon Ramey, Professor Sylvana Cote, and Professor Richard Tremblay) for their work in bringing this report to life. The *PFL* Evaluation Team also would like to thank Caitriona Logue and Seong Moon, for contributing to the report and Mark Hargaden for IT support. We also would like to thank all the local organisations, including Sphere 17, the Darndale Belcamp Village Centre, Project West, Finglas and St. Helena's Resource Centre, Finglas, who offered work space in which to conduct interviews with participants.

We thank our funders The Atlantic Philanthropies and the Department of Children and Youth Affairs and acknowledge the advice and guidance given by staff of both organisations. We also thank our Expert Advisory Committee for their support and guidance.

We are most grateful to the Northside Partnership Board, CEO and staff team as well as the *PFL* Board Sub-Committee, *PFL* Planning, Steering and Working Groups for their on-going encouragement and support. We appreciate the generosity of all the agencies, organisations and individuals who have given so willingly of their time and expertise in supporting *PFL* throughout the planning and implementation phases.

Finally, we thank the members of the *PFL* staff team over the past eight years (Val Smith, Sarah Jane Leonard, Melanie Murphy, Susan Cullen, Eva Rigo, Victoria Monkhouse, Sandra O'Neill, Cindy Lawson, Ann Loughney, Gemma Cooper, and Bianca Toeneboen) who have brought the *PFL* programme to life.

Noel Kelly,

Manager, *Preparing For Life*

Executive Summary

Preparing for Life (PFL) is a prevention and early intervention programme which aims to improve the life outcomes of children and families living in North Dublin, Ireland, by intervening during pregnancy and working with families until the children start school. The PFL programme is being evaluated using a mixed methods approach, incorporating a longitudinal randomised control trial design and an implementation analysis. The experimental component involves the random allocation of participants from the PFL communities to either a high support treatment group or a low support treatment group. Both groups receive developmental toys, as well as access to preschool, public health workshops, and a support worker. Participants in the high treatment group also receive home visits from a trained mentor and have group parent training via the Triple P *Positive Parenting Programme. The PFL treatment groups are also being compared to a 'services as usual' comparison group (LFP), who do not receive the supports of the PFL programme. This is a summary of the findings of the evaluation when the PFL children were approximately forty-eight months of age and were preparing to leave the programme.*

Recruitment and Baseline Characteristics

In total, 233 pregnant women were recruited into *PFL* between January 2008 and August 2010. Randomisation resulted in 115 participants assigned to the high treatment group and 118 participants assigned to the low treatment group. In addition, 99 pregnant women were recruited into the comparison group. The population-based recruitment rate was 52%. Baseline data, collected before the programme began, was available for 104 high and 101 low *PFL* treatment group participants respectively, and 99 comparison group participants. Tests of baseline differences between the high and low *PFL* treatment groups found that the groups did not statistically differ on 97% of the measures analysed, indicating that the randomisation process was successful. The aggregate *PFL* group and the LFP comparison group did not statistically differ on 75% of the measures; however, the comparison group was of a relatively higher socioeconomic status.

Previous Findings to Date

The findings from six to thirty-six months are discussed below and presented in Table ES.1.

CHILD DEVELOPMENT

The programme's impact on child development increased gradually over time, from no significant differences between the high and low treatment groups at six months to differences on a third of all measures at thirty-six months. The number of significant differences grew from 7% at twelve months, including effects for fine motor skills and overall social competence, to 16% at eighteen months, which saw gains for the children's gross motor skills and personal-social development. Eighteen months also saw the emergence of an effect for the children's cognitive development, a result which persisted to thirty-six months. Twenty-four months saw the largest number of significant findings in the child development domain. There were significant differences on 34% of the tests, including new domains such as problem solving skills and child behaviour. By thirty-six months, the proportion and areas of significance remained relatively consistent, with identified treatment effects for problem solving, behavioural problems, and cognitive development.

CHILD HEALTH

The programme has consistently had an impact on child health, from a low of 10% at six months to a peak of 47% at twenty-four months. High treatment children displayed significant differences in nutrition and short-term health outcomes across each period compared to the low treatment group. At six months, they were more likely to eat at the appropriate frequency and have their necessary 4 month immunizations. By twelve months, 17% of the tests were significant, and the immunization and dietary results were maintained, while the high treatment group were also less likely to develop chest infections. By eighteen months, there were differences between the high and low treatment group on almost one-quarter of all measures, including overnight hospital stays and diet. At twenty-four months, significant differences were found on almost half of all the child health measures and many of the previously significant findings were either maintained or regained, such as diet and chest infections, while new effects were identified for asthma. By thirty-six months, there was a drop in the number of significant health findings to 24%, yet the results for asthma, dietary intake, and hospitalisations remained.

PARENTING

The number and type of significant findings for the parenting measures varied over each period. At six months, there were significant differences between high and low treatment groups on 23% of the parenting measures. These findings pertained to dysfunctional, and overall, interactions with the child, the baby comparison score, parental hostile reactive behaviour, and parental stress. At twelve months, there were no significant findings in the hypothesised direction across any of the parenting measures. The eighteen month evaluation saw some improvement, with 20% of the measures being significant, including interactions with the child. The twenty-four month results were similar to the eighteen month results with significant differences between the high and low treatment groups on 18% of all measures, including the baby comparison score and parental stress. The high treatment group also displayed a significant difference in parental self-efficacy. By thirty-six months, the highest proportion of significant findings were found (26%), including previously identified effects for parental self-efficacy and the baby comparison score, as well as effects on new measures of parenting styles, maternal education attitudes, and TV habits.

HOME ENVIRONMENT

Findings in relation to the home environment changed from wave to wave, partially due to the use of different instruments, as well as different programme effects at each wave. For example, children in the high treatment group were significantly less likely to be exposed to cigarette smoke at eighteen and thirty-six months, yet this finding was not replicated at six or twenty-four months. At six months, there were significant differences between the high and low treatment groups on 36% of the home environment measures, including improvements in the quality of the home environment concerning variety in the child's life, the physical environment, and learning materials. However, at twelve months there were no significant findings using different measures of the home environment. By eighteen months, one-third of all measures were significant, including measures which were significant at six months such as the physical environment and learning materials, and the use of electrical socket covers. Newly significant findings also included acceptance of the child by the parent and fewer restrictions in the child's life. At twenty-four months, only one of the two measures assessed was significant: whether a social worker was working with the family. This finding was not significant at previous time points. By thirty-six months, 40% of the measures were significant, including measures of organisation, involvement and acceptance. Thus, the majority of the effects in this domain were driven by improvements in the quality of the home environment as measured by the HOME instrument.

MATERNAL HEALTH AND WELLBEING

There were relatively few significant differences between the high and low treatment groups on maternal health and wellbeing until the thirty-six month period. At six months, only one of 20 measures (5%) was significant: whether the mother was hospitalised immediately after the birth for medical care. Similarly at twelve months, only one of 28 measures (4%) was significant concerning maternal alcohol consumption. By eighteen months, only one result was significant in the hypothesised direction (5%), and there were multiple effects in the non-hypothesised direction. For example, women in the high treatment group were less likely to binge drink, but engaged more with health services and hospitals, and were more likely to be pregnant. Again, at twenty-four months, only one effect was significant in the hypothesised direction, and one effect was significant in the non-hypothesised direction. Mothers in the treatment group were more likely to have a planned pregnancy if they were pregnant. However, contrary to our hypothesis, they visited the GP more frequently. Thirty-six months saw the greatest proportion of significant results in the maternal health and wellbeing domain with almost one-quarter of the results reaching significance. High treatment mothers displayed less postnatal depression, a higher level of subjective wellbeing, and had reduced the number of cigarettes they smoked. They were also significantly less likely to have smoked in the previous 12 months.

SOCIAL SUPPORT

The findings in relation to social support decreased non-linearly over time, from significant differences between the high and low treatment group on 38% of measures at six months to 5% at thirty-six months. The majority of significant findings concerned interactions with family and friends. At six months, high treatment mothers differed from low treatment mothers on over one-third of the social support measures, including programme discussion, and the frequency of parent and friend encounters. By twelve months, 43% of the social support measures bore significant differences in areas concerning social support from friends. At eighteen and twenty-four months, the number of significant findings had decreased to 8% and 10% respectively. However, the areas of significance were maintained concerning support from parents and relatives. Thirty-six months saw the lowest number of social support effects, with only one of the 19 measures showing significance, and all previous findings concerning family and friend support were no longer significant.

CHILDCARE

There have been almost no significant programme effects in relation to childcare at any time point.

HOUSEHOLD FACTORS AND SES

There has been a small but gradual increase in the number of significant findings related to household factors and socioeconomic status (SES) – moving from no significant differences at six months to 14% by thirty-six months. Over time, maternal employment in particular saw consistent significant differences between the groups. At twelve months, significantly fewer high treatment participants were long-term unemployed. At eighteen months, high treatment participants were more likely to be employed. Significant differences were maintained at twenty-four months concerning maternal unemployment and the proportion of mothers looking after the family. Some new measures also reached significance, including measures concerning maternal education. By thirty-six months, 14% of the findings were significant – once again significance was mostly found in measures concerning maternal employment status. However, the high treatment group also saw significant improvements in their financial situations and with respect to housing.

Forty-Eight Month Report

The aim of this report is to test whether the *PFL* programme had an impact on parent and child outcomes at forty-eight months, and to provide a detailed review of implementation practices regarding participant attrition and engagement.

Impact of *PFL* at Forty-Eight Months: Main Results

In total, 217 forty-eight month interviews (n*High* = 74; n*Low* = 73; n*LFP* = 70) were completed. The main analyses compared the outcomes of the high treatment group to the outcomes of the low treatment group across eight domains: child development, child health, parenting, home environment, maternal health and wellbeing, social support, childcare, and household factors and (SES), incorporating 191 outcome measures.

Table ES.1 summarises the *PFL* results at six, twelve, eighteen, twenty-four, thirty-six, and forty-eight months. Based on the literature, we anticipated moderate positive effects in the domains of child development, parenting, maternal health and wellbeing, and household factors and SES at forty-eight months. As reported in other studies, we expected to find limited programme effects in the realms of child health, the home environment, social support, and childcare. Consistent with the evaluation design, we anticipated few significant differences between the low treatment group and the comparison group.

At forty-eight months we saw a drop in the number of positive findings compared to the previous time point at thirty-six months. In particular, we found limited positive effects in the domains of child development, parenting, maternal health and wellbeing, and household factors and SES. The findings in the other domains were also relatively limited, as hypothesised. Positive significant differences between the high and low treatment groups were observed on 12% of the individual measures, and 4 of the 32 step-down categories (13%) remained significant in the multiple hypothesis analyses. In order to account for potential bias which differential attrition may introduce, the main outcome analyses were re-estimated using an Inverse Probability Weighting (IPW) technique. When IPW was applied, the number of individual significant findings increased from 12% to 18%, while the number of significant step-down findings increased from 13% to 25%. This suggests that the mothers from the high or low treatment groups who did not participate in the forty-eight month interview had differential characteristics, leading to an under-estimation of the main treatment effects. Thus, when IPW was applied, the results were more consistent with previous time points.

Individual factors found to be significant in the non-IPW analysis include the following. Children in the high treatment group exhibited stronger fine motor skills and cognitive development than the low treatment group, in addition to a reduced incidence of clinically significant levels of externalising and internalising problems. Furthermore, they were less likely to have asthma. High treatment children were more likely to consume the recommended daily amount of vegetables, and were less likely to be overweight according to their BMI. They reportedly slept for longer each day, had fewer sleep problems and were more likely to be fully toilet trained than low treatment children. Regarding parenting, mothers in the high treatment group reported fewer permissive parenting behaviours than those in the low treatment group, and also reported that their children spent less time watching TV alone. High treatment children were less likely to be exposed to cigarette smoke at home, and their families were less likely to be working with a social worker. High treatment mothers were more likely to report being in good health, they consumed less alcohol, and were less likely to binge drink. In addition, they were more likely to report having voted in the last local, European and general elections than low treatment mothers. There were no differences between the high and low treatment groups on any of the childcare measures. In terms of household factors and SES, high treatment mothers reported fewer mental health issues in their family.

Table ES.1 - Summary of Main Findings at Six, Twelve, Eighteen, Twenty-Four, Thirty-Six, & Forty-Eight Months

PFL Low – *PFL* High	Proportion of Measures Significantly Different											
	Six Months		Twelve Months		Eighteen Months		Twenty-Four Months		Thirty-Six Months		Forty-Eight Months	
	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests
Child Development	0% (13)	0% (2)	7% (28)	20% (5)	16% (25)	0% (6)	34% (41)	22% (9)	33% (39)	38% (8)	19% (32)	17% (6)
Child Health	10% (30)	0% (3)	17% (23)	0% (4)	24% (17)	67% (3)	47% (17)	50% (2)	24% (21)	33% (3)	17% (35)	0% (5)
Parenting	23% (22)	20% (5)	0% (16)	0% (2)	20% (10)	50% (2)	18% (17)	0% (3)	26% (34)	43% (7)	6% (36)	0% (8)
Home Environment	36% (22)	50% (2)	0% (6)	0% (1)	33% (21)	67% (3)	50% (2)	~	40% (15)	50% (2)	50% (4)	~
Maternal Health & Wellbeing	5% (20)	25% (4)	4% (28)	25% (4)	5% (19)	0% (3)	6% (16)	0% (3)	24% (17)	33% (3)	14% (21)	25% (4)
Social Support	38% (13)	0% (2)	43% (7)	0% (2)	8% (12)	0% (3)	10% (19)	0% (4)	5% (19)	0% (4)	14% (14)	33% (3)
Childcare	7% (14)	0% (2)	~	~	0% (16)	0% (2)	0% (7)	0% (1)	0% (17)	0% (2)	0% (8)	0% (1)
Household Factors & SES	0% (26)	0% (5)	3% (32)	0% (5)	8% (23)	0% (5)	13% (47)	29% (7)	14% (42)	0% (6)	5% (41)	20% (5)
Total Statistically Different	14% (23/160)	12% (3/25)	8% (11/140)	9% (2/23)	14% (21/152)	19% (5/27)	21% (34/166)	17% (5/29)	22% (44/204)	26% (9/35)	12% (23/191)	13% (4/32)

The findings at forty-eight months represent a decrease from those found at thirty-six months. Explanations for this decrease are further explored in the report. It is important to note however, that the majority of the effects identified at forty-eight months were also found at previous time points, which indicates some level of consistency, particularly in the area of child development which is one of the primary outcomes of the programme.

PFL Implementation Analysis at Forty-Eight Months

ATTRITION

The level of official attrition from *PFL* between baseline and forty-eight months was 16% across the whole sample. Official attrition was slightly higher among the high treatment group (19%) than among the low treatment group (17%), and was lowest among the comparison group (12%). In addition to those who officially dropped out, 18% of the sample did not complete a forty-eight month interview, either because the interview could not be scheduled at a suitable time, or because the participants disengaged from the study. The non-completion rates across the high and low treatment groups were 17% and 21% respectively, while the corresponding rate for the comparison group was 17%. Total non-completion (attrition and disengagement) at forty-eight months was very similar among the high (36%) and low (38%) treatment groups, and slightly lower among the comparison group (29%).

In order to test for non-random attrition and disengagement, we compared the baseline characteristics of those who participated in the forty-eight month survey to those who did not. Overall, there was some evidence of systematic differences between these groups. In general, more disadvantaged participants did not participate in the forty-eight month interview. In order to account for any potential bias which differential attrition may introduce, the main outcome analyses were re-estimated using an Inverse Probability Weighting (IPW) technique. There were more significant differences between the high and low treatment groups when the weighting was applied, particularly in the domains of child development (from 19% to 56%), child health (from 17% to 23%) and maternal health and wellbeing (from 14% to 19%). There was also a very slight increase in the number of significant findings in the domain of parenting (from 6% to 8%). However, the weighting led to a reduction in the number of significant results in the home environment (from 50% to 25%), social support (from 14% to 0%), and household factors and SES (from 5% to 2%) domains. There was no change to the childcare findings when weighting was applied. In the multiple hypothesis tests, the weighting led to an increase in the areas of child development (17% to 67%), child health (0% to 40%), and maternal health and wellbeing (25% to 50%), a decrease in social support (33% to 0%), and household factors and SES (20% to 0%), while there was no change in parenting or childcare. In sum, the IPW procedure had most impact on the primary outcomes of the programme, and child development, in particular.

ENGAGEMENT

Families in the high treatment group received an average of 54 home visits from the *PFL* mentors between programme intake and forty-eight months, with each visit lasting slightly under one hour on average. Having been relatively consistent for the first two years of the programme, and increasing between twenty-four and thirty-six months, the average number of home visits between thirty-six and forty-eight months fell significantly. This can be partly attributed to participant fatigue during the last year of the programme, as well as reports from mentors on reducing the amount of contact time with families in the final year of the programme to ensure a successful transition to programme exit.

Relatively few participant characteristics were associated with engagement. Mothers who smoked during pregnancy had fewer home visits and spent less total time in visits, and mothers with higher cognitive resources experienced more home visits and had a longer total duration of time spent in visits. Overall, we found little evidence that the factors which are often identified as determinants of engagement in the literature are present in this sample. For example, factors such as age, marital status, socioeconomic factors, parenting, and socio-emotional functioning were not associated with engagement in *PFL*. These findings are consistent with the analysis of engagement reported in the previous reports, with very few individual characteristics associated with engagement at any time point.

MISREPORTING

It is possible that the participants chose to answer the interview questions in a way that they felt was socially acceptable, or favourable to the researcher. The possibility of misreporting was tested in the forty-eight month interview using a "bogus question", in this case, asking participants if they had ever heard of the phrase *'Ladder of Learning'*. The high and low treatment group participants were approximately equally likely to report having heard of this phrase. This suggests that the results are unlikely to be biased by differential misreporting among the high and low treatment groups.

CONTAMINATION

A contamination analysis was conducted to determine whether the low treatment group received all or part of the services designed for the high treatment group. The indirect measures of contamination indicated that the potential for contamination was high as participants in both the high and low treatment groups reported knowing multiple neighbours in the *PFL* programme with children of similar ages to their own. In addition, the proportion of parents reporting sharing their *PFL* materials was higher in the high treatment group than in the low treatment group. This result was as expected as the high treatment group receive more materials from the programme. The results of the direct measure of contamination indicated that the high treatment group had a greater knowledge of the phrase, the *'Feeling Wheel'*, than the low treatment group and the comparison group. The mentors describe the feeling wheel to participants in the high treatment group as part of the programme and there is a Tip Sheet describing the meaning of this phrase. This suggests that contamination may not have occurred between the high and low treatment groups. The fact that a similar proportion of the low treatment group and the comparison group reported knowledge of the phrase is further evidence of the absence of major contamination.

Additional Forty-Eight Month Results

Additional analyses were conducted to explore different aspects of the data not captured in the main analysis. These included a comparison of the forty-eight month outcomes of the low treatment group to the no-treatment comparison group, and a dynamic analysis which examined changes in child and parent outcomes over time. Interaction and subgroup analyses were also carried out to investigate whether the programme was more effective for specific groups of participants.

The findings comparing the low treatment group to the comparison group support the study design as it suggested that the low treatment group was not performing systematically better than the comparison group across most domains. The low treatment group outperformed the comparison group on 7% of measures, while the comparison group outperformed the low treatment group on 8% of measures, with the majority of the differences being clustered around household factors and SES outcomes. The lack of a coherent pattern within the findings suggests that the low treatment group did not receive the supports designed for the high treatment group, and that the low treatment supports had little impact on their outcomes.

The results of the dynamic analysis, which examined changes in the impact of the programme over time, found that only two (4%) of the 50 measures upon which dynamic analyses were conducted yielded significant findings, suggesting little evidence of dynamic effects over time.

In order to investigate differential effects of the *PFL* programme for specific types of participants, interaction and subgroup analyses were conducted for categories commonly found to be of relevance to early childhood and family interventions - child gender, maternal parity status, and maternal cognitive resources. These analyses focused on the child development outcomes. Overall, the interaction and subgroup analyses indicated that the *PFL* programme impacted certain participants differently depending on their characteristics. Regarding child development, the programme was most beneficial for the children of first time mothers and the children of mothers with lower cognitive resources. The programme did not affect boys and girls differently. A further analysis of child development and parenting outcomes was conducted to compare the outcomes of participants who took part in the Triple P programme and those that did not. While Triple P participants performed slightly better than non-Triple P participants, the differences between the two subgroups were relatively small.

Conclusion

The forty-eight month evaluation of *Preparing For Life* found a reduction in positive programme effects since the thirty-six months evaluation. Twelve pecent of the individual tests and 13% of the step-down tests were statistically different across the high and low treatment groups, representing a decrease in significant findings. Counter to our hypothesis, findings were limited in all eight domains. This reduced programme impact may be due to the reduction in mentor contact time between thirty-six and forty-eight months. However, the strength of the IPW results must be taken into account when interpreting these findings. The weighted findings were more consistent with our hypotheses and, particularly in the areas of child development, were similar to findings at previous time points. This suggests that failing to account for differential attrition may lead to some under-estimation of the impact of the programme. The subgroup analysis also revealed important programme effects for different types of families depending on the domain. These results may be informative with respect to future programme planning and rollout. The final *PFL* report will include an overview of the 0-48 month results, in addition to the final results based on children's school readiness skills on school entry.

The reports of the six, twelve, eighteen, twenty-four, thirty-six, and forty-eight month *PFL* evaluations can be found at the following website under publications: http://geary.ucd.ie/preparingforlife

The Life of Kirsty, an Average PFL Child at Forty-Eight Months

Kirsty is now four and is soon to leave the *PFL* programme and start school. She lives with her Mam and Dad who are unmarried, her big brother and her granny. Kirsty and her Mam still see their mentor, but they see her less often than before, usually once every six weeks. Kirsty's Mam is in good health and does not drink too much. However, life has its difficulties: her Mam sometimes feels very down, although she does not have any diagnosed mental health issues, she has a medical card, and Kirsty's Dad is unemployed.

At the moment, Kirsty spends a good part of the week in formal childcare which is helping her prepare for the important step of starting school. Her Mam feels that she has all the mental skills needed to move into a school setting, and she also has good fine motor skills which will help her with day-to-day classroom tasks like handwriting. She is fully toilet trained which is important for school. Kirsty is typically in good form and does not get depressed, anxious or act out in a way that makes her Mam concerned. Her parents always set rules for Kirsty, for example, during the day Kirsty likes to watch TV but her Mam will always watch it with her. Unlike some of her friends, she does not have asthma and nobody in her house smokes around her. She eats all her vegetables at dinner time, is the right weight for her age group, and sleeps well. Although Kirsty and her Mam will miss their *PFL* mentor, they are looking forward to Kirsty starting school and believe she has all the skills needed to help her embark on this new phase of her life.

Chapter One

Background of the *PFL* Programme Forty-Eight Month Evaluation

1.1 Introduction

This report is the sixth and penultimate in a series of reports which present the results of the *PFL* evaluation. The report '*Preparing For Life* Early Childhood Intervention: Assessing the early impact of *Preparing For Life* at Six Months' contains relevant background information about the programme and serves as the foundation for this report[1]. This report included a detailed description of the *PFL* intervention and evaluation, the *PFL* logic model, and an explanation of the theoretical underpinnings of home visiting interventions. The six, twelve, eighteen, twenty-four, and thirty-six month reports included a discussion of the outcomes at those time points for other home visiting interventions, in addition to the results of the *PFL* impact and implementation evaluation at those time points. The present report focuses on information specific to the forty-eight month evaluation, including new measures utilised as part of the forty-eight month interview, the results of the impact evaluation at forty-eight months, and new implementation data collected between thirty-six and forty-eight months. This report also includes interaction and subgroup analyses to determine whether the programme had differential effects for different groups of participants.

Chapter 1 provides a brief summary of the recruitment process, the analysis of baseline data, and the results of the evaluation at six, twelve, eighteen, twenty-four, and thirty-six months. It then presents a table of relevant findings from the literature on the impact of home visiting programmes at forty-eight months of age. Updated hypothesis are presented, as well as information regarding the collection of the forty-eight month interview data. A description of the remainder of the report concludes this chapter.

1.2 Recruitment & Baseline Analysis

In total, 233 pregnant women were recruited into the *PFL* programme between January 2008 and August 2010. Randomisation resulted in 115 participants assigned to the high treatment group and 118 participants assigned to the low treatment group. In addition, 99 pregnant women were recruited into the comparison group. The population-based recruitment rate was 52%. Baseline data, collected before the programme began, were available for 104 and 101 high and low *PFL* treatment group participants respectively, and for 99 comparison group participants. Tests of baseline differences between the high and low *PFL* treatment groups found that the groups did not statistically differ on 97% of the measures analysed, indicating that the randomisation process was successful. The aggregate *PFL* group and the LFP comparison group did not statistically differ on 75% of the measures; however, the comparison group was of a relatively higher socioeconomic status.

Full details of the recruitment methods and baseline analysis are available in Chapter 2 of '*Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at Six Months'.

1.3 Summary of the Six Month Evaluation

The six month evaluation suggested that the programme made a promising initial impact. In total, 257 six month interviews (n_{Low} = 90; n_{High} = 83; n_{LFP} = 84) were completed, representing 77% of the original sample recruited. Analysis across eight domains revealed there were limited significant differences between the high and low treatment groups (14%). In addition, three of the 25 step-down categories (12%) remained significant in the multiple hypothesis analysis. These results were consistent with the programme evaluation literature which finds few treatment effects at this early stage. Many of the relationships were in the hypothesised direction, with the high treatment group reporting somewhat better outcomes than the low treatment group. There were significant findings in the domains of parenting, the quality of the home environment, and social support, which correspond directly to information on the *PFL* Tip Sheets delivered to participants during this period. Participants received 14 home visits between baseline and six months on average, thus the intervention may not have been sufficiently intensive to generate significant treatment effects at this early stage. However, there was high participant satisfaction with the programme, indicating that engagement was high which may impact on future outcomes.

[1] This report can be found at the following website under publications: http://geary.ucd.ie/preparingforlife.

1.4 Summary of the Twelve Month Evaluation

The twelve month evaluation indicated that the programme was progressing well regarding the retention of participants and programme satisfaction. In total, 247 twelve month interviews (n*Low* = 83; n*High* = 82; n*LFP* = 82) were completed, representing 74% of the original sample recruited. Of the variables measured, 8% were statistically significant in the hypothesised direction and 6% were significant in a non-hypothesised direction. Two of the 23 step-down categories remained significant in the multiple hypothesis analysis, including child development and maternal health and wellbeing. These limited results were in line with evaluations of other home visiting programmes, which typically identify few effects at this time period. Although there were fewer significant differences reported at twelve months than at six months, measures which focus on different aspects of the domains of interest were utilised at each time point. Therefore, it is not possible to make a direct comparison between findings from the two reports on some domains, most notably parenting and the home environment. On average, participants received 7 home visits between six and twelve months, which was lower than anticipated yet equivalent to the number of visits delivered during the first six months. This suggests that while engagement among participants is lower than anticipated, they were satisfied with the level of support they received and they are chose to remain in the programme.

1.5 Summary of the Eighteen Month Evaluation

The eighteen month findings were consistent with similarly timed evaluations of other home visiting programmes, which typically identify limited significant effects at this time point. In total, 225 eighteen month interviews (n*Low* = 80; n*High* = 74; n*LFP* = 71) were completed, representing 68% of the original sample recruited. A similar number of significant differences were found as at six months, both of which were higher than the number reported at twelve months. This is most likely due to the use of similar measures at both time points (e.g. the HOME). In total, 14% of the outcomes analysed were significant in the hypothesised direction, while 6% were statistically significant in the non-hypothesised direction. However, the majority of the relationships were in the hypothesised direction, with the high treatment group reporting better outcomes than the low treatment group. Significant effects were found for gross motor skills, personal-social competence and cognitive development, despite evaluations of other home visiting programmes failing to identify significant findings in these areas at eighteen months. Of the 27 step-down categories tested, five (18.5%) were statistically significant. The participants in the high treatment group received an average of 27 visits between recruitment and when the infant was eighteen months old. The number and timing of home visits indicated that the average number of home visits was broadly similar over time.

1.6 Summary of the Twenty-Four Month Evaluation

The twenty-four month findings indicated that the programme was continuing to progress well and to have a positive impact on many domains. A total of 239 (n*Low* = 84; n*High* = 82; n*LFP* = 73) interviews were completed, representing 72% of the original sample recruited. Of the 166 outcome measures analysed, 59% were in the hypothesised direction and 21% were statistically significant. In addition, five of the 29 step-down categories (17%) remained significant in the multiple hypothesis analysis, including aspects of child development, child health, and household factors and SES. In terms of effect sizes, moderate effect sizes of between 0.20 and 0.40 were identified on the majority of significant results. At this time point, five (3%) of the variables were significantly different in the non-hypothesised direction. Families in the high treatment group received an average of 33 home visits between programme intake and twenty-four months, with each visit lasting approximately one hour. On average, participants met their mentor just under once a month between eighteen and twenty-four months.

1.7 Summary of the Thirty-Six Month Evaluation

The thirty-six month findings were consistent with, and slightly exceeded, those found at twenty-four months and represented the largest proportion of significant individual tests and step-down tests found to date. In total, 217 interviews (n*Low* = 76; n*High* = 75; n*LFP* = 66) were completed, representing 65% of the original sample recruited. Of the 204 outcome measures tested, 22% of the individual tests were statistically significant. In addition, 9 of the 35 step-down categories (26%) remained significant in the multiple hypothesis analysis, including aspects of child development, child health, parenting, the home environment, and maternal health and wellbeing. Small to medium effect sizes of between 0.20 and 0.40 were identified for the majority of significant results. Six (3%) of the measures were significantly different in the non-hypothesised direction. Families in the high treatment group received an average of 46 home visits between programme intake and thirty-six months, with each visit lasting slightly over one hour. In addition, 62% of the intervention participants who completed the thirty-six month assessment took part in some form of Triple P, with the majority availing of Group Triple P which consists of 5 two-hour group discussion sessions and 3 individual phone calls facilitated by the mentors.

1.8 Evidence on the Effectiveness of Home Visiting Programmes

Previous reports reviewed the evidence on the effectiveness of home visiting programmes on outcomes observed up to thirty-six months of age. This section reviews the literature on outcomes reported between thirty-six and forty-eight months. Several evaluations of home visiting interventions assess outcomes at this time point and favourable results are observed predominantly in the domain of child development and school readiness; however the results are not consistent across programmes.

Table 1.1 reflects the outcomes from home visiting programmes from thirty-six months to four-five years. The primary source of information for the table was the Home Visiting Evidence of Effectiveness (HomVEE) website (http://homvee.acf.hhs.gov/). This site was launched by the U.S. Department of Health and Human Services to conduct a thorough and transparent review of the home visiting research literature and provide an assessment of the evidence of effectiveness for home visiting programme models that target families with pregnant women and children from birth to age five. Trained reviewers evaluated randomised controlled trials and quasi-experimental designs for each model and authors were given the opportunity to respond to missing information.

The table contains results from studies relating to the domains which are closely linked to those investigated by the *PFL* programme: child development and school readiness, child health, positive parenting, reductions in child maltreatment, the home environment, maternal health and wellbeing, social support, household factors and SES, and reductions in juvenile delinquency, family violence, and crime. All were rated by HomVee as either:

(1) High quality: random assignment studies with low attrition of sample members and no reassignment of sample members after the original random assignment, and single case and regression discontinuity designs that meet the What Works Clearinghouse (WWC) design standards, or
(2) Moderate quality: random assignment studies that due to flaws in the study design or analysis (e.g. high sample attrition) do not meet the criteria for the high rating, matched comparison group designs, and single case and regression discontinuity designs that meet WWC design standards with reservations.[2]

In addition, the *PFL* evaluation team conducted a supplementary literature search according to the criteria outlined by HomVEE, however no other relevant studies were identified. The table overleaf presents the findings observed between thirty-six months to four-five years from the HomVEE website from studies published after 1989. The results reported below are based on comparisons between home visiting intervention groups and control groups. Table 1.2 contains a summary of the main findings as they apply to the *PFL* evaluation. These will be discussed in greater detail in Chapter 2.

[2] Studies rated as "low" by HomVEE have not been included.

Table 1.1 - Evaluations of Outcomes for Home Visiting Programmes.

Outcome	Author	Sample Size	Programme
Child Development & School Readiness	Connell et al. (2008)	627	Family Check-Up
	Shaw et al. (2009)	619	Family Check-Up
	Shaw et al. (2006)	109	Family Check-Up
	Olds et al. (1994)	236	Nurse Family Partnership
	Olds et al. (2004)	407 - 424	Nurse Family Partnership (nurse-visited mothers)
		409 - 441	Nurse Family Partnership (paraprofessional-visited mothers)
	Landry et al. (2008)	166	Play and Learn Strategies (PALS)
	Drazen & Haust (1993)	24-40	Parents as Teachers
Child Health	Klinnert et al. (2007)	149	Childhood Asthma Prevention Study
	Olds et al. (1994)	221	Nurse Family Partnership
Positive Parenting Practices	Olds et al. (1994)	211-238	Nurse Family Partnership
	Olds et al. (2004)	635	Nurse Family Partnership (paraprofessional-visited mothers)
	Landry et al. (2008)	166	Play and Learn Strategies (PALS)
	Landry et al. (2012)	166	Play and Learn Strategies (PALS)
Reductions in Child Maltreatment	Olds et al. (1994)	221-253	Nurse Family Partnership
Home Environment	Olds et al. (1994)	209-238	Nurse Family Partnership
	Olds et al. (2004)	635	Nurse Family Partnership
	Shaw et al. (2006)	109	Family Check-Up
Maternal Health and Well-Being	Olds et al. (1988)	203-216	Nurse Family Partnership
	Olds et al. (2004)	424 - 441	Nurse Family Partnership
	Klinnert et al. (2007)	149	Childhood Asthma Prevention Study
Social Support	Olds et al. (1988)	183	Nurse Family Partnership
Household Factors and SES	Olds et al. (2004)	635	Nurse Family Partnership
	Olds et al. (1988)	95 - 208	Nurse Family Partnership
Reductions in Juvenile Delinquency, Family Violence, and Crime	Olds et al. (2004)	635	Nurse Family Partnership

Favourable impact. A statistically significant impact on an outcome measure in a direction that is beneficial for children and parents

	Measures used	Sig. Finding [Effect Size *d*]	Effect	Timing
	CBCL internalising (including emotional reactivity, anxiety and depression, somatic problems and withdrawal subscales) and externalising (including aggression and attention problems subscales), transitions from internalising/ externalising problems to "normal" from age 2 to 4)	CBCL Internalising problems, co-occurring CBCL internalising and externalising problems	Favourable	48 months
	CBCL (internalising and externalising), Eyberg Child Behaviour Inventory (problem factor)	CBCL internalising behaviours (0.23) CBCL externalising behaviours (0.21) Eyberg problem behaviour (0.23)	Favourable	48 months
	CBCL (internalising and externalising) with a focus on destructive and aggressive behaviours	Time X Treatment interaction in the destructive scale (*d*= .45 standard deviations)	Favourable	48 months
	Stanford-Binet, child behaviour problems noted in physician records	Child behaviour problems	Favourable	48 months
	Children's externalising behaviour (rule breaking and aggressive behaviour), Pre-School Language Scales-3, executive functioning, and ability to regulate behaviour and emotions	None)	None	48 months
	Pre-School Language Scales-3, behavioural adaption, emotional regulation, executive functioning, externalising behaviour problems	None	None	48 months
	PPVT-III receptive vocabulary, PLS: auditory comprehension and expressive comprehension, cooperation, coordinating attention and words, social engagement, and use of words	PPVT-3 receptive vocabulary (0.36), cooperation (0.30), social engagement (0.32), and use of words (0.37)	Favourable	38 months
	Achievement and mental processing (Kaufman Assessment Battery for Children), fine and gross motor skills, language (Zimmerman Preschool Language Scale)	Gross motor delay (-0.77), language acquisition (0.57), and mental processing delay (-1.27)	Favourable	4-5 years
	Asthma status, medication use, emergency department visits	None	None	48 months
	Paediatric and hospital records of: the number of health supervision visits, number of physician visits at which a separate illness was detected, number of injuries and ingestions, visits to emergency departments, and hospitalisations	Number of injuries and ingestions, visits to emergency departments	Favourable	48 months
	Physician records of parental coping problems, parental warmth, control and involvement	Parental coping problems	Favourable	46-50 months
	Sensitive/responsive interaction	Sensitive/responsive interaction (0.23)	Favourable	48 months
	Contingent responsiveness, maintaining child foci, positive affect of mother with child, redirecting child foci, verbal encouragement, verbal scaffolding, warm sensitivity	Verbal encouragement (*PALS II* only group) (0.25), contingent responsiveness (*PALS I* + *PALS II*) (0.51), redirecting child foci (*PALS I* + *PALS II*) (0.39)	Favourable	38 months
	Comments on books, general verbal support, language facilitation, open prompts, percent time only reading text, praise and encouragement, responsiveness	Open prompts (*PALS I* + *PALS II*) (0.38)	Favourable	38 months
	New substantiated cases of child abuse and neglect, number of emergency department visits for ingestion/injuries, number of injuries/ingestions in physicians' records	Number of injuries/ingestions in physicians' records	Favourable	48-50 months
	Home and car safety, HOME inventory, presence and availability of poisonous substances, observed hazards in home	Observed hazards in home	Favourable	46 months
	HOME inventory	None	None	48 months
	HOME involvement subscale	HOME involvement (0.49)	Favourable	48 months
	Number of subsequent pregnancies, number of spontaneous abortions, number of therapeutic abortions, number of births including first born, number of months between first and second child	Subsequent pregnancies	Favourable	46 months
	Number of subsequent pregnancies, number of subsequent live births, months between births of 1st and 2nd children, mastery, maternal mental health, subsequent miscarriage, subsequent abortion, currently using marijuana, moderate to heavy drinker, maternal behavioural problems attributable to substance abuse	Months between births of first and second children (0.32) *(nurse visited mothers)* Subsequent miscarriage (0.42), mastery (0.20), and mental health (-0.03) *(paraprofessional visited mothers)*	Favourable	48 months
	Caregiver quality of life	Caregiver quality of life	Favourable	48 months
	Help with childcare	None	None	46 months
	Months employed, months with current partner, use of welfare services, educational achievement, married, living with father of child	Months employed (0.11), living with father of child (0.64) *(paraprofessional visited mothers)*	Favourable	48 months
		Married (0.61) *(paraprofessional visited mothers)*	Unfavourable	48 months
	Number of days on public assistance, employment, educational achievement	Employment	Favourable	46 months
	Domestic violence	Domestic violence in the past 6 months (0.47) *(nurse visited mothers)*	Favourable	48 months

Unfavourable or ambiguous impact: A statistically significant impact on an outcome measure in a direction that may indicate potential harm to children and/or parents.

Effect size statistics are only included for those studies which report a relevant figure. When an effect size is reported and not precisely described it is assumed to be a Cohen's *d* statistic calculated using pooled variance. Odds ratios have been transformed to Cohen's *d* effect sizes according to the following formula: $d = LogOddsRatio \frac{\sqrt{3}}{\pi}$

Table 1.2 - Summary of the Main Findings Outlined in Table 1.1

Domain	No. of programmes identifying favourable effects	Summary of positive findings
Child Development	6 out of 7 programmes	• Fewer internalising and externalising problems • Improved receptive and expressive language • Improved cooperation • Improved social engagement • Better gross motor skills • Enhanced mental processing
Child Health (including reductions in child maltreatment)	1 out of 2 programmes	• Fewer ingestions and injuries • Reduction in visits to emergency departments
Parenting	4 out of 4 programmes	• Improved parenting coping • Increase in sensitive/responsive interactions • Increased verbal encouragement • Enhanced joint reading behaviour
Home Environment	2 out of 3 programmes	• Reduction in number of observed hazards in home • Improve HOME involvement scores
Maternal Health	3 out of 3 programmes	• Mother less likely to have subsequent pregnancies • Subsequent pregnancies less likely to end in miscarriage • Enhanced sense of mastery and mental health • Improved caregiver quality of life
Maternal Social Support	0 out of 1 programme	N/A
Childcare	0 out of 1 programme	N/A
Household Factors & SES	2 out of 2 programmes	• Increased rates of employment • Mother more likely to live with child's father

1.9 Hypotheses

The primary aim of the *PFL* programme is to change parental knowledge, attitudes and feelings leading to improved parenting behaviour, which will then positively impact on child development, ultimately increasing a child's school readiness. *PFL* also hypothesises that the programme will have an effect on other child and family outcomes (e.g. social support, maternal health and wellbeing, and household factors and SES). Therefore, *PFL* may affect both primary and secondary outcomes. In effect, secondary outcomes may serve as mediators or explanatory factors that may help to clarify the relationship between the *PFL* programme and any observed effects on parenting skills or child school readiness.

For the main results (high versus low treatment groups), our hypotheses regarding the effectiveness of the *PFL* programme at forty-eight months are informed by the evidence described above. Results from previous studies indicate that at forty-eight months, home visiting programmes had moderate positive effects in the domains of child development, parenting, maternal health and wellbeing, and household factors and SES. Therefore, we would expect that, consistent with these findings, the impact of *PFL* on these domains will be moderate. As reported in other studies, we expect to find limited programme effects in the realms of child health, social support, and the home environment. As effects on childcare are rarely examined in home visiting evaluations it is difficult to ascertain the impact of *PFL* on this area, however we anticipate that it will be limited. Consistent with the evaluation design, we anticipate few significant differences between the low treatment group and the comparison group.

1.10 Description of the Forty-Eight Month Survey & Data Collection Process

Between July 2012 and June 2015, a seventh research interview was conducted by the *PFL* evaluation team. The interviews took place when the *PFL* child was between two weeks before their fourth birthday and up to six months after their birthday. In total, 217 forty-eight month interviews (n*High* = 74; n*Low* = 73; n*LFP* = 70) were completed. The average age of the target child at the time of interview was 49.04 months old (*SD*= 1.34 weeks). These 217 participants represent 65% of the original sample recruited into the study (n*High* = 115; n*Low* = 118; n*LFP* = 99). The forty-eight month completion rate was very similar for the high (64%) and low (62%) treatment groups, and slightly higher for the comparison group (71%). The official dropout rate between thirty-six months and forty-eight months was minimal. All of the high treatment and comparison groups were retained during this period, and one member of the low treatment group dropped out between thirty-six and forty-eight months. However, the level of disengagement was at its highest level, in the forty-eight month interview. A comprehensive analysis of attrition rates may be found in Chapter 3.

The forty-eight month interview lasted approximately 2 hours and was conducted using a Computer Assisted Personal Interviewing (CAPI) technique on tablet laptops. The interviews were conducted by trained interviewers who were blinded to participant treatment status. Immediately prior to the interview, participants were asked to complete the Achenbach Child Behaviour Checklist (CBCL) on paper. Although home interviews were encouraged, participants also had the option of conducting the interview in a local community centre. The vast majority of participants completed the interview in their homes (81% in the high treatment group, 89% in the low treatment group, 100% in the comparison community), while others completed them in the community centre (19% in the high treatment group, 10% in the low treatment group), or another home (1% of the low treatment group). Each participant was given a €20 shopping voucher after the forty-eight month interview was completed as a thank you for taking part.

A number of questions/measures included in the interview were used at previous time points, while a number of new questions/measures were added. The repeated questions related to family demographics and socioeconomic profile, maternal physical and psychological health, substance use, family risk factors, safety in the home, parenting stress, parental monitoring of TV, use of childcare, child motor skills, cognitive development, behavioural, and emotional functioning, social-emotional development, and child health. New questions added to the forty-eight month interview included items related to primary school attendance, school readiness traits, activities with child, peer problems and pro-social behaviour, and children's sleep habits.

The forty-eight month survey was divided into 9 modules, each containing questions with a common theme.

1. Your Child's Development: Part 1
2. Update on Your Life
3. Your Thoughts on Parenting
4. Your Social Support Network
5. Your Health
6. Your Child's Development: Part 2
7. Family Environment
8. Your Child's Health
9. Closing

Similar to previous reports, this report focuses on eight domains incorporating 33 categories and 191 outcome measures. The domains and categories within each domain are – child development (Ages & Stages Questionnaire, Developmental Profile-3, Achenbach Child Behaviour Checklist, Strengths and Difficulties Questionnaire), child health (child physical health, mother's health decisions for her child, Children's Sleep Habits Questionnaire), parenting (Parenting Daily Hassles Scale, Parenting Stress Index, Parenting Styles and Dimensions Questionnaire, Home Learning Environment, child protective services involvement, parental monitoring of TV), home environment (Injury Prevention Program, Framingham Safety Survey), maternal health and wellbeing (maternal physical health, maternal mental health, drug and alcohol use, Pearlin Self-efficacy Scale, Rosenberg Self-esteem Scale, Edinburgh Postnatal Depression Scale, WHO-5 Index), social support (voting, father involvement, and support from relatives, friends and neighbours childcare (childcare use, type & satisfaction), and household factors and SES (household factors, parental education, parental employment, household finances, and expectations of future finances). Note that while the same domains were investigated at each time point the measures included in the forty-eight month report may differ to those included at previous time points.

1.11 Overview of Report

The report is organised as follows. Chapter 2 presents the results comparing the *PFL* high treatment group and the *PFL* low treatment group on all primary outcome domains (child development, child health, parenting) and secondary outcome domains (home environment, maternal health and wellbeing, social support, childcare, household factors and SES). Chapter 3 presents an implementation analysis of the *PFL* programme between programme intake and forty-eight months. Chapter 4 presents a summary of the results comparing the *PFL* low treatment group to the community comparison group and a summary of the results from the dynamic analysis examining changes in child and parent outcomes over time. Chapter 5 presents interaction and subgroup analyses examining the main results according to gender, parity, maternal cognitive resources, and Triple P status. Chapter 6 summarises and interprets the findings.

Chapter Two

Main Results
High and Low Treatment Groups

2.1 Introduction

This chapter presents the main results comparing the forty-eight month outcomes of the high treatment group to those of the low treatment group. As there were no statistical differences, on average, between these groups before the programme began, any identified statistical differences between the two groups at forty-eight months are indicative of a programme effect. The analysis focused on eight domains - child development, child health, parenting, the home environment, maternal health and wellbeing, social support, childcare, and household factors and SES. Although each report contains the same overarching eight domains, measures which focus on different aspects of these domains were utilised at each time point. Therefore, it is not always possible to make a direct comparison between the present findings and findings outlined in the five previous reports. This chapter contains relevant literature for the new measures which were not included in previous reports and considers the relevance and impact of previous home visiting programmes on all measures at forty-eight months. Each section also includes a description of the instruments used to measure the domain and the statistical results, in both text and table format, comparing the high and low treatment groups on that domain. Each section should be read in conjunction with the corresponding section in Chapter 3 of '*Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at Six Months' and Chapter 2 of our previous reports as these will be referenced where relevant. These reports can be found at the following website under publications: *http://geary.ucd.ie/preparingforlife*.

This chapter proceeds as follows: Section 2.2 outlines the methods used to conduct the analyses and information on how to interpret the outcomes tables. Sections 2.3 to 2.10 present the results for each of the eight main domains under analysis. Section 2.11 provides a summary of the main results of the *PFL* evaluation at forty-eight months.

2.2 Methods & Description of Outcome Tables

A full description of the methodology used to analyse each wave of outcome data may be found in '*Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at Six Months'. It describes the permutation method used for hypothesis testing[1], including conditional permutation testing, the step-down procedure which is used for multiple hypothesis testing, and the procedure for dealing with missing data due to item non-response[2].

The following information is included in the outcome tables and provides a reference for interpreting the results.

[1] Note that due to an improvement in computing power, the permutation testing is now conducted with 100,000 replications.

[2] Overall, the extent of missing information in the forty-eight month data is low; less than 10% of data were missing in each psychometric scale, with the majority of scales missing less than 1% of data . In order to account for missing data, interpolation methods were used. Note that such methods were only used for standardised psychometric scales, as it is possible to utilise information within that scale to replace the missing data. In cases where data were missing on single item measures, these observations were excluded from the analysis. On average, over 99% of data were present for single item measures.

N	*N* represents the number of respondents who are included in the analysis.
M	*M* is the mean, or average value, of responses. This statistic represents the average response of all participants who answered the question of interest. For binary variables, this value can be interpreted as the proportion of the sample who reported being in the category described.
SD	*SD* is the standard deviation. This is calculated by summing the squared difference between each observed response and the average response. This sum is then divided by the total number of observations to derive the average squared difference between responses and the mean. The square root of the resulting figure gives the standard deviation. It serves as a useful indicator of how varied the responses were.
Low/High/ LFP	*Low/High/LFP* subscripts attached to the summary statistics *(N, M, and SD)* indicate the subgroups for which the summary statistics have been calculated.
Individual Test p^1	In this chapter the data are first grouped by *PFL* treatment status (low treatment and high treatment) to examine forty-eight month differences within the *PFL* cohort. In Chapter 4 the low treatment group is compared to the comparison group. Classical statistical tests rely on the assumption that sample sizes are large, and produce inferences based on *p*-values that are only valid for large samples. These tests can be unreliable when the sample size is small. As the sample size of *PFL* is relatively small, particularly as the evaluation progresses over time, all the analyses comparing the forty-eight month outcomes of the high and low treatment groups use an alternative approach called permutation-based hypothesis testing. This approach is appropriate for small samples and was used to analyse data for a similar evaluation of the Perry Preschool Program by Heckman and colleagues (2010). The individual one-tailed *p*-value represents the probability of observing differences between the two groups by chance. In cases where there is a statistically significant difference between the two groups, a *p*-value is presented which indicates the likelihood that the group difference could have randomly occurred. A *p*-value of less than 0.10 is considered to be statistically significant and conveys that the probability of the difference between the two groups being due to chance is less than 10%. Similarly, *p*-values of less than 0.05 and 0.01 indicate that the probability of the difference between the two groups being due to chance is less than 5%, or 1% respectively. Low *p*-values (i.e., significant results) indicate that the high treatment group is outperforming the low treatment group. *p*-values are presented for significant differences only. Differences that are significant in the non-hypothesised direction are denoted by s~. Non-significant differences are denoted by 'ns'.
Step-down Test p^2	As 191 outcome measures are considered in this report, it is possible that we may reject some of these null hypothesis by chance (i.e. we may identify a significant difference between the high and low treatment groups on certain outcomes when there is, in fact, no significant difference). Multiple hypothesis testing allows us to test for the joint significance of multiple outcomes at the same time, thus minimising the likelihood of finding treatment effects that are false. The multiple hypothesis method we use is called the step-down procedure. We apply the step-down procedure of Romano and Wolf (2005) and its extension by Heckman et al., (2010). Their methods control for overall error rates for vectors of hypothesis using the family-wise error rate (FWER), the probability of yielding one or more false positives out of a set of hypothesis tests, as a criterion. The *p*-value from the step-down test may be interpreted in the same manner as the individual *p*-value discussed above. Each *p*-value in the step-down test represents the joint test of all outcomes included in that category. For example, the *p*-value corresponding to the first outcome represents a test of the joint significance of all outcomes included in that category. The next *p*-value corresponding to the second outcome in that category represents the test that all remaining outcomes in that category are jointly significant, excluding the first outcome. Similarly, the *p*-value corresponding to the third outcome in that category represents a test of the joint significance of all the outcomes remaining in that category, excluding the first two outcomes. Note that all outcomes in the tables are organised according to their individual test-statistic, such that the measure with the largest test-statistic is listed first and the outcome with the smallest test-statistic is listed last within that category. Thus, the ordering of the outcomes in the tables (within categories) is indicative of the strength of the treatment effects.
Effect Size *d*	Effect size *(d)* illustrates the magnitude of the difference between the groups. While the p-value allows the reader to determine whether or not there is a statistically significant difference between groups, it does not indicate the strength of the difference. As the strength of a relationship can provide valuable information, the effect size was calculated using Cohen's d. A Cohen's d ranging from 0.0 to 0.2 is deemed a small effect; values ranging from 0.2 to 0.8 represent a medium effect; and values greater than 0.8 illustrate a large effect (Gravetter & Wallnau, 2004).

2.3 Child Development

The early childhood period is a critical developmental stage as children develop rapidly across multiple domains. Healthy physical, cognitive, and socio-emotional development in early childhood are key determinants of later life outcomes and are predictive of wellbeing, mental health, physical health, competence in literacy and numeracy, criminality, and economic participation throughout the life course (Illig, 1998; Irwin et al., 2007). A child's developmental trajectory is also strongly influenced by epigenetic factors such as a family's economic, educational, social, emotional and environmental circumstances, as well as parenting resources and practices (Illig, 1998; Tremblay, 2010). In particular, social disadvantage is consistently identified as a predictor of poorer life outcomes (McLoyd, 1998). However, meta-analytic evidence suggests that early intervention programmes may be an effective strategy to reduce socioeconomic inequalities and promote healthy child development (Gomby, 2005; Kahn & Moore, 2010). In this section we describe different areas of development and review the impact of home visiting interventions on each one.

PHYSICAL DEVELOPMENT: GROSS AND FINE MOTOR SKILLS

Physical development refers to growth and also the ability to use muscles and body parts to execute large (gross motor) and small (fine motor) muscle movements (Carr, 2006). The acquisition of these skills is fundamental for daily life tasks from walking/running to writing (Cools, De Martelaer, Samaey, & Andries, 2009). At four years of age most children have developed skilled control over their own movement; they can turn sharply, climb ladders and trees, and typically navigate stairs using only one foot to a step (American Academy of Pediatrics; 2013, Sheridan, Frost, & Sharma, 2004). They can run and walk on tiptoe, and demonstrate increasing skill at ball games as their co-ordination develops (Sheridan et al., 2004). In terms of fine motor skills, they can generally hold pencils in a tripod grasp like an adult with good control. They can copy basic patterns such as crosses and some letters and usually include features such as head, legs, and arms when drawing people (Sheridan et al., 2004).

COGNITIVE AND LANGUAGE DEVELOPMENT

Although somewhat independent, children's language and cognitive development are interrelated (Gopnik & Meltzoff, 1987 as cited in Carr, 2006). Cognitive development relates to the ability to think, reason, and understand. Language development is intertwined with cognitive development and provides children with an avenue to communicate with and understand others (Goswami, 2002). Four year old children typically speak in a grammatically correct and completely intelligible manner with only small sound substitutions and shortened words (Sheridan et al., 2004). They can listen to and tell stories and give connected accounts of recent events and experiences. Children of this age also show a sense of humour in conversation and social activities, they understand jokes and some abstract concepts. They know and repeat several nursery rhymes and regularly ask why, how, and when questions (American Academy of Pediatrics, 2013; Sheridan et al., 2004).

PERSONAL, SOCIAL, AND EMOTIONAL DEVELOPMENT

Socio-emotional development is the cornerstone of effective communication (Bartolotta & Shulman, 2009). Children who can effectively experience, manage, and express emotion form close and secure relationships and are better able to cooperate with others, pay attention, and transition between activities (Nisha, 2006; Zero to Three, 2001). In contrast, children who have socio-emotional difficulties are at greater risk of behavioural problems and poor academic performance (Denham, Wyatt, Bassett, Echeverria, & Knox, 2009). By four years of age children's general behaviour is more independent and strongly self-willed. They engage with other children and will alternatively co-operate and argue with them. They understand the need to use words rather than physicality in arguments and will tend to quarrel with playmates and adults when they disagree (Sheridan et al., 2004). They usually favour make-believe play and dressing up, will show concern for younger siblings and sympathy for distressed playmates (Sheridan et al., 2004).

IMPACT OF HOME VISITING INTERVENTIONS ON CHILD DEVELOPMENT AT FORTY-EIGHT MONTHS

Overall, evidence suggests that home visiting programmes can positively influence child development at forty-eight months. Evaluations of four home visiting programmes all report favourable child development outcomes between thirty-six and forty-eight months. A series of three papers evaluating the Family Check-Up model, which incorporates three home visits in advance of an intensive parenting support intervention, all report more favourable behavioural outcomes for children in the intervention group than those in the control group (Connell et al., 2008; Shaw et al., 2006; Shaw et al., 2009). Similarly, results from two studies investigating the effectiveness of the Nurse Family Partnership (NFP) programme, which includes one-to-one home visits by trained public health nurses, report favourable outcomes at forty-eight months (Olds et al., 1994; Olds et al., 2004). The first study conducted by Olds et al. (1994) reported better outcomes for intervention group children over control group children in relation to child behaviour problems. In the second study, Olds et al. (2004) found that nurse-visited participants had better behavioural adaptation and executive functions. However, when an alternative implementation of NFP using paraprofessionals rather than nurses was assessed, there were no significant differences on child development outcomes. Results from two other home visiting evaluations also report favourable outcomes for intervention families. Landry et al. (2008) found that recipients of the Play and Learn Strategies (PALS) programme had better receptive vocabulary, better use of words, were more cooperative, and showed greater social engagement. Similarly, Drazen and Haust (1993) in their evaluation of the Parents as Teachers (PAT) programme, found that intervention children were less likely to have gross motor delay, and had better mental processing and language acquisition than their control group counterparts.

2.3.1 Child Development Instruments

AGES AND STAGES QUESTIONNAIRE

Child development in the *PFL* evaluation was assessed using the forty-eight month version of the Ages and Stages Questionnaire (ASQ; Squires et al., 1999). The ASQ was designed as an effective screening measure for young children who were considered to be at risk for developmental delay. The ASQ child monitoring system consists of 19 screening questionnaires at specific age intervals ranging from four to sixty months of age and provides scores across five domains of child development, with each domain comprising six items. Communication (α=0.61) measures the child's understanding of language, naming of items, and word combinations. The Gross Motor domain (α=0.47) measures the child's walking, running, and jumping movements. The Fine Motor domain (α=0.54) assesses the child's finger and hand movements, including stacking and threading. Problem Solving (α=0.47) measures the child's ability to follow instruction, pretence, and problem solving. Finally, the Personal-Social domain (α=0.57) provides a rating of eating skills, solitary play, and self-awareness. Cronbach's alpha (α) is used as an internal consistency estimate of reliability for test measures. Generally, an α greater than 0.7 indicates acceptable internal consistency, an α between 0.5 and 0.7 represents questionable to poor reliability, and an α of less than 0.5 indicates an unreliable measure. As such, the low alphas for the above ASQ measures should be taken into account when interpreting results. During the interview, the interviewer asked the mother questions related to different activities her child was capable of at that time. The mother responded by indicating whether her child exhibited the behaviour regularly, sometimes, or not yet. If the mother did not know whether her child was capable of the behaviour, where possible, the interviewer asked her to test the behaviour with the child during the interview using the ASQ toolkit. Domain scores represent the sum of all six items in that domain, resulting in a possible range of 0 to 60 with higher scores indicative of more advanced development.

In addition, the ASQ provides age-specific standardised cut-off points for each domain (Communication = 39.1; Gross Motor = 32.9; Fine Motor = 30.7; Problem Solving = 35.0; and Personal-Social = 23.4). In line with these cut-off scores, a binary variable was calculated for each domain illustrating if the child scored below the cut-off point. Those children who scored below the cut-off point were considered to be at risk of developmental delay in that domain. Furthermore, an ASQ standardised total score was created, with a mean of 100 and standard deviation of 15, for each domain. These standardised scores for Communication,

Gross Motor, Fine Motor, Problem Solving and Personal-Social were then summed and standardised again within the sample, to a mean of 100 and standard deviation of 15, to produce the ASQ standardised total score.

AGES AND STAGES QUESTIONNAIRE: SOCIAL-EMOTIONAL

Children's social-emotional development was assessed using a modified 30-item version of the Ages and Stages Questionnaire: Social-Emotional (ASQ:SE; Squires et al., 2003). The ASQ:SE (α=0.79) is a screening tool used alongside the ASQ to identify children from six to sixty months of age who are in need of further social and emotional behavioural assessment. Questions on the ASQ:SE pertain to self-regulation, compliance, communication, adaptive functioning, autonomy, affect, and interaction with people. During the interview, the interviewer asked the mother questions related to different behaviours the child displays. The mother responded by indicating whether her child exhibited the behaviour *most of the time, sometimes,* or *never.* Additionally, the mother indicated if the behaviour was a concern for her. Scores to each item were rated on a 0 to 10 scale and an additional five points were added to the score for every indication that the behaviour was a concern for the mother. Scores were summed to provide a total ASQ:SE score, with a possible range of 0 to 450. Higher scores indicated that the child may be at risk of poor social-emotional development. In addition, the ASQ:SE provides a cut-off score of 70 and suggests that children with scores above this cut-off may be at risk. In line with this cut-off, a binary variable was calculated to illustrate if the child was at risk of poor socio-emotional development.

STRENGTHS AND DIFFICULTIES QUESTIONNAIRE: PEER PROBLEMS AND PRO-SOCIAL SUBSCALES

The Strengths and Difficulties Questionnaire (Goodman, 1997) is a 25-item questionnaire assessing behaviours, emotions, and relationships of 4 to 16 year olds. The questionnaire covers five dimensions: conduct problems, emotional symptoms, hyperactivity, peer problems, and pro-social behaviour. The 5-item Peer Problems (α=0.48) and 5-item Pro-Social (α=0.72) subscales were used at forty-eight months. Items were scored 0 for *not true*, 1 for *somewhat true*, and 2 for *certainly true*. Two items from the Peer Problems subscale were reverse scored. The five items for each subscale were summed giving a total score of 0 to 10 for each subscale.

CHILD BEHAVIOR CHECKLIST

The Child Behavior Checklist for Ages 1½ -5 (CBCL; Achenbach & Rescorla, 2000) is a parent report instrument for assessing behaviour in children. It provides scores for a range of internalising and externalising problems for children aged eighteen months to five years. The CBCL consists of seven syndromes; emotionally reactive (α=0.75), anxious/depressed (α=0.76), somatic complaints (α=0.65), withdrawn (α=0.76), sleep problems (α=0.76), attention problems (α=0.71), aggressive behaviour (α=0.91), and one 'other problems' (α=0.83) category. These eight categories map onto two subscales, Internalising (α=0.90) and Externalising Problems (α=0.92), and also a Total Problems score (α=0.96). Mothers were asked to complete the CBCL with pen and paper before beginning the main part of the interview. This consisted of 100 questions with the response options *not true, somewhat/sometimes true, or very true/often true.* These were scored as 0, 1, and 2 respectively. From the 100 questions, eight raw scores were produced (seven syndromes and other category, as above). The raw scores of the emotionally reactive, anxious/depressed, somatic complaints, and withdrawn subscales were totalled as an Internal Problems score. Correspondingly, the attention problems and aggressive behaviour syndromes were totalled to produce the External Problems score. The Sleep Problems subscale is separate to both of these categories. Finally, the totals of all seven syndromes plus the other problems subscale were combined to produce a Total Problems score. The clinical cut-off range was identified for each domain as follows: an Internal Problems score of above 17, an External Problems score of above 24, and a Total Problems score of above 60 indicating that the child was at risk of clinically significant problems. The CBCL produces a total of 14 scores: three domains, three domain cut-offs, and eight sub-domains.

DEVELOPMENTAL PROFILE-3: COGNITIVE SECTION

The Developmental Profile-3 (DP-3; Alpern, 2007) is a parent-report measure of child development from birth to age twelve years and eleven months. The *PFL* evaluation includes the DP-3 cognitive section which measures cognitive abilities (α=0.79) using a 38-item scale. Each of the items refers to tasks which require cognitive skill and were arranged in order of difficulty, for example: *'Does your child say size words (large or big, and little or small) correctly'.* For each item, mothers were asked whether their child had carried out the task and responded *yes* or *no* accordingly. The yes responses were tabulated to create a continuous score whereby higher values indicated greater cognitive development. These scores were standardised by age according to the normative sample provided in the DP-3 manual, with a mean of 100 and standard deviation of 15. In addition, a binary variable was created to indicate those scoring above average, that is, a score of above 115.

SERVICES RECEIVED

Participants were asked to indicate if their child was receiving any special services, specifically any services to help them catch up in areas such as speech or physical development. A binary variable was created using this question.

2.3.2 Child Development Results

ASQ SCORES

Table 2.1 presents the results comparing the high and low treatment groups on the child development domain.

Within the ASQ scores category, four of the six child development measures were in the hypothesised direction and one of these, ASQ Fine Motor score, was statistically significant. The high treatment group scored an average of 47.30 on this subscale while the low treatment group scored an average of 43.84 (p<.10, d=0.27) indicating that children in the high treatment group were more likely to display developmentally appropriate fine motor skills than children in the low treatment group. The step-down test showed that the joint effect of all six measures in the ASQ scores category was not statistically significant.

ASQ CUT-OFF SCORES

Within the ASQ cut-off scores category, which measures the proportion of children at risk of developmental delay, four of the six measures were in the hypothesised direction. One of these differences was statistically significant. Fourteen percent of children in the high treatment group were at risk of developmental delay regarding fine motor skills, compared with 26% of children in the low treatment group (p<.05, d=0.32). The step-down test showed that the joint effect of all six measures in the ASQ cut-off Scores category was not statistically significant.

SDQ

Within the SDQ category, both measures were in the hypothesised direction, however neither was statistically significant. Furthermore, the step-down test showed that the joint effect of both measures was not statistically significant.

CBCL DOMAINS

Within the CBCL category, all three measures were in the hypothesised direction, however none were statistically significant. Furthermore, the step-down test showed that the joint effect of the CBCL scores was not statistically significant.

CBCL DOMAINS CUT-OFF

Within the CBCL domains cut-off category, all three measures were in the hypothesised direction, and two of these differences were statistically significant. In terms of External Problems, none of the high treatment group were rated as having problems at the clinical level, compared with 6% of the low treatment group ($p<.05$, $d=0.35$). Regarding Internal Problems, 4% of the high treatment group were rated as having problems at the clinical level, compared with 11% of the low treatment group ($p<.10$, $d=0.27$). In addition, the step-down test showed that the joint effect of the CBCL domain cut-off scores was statistically significant. The joint effect finding was driven by the significant results found for the CBCL External Problems cut-off ($p<.05$) and Internal Problems cut-off ($p<.10$).

CBCL SUBDOMAINS

Within the CBCL subdomains category, all eight measures were in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the CBCL subdomains was not statistically significant.

NON STEP-DOWN MEASURES

Three of the four non step-down measures were in the hypothesised direction. In two of these cases the differences between the high and low treatment groups reached statistical significance. The high treatment group scored an average of 108.20 on the DP-3 cognitive development standardised score, compared to the low treatment score of 102.88 ($p<.05$, $d=0.36$). Furthermore, 34% of the high treatment group were scored as above average on the DP-3 compared to 19% of the low treatment group ($p<.05$, $d=0.34$). This indicates that the high treatment group were displaying more advanced cognitive abilities than the low treatment group.

Table 2.1 - Results for High and Low Treatment Groups: Child development

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size d
ASQ Scores									
ASQ Fine Motor Score	147	(74/73)	47.30	(12.47)	43.84	(13.48)	$p<.10$	ns	0.27
ASQ Personal-Social Score	147	(74/73)	51.95	(10.85)	49.86	(11.52)	ns	ns	0.18
ASQ Communication Score	147	(74/73)	53.92	(8.16)	53.15	(9.37)	ns	ns	0.08
ASQ Problem Solving Score	147	(74/73)	53.45	(9.03)	52.95	(9.31)	ns	ns	0.05
ASQ Gross Motor Score	147	(74/73)	53.92	(8.53)	54.04	(8.65)	ns	ns	0.01
* ASQ Social-Emotional Score	147	(74/73)	33.18	(27.59)	32.67	(30.20)	ns	ns	0.02
ASQ cut-off scores									
* ASQ Fine Motor cut-off	147	(74/73)	0.14	(0.34)	0.26	(0.44)	$p<.05$	ns	0.32
* ASQ Personal-Social cut-off	147	(74/73)	0.03	(0.16)	0.04	(0.20)	ns	ns	0.08
* ASQ Gross Motor cut-off	147	(74/73)	0.03	(0.16)	0.04	(0.20)	ns	ns	0.08
* ASQ Communication cut-off	147	(74/73)	0.05	(0.23)	0.07	(0.25)	ns	ns	0.06
* ASQ Social-Emotional cut-off	147	(74/73)	0.07	(0.25)	0.07	(0.25)	ns	ns	0.00
* ASQ Problem Solving cut-off	147	(74/73)	0.05	(0.23)	0.05	(0.23)	ns	ns	0.00
Strengths and Difficulties Questionnaire (SDQ)									
SDQ Pro-Social Behaviour	147	(74/73)	8.47	(1.59)	8.18	(1.85)	ns	ns	0.17
* SDQ Peer Problems	147	(74/73)	1.37	(1.48)	1.45	(1.51)	ns	ns	0.06
Child Behaviour Checklist (CBCL) domains									
* CBCL External Problems	146	(74/72)	7.45	(5.65)	8.64	(8.66)	ns	ns	0.16
* CBCL Total Score	146	(74/72)	22.15	(14.53)	24.69	(23.65)	ns	ns	0.13
* CBCL Internal Problems	146	(74/72)	6.50	(5.03)	6.86	(7.40)	ns	ns	0.06
Child Behaviour Checklist (CBCL) domains cut-off scores									
* CBCL External Problems cut-off	146	(74/72)	0.00	(0.00)	0.06	(0.23)	$p<.05$	$p<.05$	0.35
* CBCL Internal Problems cut-off	146	(74/72)	0.04	(0.20)	0.11	(0.32)	$p<.10$	$p<.10$	0.27
* CBCL Total Score cut-off	146	(74/72)	0.03	(0.16)	0.07	(0.26)	ns	ns	0.20
Child Behaviour Checklist (CBCL) subdomains									
* CBCL Aggressive Behaviour	146	(74/72)	5.70	(4.77)	6.81	(7.02)	ns	ns	0.19
* CBCL Other Problems	146	(74/72)	6.07	(4.21)	6.85	(6.67)	ns	ns	0.14
* CBCL Anxious/Depressed	146	(74/72)	1.87	(1.97)	2.13	(2.18)	ns	ns	0.13
* CBCL Sleep Problems Behaviour	146	(74/72)	2.14	(1.97)	2.35	(2.49)	ns	ns	0.10
* CBCL Attention Problems	146	(74/72)	1.74	(1.46)	1.83	(2.16)	ns	ns	0.05
* CBCL Withdrawn	146	(74/72)	1.28	(1.74)	1.35	(2.01)	ns	ns	0.03
* CBCL Somatic Complaints	146	(74/72)	1.39	(1.72)	1.42	(2.14)	ns	ns	0.01
* CBCL Emotionally Reactive	146	(74/72)	1.96	(1.68)	1.97	(2.45)	ns	ns	0.01
Non Step-down Measures									
DP-3: Cognitive Development standardised score	147	(74/73)	108.20	(13.83)	102.88	(15.75)	$p<.05$	-	0.36
DP-3: Cognitive Development above average cut-off	147	(74/73)	0.34	(0.48)	0.19	(0.40)	$p<.05$	-	0.34
ASQ Standardised Total Score	147	(74/73)	100.51	(16.04)	97.93	(15.81)	ns	-	0.16
* Child receiving special services	147	(74/73)	0.19	(0.39)	0.19	(0.40)	ns	-	0.01

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p*-value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p*-value from a Step-down permutation test with 100,000 replications. *d* is Cohen's *d* Effect Size. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not-statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. *'s~'* indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest t-statistic within each Step-down category.

2.4 Child Health

Children's early experiences are central to shaping their long term health and wellbeing. However, research has shown that children from lower SES families typically suffer worse health outcomes than children from higher SES families (Chen et al., 2002). Promoting health in early childhood is therefore critical to improving the health of the whole population and reducing inequalities in health over the longer term (Brandt et al, 2012; Campbell et al., 2014; German & Latkin, 2012; Komro et al., 2011). As well as the physical, psychological and social impacts of poor health on the individual, social inequalities in health also place a huge economic burden on health systems (Lavin & Metcalfe, 2009). Evidence suggests that home visitation can play an important role in health promotion (Koniak-Griffen et al., 2003; Larson, 1980).

GENERAL HEALTH

In terms of general health in childhood, over half of the total hospital discharges of Irish children in 2011 were under five years of age. The most commonly reported diagnosis was disease of the respiratory system (Department of Children and Youth Affairs, 2012). Unintentional injuries was the second most common diagnosis and represents, alongside congenital malformations, the leading cause of mortality in those aged between 12 and 48 months of age (Department of Children and Youth Affairs, 2012). The main place for injury to occur in children under the age of six is in the child's home, and evidence suggests that children from disadvantaged families are at greatest risk from this type of injury (CSDH, 2008; Edwards, Roberts, Green, Lutchmun, 2006; Hyder et al., 2009). Children under four years of age are the most affected, with burns and poisoning being the most commonly reported injuries (Hyder et al., 2009).

Lifelong patterns of eating and physical activity are established during early childhood and thus the greatest potential for avoiding problems such as obesity later in life lies in the early years. In Ireland, the majority (77%) of 2-4 year old children are within the normal weight range, yet 23% are defined as being overweight or obese. This is concentrated among 2 and 3 year olds, with only 8% of 4 year olds being classified by the UK-WHO criteria as overweight or obese (Irish Universities Nutrition Alliance, 2012). However, research has shown that children living in socially disadvantaged areas are at increased risk of weight problems (Greves Grow et al., 2010). Indeed, the Growing Up in Ireland (GUI) study shows that by age thirty-six months, Irish children from disadvantaged backgrounds are more likely to be obese and have poorer diets than children from more advantaged families, indicating an early social gradient in health (Williams et al., 2013). At forty-eight months, it is suggested that an average boy has a weight of 16.5kg and height of 103cm, and an average girl weighs 16kg with a height of 102.5cm (Royal College of Paediatrics and Child Health: http://www.rcpch.ac.uk/child-health/research-projects/uk-who-growth-charts/uk-who-growth-chart-resources-0-4-years/uk-who-0).

TOILETING

Toilet training is both a developmental milestone and a practical skill. It is an indicator of a child's ability for independent action, and is thus one of the criteria for assessing school readiness. With appropriate training and encouragement from primary caregivers, most children begin to stay dry during the day between 17 months and three years of age (Cummins & McMaster, 2006), and full toileting skills are generally achieved by three and a half years (Schum et al., 2002). A number of factors can influence toilet training, including gender, with one study suggesting girls achieve nearly all toilet-training skills earlier than boys (Schum et al., 2002). In a large longitudinal study, Joinson et al. (2008) found that children with delayed development in motor, communication, and social skills, and those with difficult temperament traits, were more likely to experience daytime wetting and soiling into their school years. Relatedly, daytime wetting and soiling have been shown to be associated with externalising and internalising problems in school age children (Joinson et al., 2006).

SLEEP

Sleep is important for many aspects of child development. For example, sleep position has been linked to infant motor development (Davis et al., 1998). Indeed, the GUI study found that Irish infants who were placed on their backs to sleep took their steps later, on average, than those who had been placed on their stomach or side (Williams et al., 2013). Up to 50% of children experience a sleep problem, with 4% having received a formal sleep disorder diagnosis (Carter et al., 2014). Issues can arise around going to sleep, waking up during the night, and needing help to go back to sleep (Nicholson, 2015). The negative consequences of sleep problems include daytime sleepiness, irritability, behavioural problems, learning difficulties, and poor academic performance (Carter et al., 2014). One study found that sleep restriction may increase the risk of weight gain and play a role in the current epidemic of obesity (Van Cauter & Knutson, 2008).

Sleep problems can also be an indicator and prominent feature of other issues such as ADHD (Ren & Qiu, 2014). Recently, Scott et al., (2013) showed that shorter sleep duration and sleep disturbances appear early and predate the usual age of clinical diagnosis in children with ADHD, suggesting the importance of understanding the role of sleep. At forty-eight months, most children need around 11 to 12 hours of sleep and parents are advised to maintain a regular and consistent sleep schedule and have a relaxing bedtime routine in order to help preschoolers get their recommended amount of sleep. (Health Service Executive: http://www.hse.ie/eng/services/publications/Children/caringforyourchild2to5yrs.pdf).

IMPACT OF HOME VISITING INTERVENTIONS ON CHILD HEALTH AT FORTY-EIGHT MONTHS

At forty-eight months two studies report on the impact of home visiting on child health. Firstly, Klinnert et al. (2007) investigated whether the Childhood Asthma Prevention Study decreased the number of children being diagnosed with asthma, the amount of medication being used for asthma, and the number of hospital visits related to the condition. They found that the intervention had no effects at forty-eight months. A study reporting on the Nurse Family Partnership programme found that nurse-visited children had 40% fewer injuries and ingestions and 35% fewer visits to the emergency department than children in the comparison group at forty-eight months (Olds et al., 1994). However, they found no effects on the rates of child abuse and neglect, a result which was previously observed at twenty-four months (Olds et al., 1994).

2.4.1 Child Health Instruments

CHILD HEALTH IN LAST 12 MONTHS

A number of variables were used to assess child health. A variable representing the overall general health of the child in the previous 12 months was asked of the mother with response options given on a 5-point scale ranging from *excellent* to *poor*. This measure was dichotomised to create a binary variable denoting whether the child had good health *(good, very good, excellent)* or not *(poor, fair)*. The number of health problems the child had in the last 12 months was assessed by asking the mother whether her child had been taken to the GP, health centre, or hospital accident and emergency department for any problems on a list of 13 possible options. A variable denoting the total number of health problems was created by summing the number of problems endorsed by the mother. Binary variables were also created based on whether or not the child had stayed overnight in hospital in the last twelve months for any illness, or required medical attention in the past 12 months for an accident, chest infection, asthma, skin problems, or an ear infection.

LONG TERM CHILD HEALTH

Three binary variables were created based on a) whether the child had any ongoing diagnosed chronic illness other than asthma, b) whether the child was diagnosed with asthma, and c) whether the child had any diagnosed physical disability.

APPROPRIATE FOOD

Mothers were asked how often their child ate grains, dairy, protein, fruit, vegetables, and other foods (including sugars and fats, sweets, crisps, etc.). These were scored as a continuous variable with 1 representing *never*, up to 9, representing *more than six times a day*. A binary variable was created using the continuous measure to reflect whether or not the child had met the dietary requirements for each food category (Department of Health and Children, 2004). For each food group, children needed to consume two to three portions per day or more to meet the guidelines. If participants reported that their child met all of the individual food group guidelines they were coded as meeting all dietary recommendations. The sugars and fats category was reverse scored to indicate that more of these foods were not beneficial. A diet quality score was also calculated. This was a cumulative measure which assigned a value for consumption of each of the food groups (i.e. more was better for protein/vegetables/fruits/dairy/grains; less was better for other foods such as sugars and fats).

CHILDREN SLEEP HABITS

Children's sleep habits were examined using the Children's Sleep Habits Questionnaire (CSHQ; Owens, Spirito, & McGuinn, 2000). The CSHQ is a retrospective, 45-item parent-report questionnaire that has been used to examine sleep behaviour in young children. It includes items relating to a number of key sleep domains that encompass the major presenting clinical sleep complaints in this age group. For this study, a modified 22-item version developed by the NICHD SECCYD study (Vandell & Wolfe, 2000) was used and the response scale was expanded to 5 points (1 = *always*, 5 = *never*). A total sleep disturbance score was computed by summing these individual items. Some items were reverse scored to ensure a higher score was indicative of more disturbed sleep. A number of other sleep related questions were also included which recorded the child's normal wake-up, nap, and bed times. These questions were used to determine whether the child had regular wake-up and bed times, the number of hours the child slept for every day, and the occurrence and length of daily naps.

TOILET TRAINING

Developmental Milestones (Cowen, Work, Wyman, & Jarrell, 1994) is a parent-report measure that looks retrospectively at a child's mastery of twelve developmental milestones. Parents were first asked if their child was completely toilet trained, and then to retrospectively report the age of occurrence. Parents then rated the age of occurrence of toilet training completion relative to other children they knew on a 5-point scale ranging from *much sooner* to *much later*.

WEIGHT AND HEIGHT

The child's current weight and height were measured by the interviewer during the forty-eight month interview. To ensure consistency, the interviewer provided weighing scales and height measurement tools. Weight was measured with the child standing comfortably, arms at their side, looking straight ahead with feet centred on the scales. Shoes and jackets were removed prior to measurement. The child's height was measured with a measuring stick and spirit level. The child was instructed to stand with his/her back to the wall, with heels together and feet at a 45 degree angle to each other. Height was measured as the point where the bottom of the spirit level met the measuring stick. Variables were created for the child's current weight (kgs), height (cms), and BMI scores. A binary variable was created to denote whether the child was overweight based on BMI. Height and weight measurement data were obtained for 71% of the *PFL* sample. It was not always possible for the interviewer to record these measurements as the *PFL* children were not always present while the interviews took place and some children refused to be measured.

2.4.2 Child Health Results

Table 2.2 presents the results comparing the high and low treatment groups on the child health domain.

CHILD HEALTH IN LAST 12 MONTHS

Seven of the eight measures in the child health in last 12 months category were in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the eight measures was not statistically significant.

LONG TERM CHILD HEALTH

One of the three measures in the long term child health category was in the hypothesised direction, and this measure was statistically significant. Twelve percent of children in the high treatment group were reported to have been diagnosed with asthma compared with 25% of the low treatment group ($p<.05$, $d=0.34$). Overall, the step-down test showed that the joint effect of the three measures in this category was not statistically significant.

MEETING DIETARY GUIDELINES

Four out of the five measures in the meeting dietary guidelines category were in the hypothesised direction, one of which was statistically significant. Thirty-eight percent of the high treatment group were found to be meeting the dietary guidelines for vegetables compared with 25% of the low treatment group ($p<.05$, $d=0.29$). The step-down test showed that the joint effect of the five measures in this category was not statistically significant.

SLEEP

Seven of the eight measures in the sleep category were in the hypothesised direction, and two were statistically significant. Children in the high treatment group were reported to sleep for 11.21 hours per day, while the children in the low treatment group were reported to sleep for 10.83 hours ($p<.05$, $d=0.31$). On the Children's Sleep Habits Questionnaire, the high treatment group scored 34.85 in terms of their sleep disturbances, while the low treatment group scored 36.80 ($p<.10$, $d=0.22$). The step-down test showed that the joint effect of the eight measures was not statistically significant.

TOILET TRAINING

One of the three measures in the toilet training category was in the hypothesised direction and statistically significant. Ninety-three percent of children in the high treatment group were reported to be toilet trained, compared to 85% of children in the low treatment group ($p<.10$, $d=0.27$). However, there was one non-hypothesised significant effect, only 32% of high treatment children were reported to finish toilet training earlier than other children, compared to 45% of children in the low treatment group ($d=0.28$).

NON STEP-DOWN MEASURES

Of the eight non step-down measures, six were in the hypothesised direction, one was in the non-hypothesised direction, and there was no difference in the eighth. Two of the measures were statistically significant. Twenty-six percent of children in the high treatment group were reported to be overweight compared to 41% of children in the low treatment group ($p<.10$, $d=0.32$). Counter to hypothesis, children in the high treatment group who were diagnosed with asthma, were diagnosed at an average of 2.51 years, while those in the low treatment group were diagnosed at 1.71 years ($p<.05$, $d=0.78$), yet this result was based on only 25 cases.

Table 2.2 - Results for High and Low Treatment Groups: Child Health

	Variable	*N*	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size *d*
Child Health in Last 12 months										
*	Had skin problems	147	(74/73)	0.07	(0.25)	0.12	(0.33)	ns	ns	0.19
*	Received asthma treatment	147	(74/73)	0.15	(0.36)	0.22	(0.42)	ns	ns	0.18
*	No. of health problems taken to GP/health centre/casualty	147	(74/73)	1.18	(1.03)	1.32	(1.09)	ns	ns	0.13
*	Had chest infection	147	(74/73)	0.24	(0.43)	0.30	(0.46)	ns	ns	0.13
	Child had good health	147	(74/73)	0.95	(0.23)	0.92	(0.28)	ns	ns	0.11
*	Had an ear infection	147	(74/73)	0.19	(0.39)	0.23	(0.43)	ns	ns	0.11
*	Had an accident	147	(74/73)	0.10	(0.30)	0.12	(0.33)	ns	ns	0.09
*	Stayed in hospital for at least one day	147	(74/73)	0.08	(0.27)	0.07	(0.25)	ns	ns	0.05
Long Term Child Health										
*	Has a physical disability	147	(74/73)	0.00	(0.00)	0.00	(0.00)	-	-	-
*	Diagnosed with asthma	147	(74/73)	0.12	(0.33)	0.25	(0.44)	$p<.05$	ns	0.34
*	Has other chronic illness	147	(74/73)	0.07	(0.25)	0.05	(0.23)	ns	ns	0.05
Meeting Dietary Guidelines										
	Vegetables	147	(74/73)	0.38	(0.49)	0.25	(0.43)	$p<.05$	ns	0.29
	Dairy	147	(74/73)	0.70	(0.46)	0.62	(0.49)	ns	ns	0.18
	Protein	147	(74/73)	0.34	(0.48)	0.27	(0.45)	ns	ns	0.14
	Grains	147	(74/73)	0.62	(0.49)	0.60	(0.49)	ns	ns	0.02
	Fruits	147	(74/73)	0.58	(0.50)	0.59	(0.50)	ns	ns	0.02
Sleep										
	Usual amount of sleep each day (hours)	147	(74/73)	11.21	(1.21)	10.83	(1.26)	$p<.05$	ns	0.31
*	Children's Sleep Habits Questionnaire	147	(74/73)	34.85	(8.43)	36.80	(9.37)	$p<.10$	ns	0.22
	Child has a regular weekend wake-up time	147	(74/73)	0.96	(0.20)	0.92	(0.28)	ns	ns	0.17
	Child naps during the day	147	(74/73)	0.12	(0.33)	0.08	(0.28)	ns	ns	0.33
	Child has a regular weekend bedtime	146	(74/72)	0.92	(0.27)	0.89	(0.32)	ns	ns	0.10
	Child has a regular weekday wake-up time	147	(74/73)	0.99	(0.12)	0.97	(0.16)	ns	ns	0.10
	Child has a regular week night bedtime	147	(74/73)	0.93	(0.25)	0.94	(0.23)	ns	ns	0.05
	Length of usual nap (minutes)	15	(9/6)	78.33	(40.93)	90.00	(32.86)	ns	ns	0.13
Toilet Training										
	Is child toilet trained?	147	(74/73)	0.93	(0.25)	0.85	(0.36)	$p<.10$	ns	0.27
	Age child was toilet trained	131	(69/62)	2.60	(0.63)	2.59	(0.52)	ns	ns	0.02
	Finish toilet training earlier than other children	131	(69/62)	0.32	(0.47)	0.45	(0.50)	s~	ns	0.28
Non Step-down Measures										
*	BMI Overweight	104	(53/51)	0.26	(0.45)	0.41	(0.50)	$p<.10$	-	0.32
*	BMI Score	104	(53/51)	16.98	(2.41)	17.06	(1.83)	ns	-	0.03
*	Child's current weight kg's	105	(54/51)	18.11	(2.67)	18.25	(2.83)	ns	-	0.05
	Child's current height cm's	105	(53/52)	103.33	(4.64)	103.15	(4.25)	ns	-	0.04
	Diet Quality Score	147	(74/73)	39.11	(11.25)	36.93	(10.91)	ns	-	0.20
	Meeting dietary guidelines	147	(74/73)	0.09	(0.29)	0.08	(0.28)	ns	-	0.04
*	Age diagnosed with asthma	25	(9/16)	2.51	(0.84)	1.71	(1.18)	s~	-	0.78
*	Daily activities limited by asthma	27	(9/18)	0.22	(0.44)	0.22	(0.43)	ns	-	0.00

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.5 Parenting

Parental behaviours and attitudes play a critical role in child development and in later child outcomes including academic achievement, behaviour, and social and emotional development (Amato & Fowler, 2002). Baumrind (1966) posited three different parenting styles, permissive, authoritarian, and authoritative. Authoritative parenting styles, characterised by a combination of high responsiveness and high control, are most often associated with positive child developmental outcomes (e.g., Baumrind, 1991; Hetherington et al., 1999; Taylor et al., 2004), while permissive and authoritarian styles, associated with low responsiveness and low or high control, are commonly related to negative developmental outcomes (Aunola & Nurmi, 2005; Petito & Cummins, 2000).

Research suggests that children may develop internalising and externalising behaviour problems through early maladaptive parent-child interaction processes (Buke, Loeber, Birmaher, 2002; Hinshaw, 2002). This can happen via a number of potential mechanisms. Attachment theory suggests that infants are biologically predisposed to use their parents as a secure base from which they can explore their world and as a safe haven in which to seek comfort and protection when they are distressed (Bowlby, 1969). Therefore, the quality of early parenting, in terms of sensitivity and responsiveness to the child's needs, is an important contributor to socialisation in the early years (Ainsworth et al., 1978; Bowlby, 1969). Furthermore, coercion theory, which is rooted in a social learning perspective, focuses on ineffective and inconsistent discipline practices and suggests that externalising problems are more likely to emerge if a child receives reinforcement for negative behaviour (Patterson, 1976; Snyder, 1995).

Through these interactions, parents influence the attitudes and behaviours of their children. For example, parents who have a positive attitude to education are more likely to promote positive beliefs about school and education in their child (Hoover-Dempsey & Sandler, 1995). Indeed, parental involvement is one of the strongest predictors of students' academic motivation (Cotton & Wikelund, 1989; US Department of Education, 1994). Research suggests that interventions aimed at increasing responsive parenting can support parents and in turn, facilitate a range of child outcomes (Juffer, Hoksbergen, Riksen-Walraven, & Kohnstamm, 1997; Landry, Smith, & Swank, 2006). As a result, parenting and parent-skills training is an area which many home visiting programmes emphasise in an attempt to impact on child outcomes.

IMPACT OF HOME VISITING INTERVENTIONS ON PARENTING AT FORTY-EIGHT MONTHS

Several home visiting evaluations have reported positive impacts on parenting at forty-eight months (Landry et al., 2008; Landry et al., 2012; Madden, O'Hara, & Levenstein, 1984; Olds et al., 1994; Olds et al., 2004). Each of these studies measured parenting through conducting observations of parent-child interactions, while Olds et al. (1994) also utilised physician reports of parental coping problems. In an evaluation of the Nurse Family Partnership, Olds et al. (1994) reported that the programme had a significant impact on this outcome. However, there was no effect on observation-based ratings of parental warmth, control and involvement. In a later evaluation of the same programme, Olds et al. (2004) reported that mothers who received home visitation from paraprofessional visitors were rated significantly higher on an observation-based measure of sensitive/responsive interactions. However, there was no statistically significant impact on maternal sensitive/responsive interactions for nurse-visited mothers.

The Play and Learn Strategies (PALS) intervention is a programme whereby expectant mothers were randomised into either an experimental group (PALS I) or a control group. These mothers were later re-randomised into either a toddler-preschool phase of the programme (PALS II) or a Developmental Assessment Sessions condition to determine at what time points the programme had the greatest impact. Landry et al. (2008) report that, at 38 months of age, observations of parental verbal encouragement were significantly higher in the group which received the PALS II intervention, while contingent responsiveness and redirecting child focus were significantly higher among the families who received both the PALS I and PALS II treatment. The programme did not have a significant impact on parents' ability to maintain child focus, positive affect of mother with child, verbal scaffolding, or sensitivity. A further evaluation of this programme, which focussed on shared book reading activities, found that those who received both

the PALS I and the PALS II intervention were more likely to provide open prompts to their child when reading than those who did not. However, there was no difference in relation to other aspects of cognitive-linguistic supports or in affective supports (Landry et al., 2012).

In summary, the research to date suggests that home visiting programmes can have a significant positive impact on parenting at forty-eight months. However, the results are mixed and there is no clear area of parenting where home visiting consistently creates improvement between thirty-six and forty-eight months of age. Additional information on the impact of parental attitudes and behaviours, parental stress and parental monitoring of television can be found in the twenty-four month and thirty-six month reports.

2.5.1 Parenting Instruments

PARENTING DAILY HASSLES SCALE

The Parenting Daily Hassles Scale (PDH; Crnic, & Greenberg, 1990) is a 20-item measure of typical everyday events in parenting and parent-child interactions, some of which may make life difficult. It assesses the frequency and intensity of these hassles. The frequency of each event gives an objective marker of how often the event occurs and the intensity or impact score indicates the caregiver's subjective appraisal of how much those events affect or hassle them. The PDH provides two global measures, a Frequency scale (α=0.85) which indicates the frequency of typical hassle events and an Intensity scale (α=0.91) which reflects the parent's subjective appraisal of how much of a hassle she finds the event to be. Two further subscales were also calculated; parenting hassles related to typical parenting tasks or duties a parent may be exposed to (α=0.82), and parenting hassles related to challenging behaviour by a child (α=0.84).

PARENTING STRESS INDEX

The short version of the Parenting Stress Index (PSI; Abidin, 1995) consists of 36 items which are completed by parents. The PSI provides a total score (36 items, α=0.82) and three subscales measuring factors related to parental stress. Difficult child (12 items, α=0.88), which indicates behavioural characteristics of the child as perceived by the mother, parenting distress (12 items, α=0.88), and parent-child dysfunctional interactions (12 items, α=0.92). Mothers were asked to rate how much they agree or disagree with each item on a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. Responses to both the overall stress score and the three subscales were summed to generate representative scores, resulting in a possible scoring range of 36 to 180 for the total stress score, and 12 to 60 for perceptions of child behavioural problems, parenting distress, and parent-child dysfunctional interactions. A binary variable was calculated to represent the proportion of mothers scoring above 90, which is indicative of clinically significant stress levels.

PARENTING STYLES AND DIMENSIONS QUESTIONNAIRE

The Parenting Styles and Dimensions Questionnaire (PSDQ; Robinson et al., 1995) is a 32-item self-report measure that assesses parenting styles in accordance with Baumrind's (1989) typologies of authoritative, authoritarian, and permissive parenting. These parenting typologies are based on parents' relative use of responsiveness and demandingness. Authoritative parenting is considered a positive outcome as it is associated with high levels of responsiveness and demandingness. Authoritarian and permissive parenting styles are considered as negative outcomes. For each item parents rated how often they react to their child in the manner described in each statement on a 5-point scale ranging from *never* to *always*. The PSDQ yields an overall mean score for each of the three categories of parenting style. The mean scores are calculated by summing the responses for the items in each category: Authoritarian (15 items; α=0.79), Authoritative (12 items; α=0.82), and Permissive (5 items; α=0.75). The parenting style with the highest mean determines the respondents' parenting style. In addition, the PSDQ yields three positive authoritative parenting subdomains: connection (α=0.66), regulation (α=0.69), and autonomy (α=0.68); as well as three negative authoritarian parenting subdomains: punitive (α=0.74), hostility (α=0.61), and coercion (α=0.74).

PARENTAL PERCEPTION OF IMPORTANT SCHOOL READINESS TRAITS

Mothers were asked to indicate what types of traits/skills they thought were important for getting children ready for school. Responses to this question were coded into the following domains of school readiness: physical health and wellbeing, social competence, emotional maturity, language and cognitive development, communication skills and general knowledge, and other. The variables thus indicate the proportion of parents who mentioned each of these five particular traits.

PRIMARY SCHOOL PLANNING AND ATTENDANCE

Mothers were asked three questions about primary school education. They were asked if their child had commenced primary school and if so at what age. If the child was not in primary school, they were then asked if their child was currently on a waiting list for primary school.

CHILDREN'S TELEVISION HABITS

Mothers were asked 11 questions in relation to their child's television habits. Participants reported the average duration, in hours and minutes, that the child spends watching television per day, watching videos/DVDs per day, watching television alone per day, and watching television with his/her mother per day. They were also asked to report how long the television is on in their home per day and the shows their child watches. Additionally, mothers were asked whether they limit their child's exposure to television or video watching, and if so, to report the limit, and whether they discussed the shows with their child.

2.5.2 Parenting Results

Table 2.3 presents the results comparing the high and low treatment groups on the parenting domain.

PARENTING DAILY HASSLES SCALE (PDH)

All four Parenting Daily Hassles subscales were in the hypothesised direction. However, none were statistically significant. Furthermore, the step-down test showed that the joint effect of the four subscales was not statistically significant.

PARENTING STRESS INDEX (PSI)

All three of the Parenting Stress Index subscales were in the hypothesised direction. However, none of these reached statistical significance. The step-down test demonstrated that the joint effect of the three subscales was not statistically significant.

PARENTING STYLES AND DIMENSIONS QUESTIONNAIRE (PSDQ)

The three PSDQ scales were in the hypothesised direction, and one indicated a statistically significant difference. Mothers in the high treatment group scored an average of 2.25 on the Permissive Parenting scale, while the low treatment group scored an average of 2.47 ($p<.05$, $d=0.26$). This indicates that mothers in the high treatment group were less likely to engage in behaviours associated with permissive parenting. The results from the step-down analysis indicated that the joint effect of the three PSDQ subscales was not statistically significant.

PSDQ AUTHORITATIVE PARENTING SUBDOMAINS

Of the three PSDQ Authoritative Parenting subdomains, two were in the hypothesised direction; however, none were significantly different. Furthermore, the step-down test showed that the joint effect of the subdomains was not statistically significant.

PSDQ AUTHORITARIAN PARENTING SUBDOMAINS

Two of the three PSDQ Authoritarian Parenting subdomains were in the hypothesised direction. None of these differences reached statistical significance. In addition, the step-down test showed that the joint effect of the three subdomains was not statistically significant.

PARENTAL PERCEPTION OF IMPORTANT SCHOOL READINESS TRAITS

Of the six domains measuring parental perceptions of important school readiness traits, none were in the hypothesised direction and none reached statistical significance. In addition, the step-down test showed that the joint effect of all of the traits was not statistically significant.

TV HABITS

Seven items measured TV habits within the home. Of these, six were in the hypothesised direction while one was statistically significant. Children in the high treatment group spent on average 0.94 hours per day watching TV alone, while those in the low treatment group spent on average 1.2 hours per day engaged in this activity ($p<.10$, $d=0.22$). The step-down test showed that the joint effect of the TV habits measures was not statistically significant.

NON STEP-DOWN MEASURES

Five of the seven non step-down measures were in the hypothesised direction and none were statistically significant.

Table 2.3 - Results for High and Low Treatment Groups: Parenting

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size d
Parenting Daily Hassles (PDH)									
* PDH Parenting Tasks Score	147	(74/73)	11.51	(3.86)	12.38	(4.70)	ns	ns	0.20
* PDH Intensity Scale Score	147	(74/73)	31.58	(11.21)	33.15	(12.01)	ns	ns	0.14
* PDH Frequency Scale Score	147	(74/73)	33.41	(6.72)	34.22	(7.94)	ns	ns	0.11
* PDH Challenging Behaviour Score	147	(74/73)	12.36	(5.15)	12.73	(5.01)	ns	ns	0.07
Parenting Stress Index (PSI)									
* Parental Distress	147	(74/73)	23.92	(8.24)	24.36	(7.41)	ns	ns	0.06
* Difficult Child	147	(74/73)	22.38	(7.50)	22.79	(7.63)	ns	ns	0.06
* Parent-Child Dysfunctional Interactions	147	(74/73)	18.35	(6.10)	18.60	(5.80)	ns	ns	0.04
Parenting Styles and Dimensions Questionnaire (PSDQ)									
* Permissive Parenting	147	(74/73)	2.25	(0.77)	2.47	(0.86)	p<.05	ns	0.26
* Authoritarian Parenting	147	(74/73)	1.56	(0.47)	1.61	(0.49)	ns	ns	0.10
Authoritative Parenting	147	(74/73)	4.07	(0.51)	4.06	(0.59)	ns	ns	0.02
PSDQ Authoritative Parenting Subdomains									
PSDQ Connection	147	(74/73)	4.68	(0.39)	4.65	(0.47)	ns	ns	0.07
PSDQ Regulation	147	(74/73)	3.91	(0.76)	3.89	(0.80)	ns	ns	0.03
PSDQ Autonomy	147	(74/73)	3.61	(0.80)	3.64	(0.79)	ns	ns	0.03
PSDQ Authoritarian Parenting Subdomains									
* PSDQ Punitive	147	(74/73)	1.69	(0.69)	1.81	(0.64)	ns	ns	0.16
* PSDQ Hostility	147	(74/73)	1.62	(0.54)	1.65	(0.67)	ns	ns	0.05
* PSDQ Coercion	147	(74/73)	1.38	(0.48)	1.37	(0.43)	ns	ns	0.00
Parental Perception of Important School Readiness Traits									
Emotional Maturity	136	(69/67)	0.23	(0.43)	0.25	(0.61)	ns	ns	0.04
Other Skills	134	(69/65)	0.10	(0.30)	0.12	(0.33)	ns	ns	0.07
Communication and General Knowledge	137	(70/67)	0.24	(0.69)	0.31	(0.91)	ns	ns	0.09
Physical Health and Wellbeing	136	(70/66)	0.36	(0.48)	0.42	(0.50)	ns	ns	0.14
Social Competence	139	(71/68)	0.66	(0.53)	0.75	(0.53)	ns	ns	0.17
Language and Cognitive Development	138	(71/67)	0.42	(0.77)	0.57	(0.78)	ns	ns	0.19
TV Habits									
* Time spent by child watching TV alone (hours)	146	(74/72)	0.94	(1.09)	1.20	(1.17)	p<.10	ns	0.22
* Time TV is on in the home (hours)	146	(74/72)	8.33	(3.97)	9.11	(4.11)	ns	ns	0.19
Child's TV time limited	147	(74/73)	0.62	(0.49)	0.53	(0.50)	ns	ns	0.18
Mother talks to child about TV	132	(67/65)	0.99	(0.12)	0.95	(0.21)	ns	ns	0.18
Time spent watching TV with child (hours)	132	(67/65)	1.40	(1.20)	1.21	(1.03)	ns	ns	0.17
* Time spent by child watching TV/videos/ DVDs per day (hours)	147	(74/73)	2.44	(1.54)	2.58	(1.55)	ns	ns	0.09
* Maximum TV time allowed per day	85	(46/39)	2.97	(1.52)	2.77	(1.12)	ns	ns	0.15
Non Step-down Measures									
Worried about child's behaviour	147	(74/73)	0.14	(0.34)	0.08	(0.28)	ns	-	0.17
Worried about child's language development	147	(74/73)	0.15	(0.36)	0.16	(0.37)	ns	-	0.04
* PSI Stress cut-off	147	(74/73)	0.09	(0.29)	0.11	(0.31)	ns	-	0.05
* PSI Total Stress Score	147	(74/73)	64.65	(19.23)	65.75	(18.54)	ns	-	0.06
Age started school (months)	8	(5/3)	50.80	(2.59)	49.67	(2.89)	ns	-	0.49
Child has started primary school	147	(74/73)	0.07	(0.25)	0.05	(0.23)	ns	-	0.05
Child is on primary school waiting list	138	(69/69)	0.75	(0.43)	0.81	(0.39)	ns	-	0.14

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '*p*<.01', '*p*<.05' and '*p*<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.6 Home Environment

Socioeconomic inequalities in developmental outcomes result from inequalities in the experiences and environmental conditions children are exposed to (Siddiqui, Irwin, Hetzman, 2007). While the physical, social and cognitive aspects of a child's home environment were measured and discussed in the earlier *PFL* reports, at forty-eight months the home learning environment specifically and children's safety in the home were also assessed.

HOME LEARNING ENVIRONMENT

Higher quality home learning environments have been linked to better outcomes in terms of children's social, behavioural and cognitive development. In a longitudinal study of pre-school education, Sammons et al., (2004) reported that the quality of the home environment had a stronger impact on child outcomes than other factors such as parental education or social class. This finding has been replicated in other studies (Desforges & Abouchaar, 2003; Gutman & Feinstein, 2007; Sylva et al., 2004). These findings suggest that improving the home-learning environment may be an important mechanism through which parenting interventions can be effective, particularly for children from disadvantaged backgrounds. Indeed, research has demonstrated that supporting parents to improve the child's home learning environment can have a significant impact on child outcomes, including school readiness as well as attainment and achievement up to the age of 16 (Desforges & Abouchaar, 2003; Feinstein et al., 2004).

A quality home-learning environment is one in which parents actively engage with their children through both play and learning activities. Increasing a parent's understanding of play and the ability to facilitate learning has predicted positive school readiness factors including independence and creativity (Parker, Boak, Griffin, Ripple, & Peay, 1999). Improvements in home learning environments from thirty-six to fifty-four months have been positively related to the level of maternal education, work-hours and household income (Son & Morrison, 2010; Votruba-Drzal, 2003). Irrespective of socioeconomic status, parents have been observed to engage in various home learning activities equally, though reading activity is an exception (Hartas, 2011).

CHILD SAFETY

Children living in low income, urban households are at an increased risk of fire and scald burns in the home due to substandard housing such as the absence of working smoke alarms and safe hot water temperatures (Gielen et al., 2012). A child-safe physical environment can be protective while facilitating opportunities for children to be independent and explore, allowing them to learn and develop physically, socially and emotionally (Carr, 2006; Edwards et al., 2010). As children develop their motor movements at four years, including learning skills such as walking up and down stairs, they gain more independence (Sheridan, 2004). As they begin to have more control over their environments and daily experiences, home safety becomes an important issue (Scarr & McCartney, 1983). The World Health Organisation highlights the importance of safety in the home, with evidence that children under four years are the most affected by injuries in the home and that the majority of injuries are preventable through environmental modification, education and regulation (Hyder et al., 2009).

IMPACT OF HOME VISITING INTERVENTIONS

Several studies have documented the positive effect of home visiting programmes on the home and family environment of young children including the Queensland Study of nurse visits to at-risk mothers, the evaluation of Healthy Families America Alaska, and the Infant Health and Development Program (Howard & Brooks-Gunn, 2009). Contrary to this, the evaluation of the Comprehensive Child Development Program – which examines a multi-purpose home visiting programme - documented no impact on the home environment (St. Pierre, Layzer, Goodson, & Bernstein, 1999). In a three-site evaluation of the Nurse Family Partnership intervention, contrasting results were reported at forty-eight months in relation to parental smoking activity, provision of appropriate play materials, the HOME inventory and the presence of hazards. The author hypothesises that this was a result of the evaluation design, execution of programme protocol, cohort size and facilitator qualification (Olds, 2006).

2.6.1 Home Environment Instruments

HOME LEARNING ENVIRONMENT

The Home Learning Environment Index (HLE; Melhuish, Phan, Sylva, Sammons, Siraj-Blachford, & Taggart, 2008) is a measure of 7 activities that investigates how often children engage in each activity with their parents: frequency of being read to, going to the library, being taught a sport, dance, or physical activity, playing with letters, being taught letters, being taught numbers, singing songs/poems/rhymes, and painting and drawing. The HLE is designed to be administered as a semi-structured interview and is suitable for children aged 3, 5, and 7. Parents are asked to rank how often someone in the household engages in each of the activities with the child from 0 *'occasionally or less than once a week'* to 7 *'seven times a week/ constantly'*. These scores were then summed to give a total score ranging between 0 and 49, with higher scores indicating increased levels of engagement with learning activities in the home.

CHILD SAFETY

The safety of the home environment was measured using 18 items from the aged one to four year version of the Framingham Safety Survey (FSS; American Academy of Pediatrics, 1991). Items were measured using questions such as, *'Do you leave your child alone at home?'* and *'Are any of your babysitters younger than 13 years?'* A summative score relating to the safety of the physical environment was derived from this instrument.

CHILD EXPOSED TO CIGARETTE SMOKE

Participants were asked whether another person in the house, other than themselves, smoked and a binary variable was created for *yes* and *no*.

SOCIAL WORKER INVOLVEMENT

Participants were asked if there was a social worker working with the family and a binary variable was created for yes and no.

2.6.2 Home Environment Results

Table 2.4 presents the results comparing the high and low treatment groups on the home environment domain.

All four of the measures were in the hypothesised direction and two were statistically significant. Children in the high treatment group were significantly less likely to be exposed to smoking in the home, with 35% of high treatment mothers indicating that another person smoked in the home, compared to 52% of low treatment mothers ($p<.05$, $d=0.35$). In addition, 6% of low treatment mothers reported that they had a social worker working with the family compared to only 1% of high treatment mothers ($p<.10$, $d=0.23$). With means of 31.66 and 31.07 ($d=0.06$) reported by the high and low treatment participants respectively regarding the Home Learning Index, there was almost no difference in the home learning environment across groups.

Table 2.4 - Results for High and Low Treatment Groups: Home and Family Environment

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size d
Non Step Down Measures									
* Child exposed to cigarette smoke	147	(74/73)	0.35	(0.48)	0.52	(0.50)	$p<.05$	-	0.35
* Social worker working with family	146	(74/72)	0.01	(0.12)	0.06	(0.23)	$p<.10$	-	0.23
Home Learning Environment	147	(74/73)	31.66	(10.05)	31.07	(8.94)	ns	-	0.06
Framingham Safety Survey	147	(74/73)	8.05	(0.95)	8.04	(0.95)	ns	-	0.02

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.7 Maternal Health & Wellbeing

Maternal physical and mental health are important determinants of child outcomes. Research has demonstrated that poor maternal physical health is associated with a number of negative child outcomes including increased rates of childhood depression (Osborn, 2007); anxiety disorders and illicit drug use (Lester et al., 2006); social isolation (Vannatta, Grollman, Noll, & Gerhardt, 2008); school absenteeism, somatic complaints, and externalising and internalising behaviour problems (Evans, Keenan, & Shipton, 2007); and poorer long term economic wellbeing (Wagmiller, Lennon, & Kuang, 2008). Furthermore, maternal depression has been linked to behavioural problems and to lower vocabulary scores in children (Brennan, Hammen, Andersen, Bor, Najman, & Williams, 2000), in addition to elevated rates of childhood depression (Downey & Coyne, 1990). The mechanisms through which maternal health impacts on child outcomes are not well known, but it has been proposed that the relationship is mediated through the impact of ill health on parenting (Armistead, Klein, & Forehand, 1995). Research has suggested that parents who are suffering from ill health have lower energy levels and as a result provide less time and attention to their children (Helseth & Ulfsaet, 2009; Mukherjee, Sloper, & Lwein, 2002; White, Mendoza, White, & Bond, 2009).

The impact of parental health and wellbeing on child outcomes is particularly relevant among disadvantaged groups as lower income parents are at increased risk of developing poorer mental health and are more likely to suffer from depression (Shonkoff & Phillips, 2000). Furthermore, research has consistently demonstrated socioeconomic inequalities in risk behaviours related to health such as smoking and lack of physical activity (Barbeau, Krieger, & Soobader, 2004; Najman, Toloo, & Siskind, 2006). Therefore, this is an important area for home visiting programmes to target as it may have an impact on parental behaviour and subsequent child outcomes.

IMPACT OF HOME VISITING INTERVENTIONS ON MATERNAL HEALTH AND WELLBEING AT FORTY-EIGHT MONTHS

At forty-eight months only two home visiting programmes measured parental health: the Nurse Family Partnership and the Childhood Asthma Prevention Study. An early evaluation of the Nurse Family Partnership reported that nurse-visited mothers had significantly fewer pregnancies by the time their child was four years of age and they postponed the birth of their second child by an average of 12 months longer than mothers who did not receive the programme (Olds et al., 1988). This finding was replicated in a later study by Olds et al. (2004), who reported that mothers who received support from nurse visitors had greater intervals between the births of their first and second children. However, nurse-visited families did not display any differences to control families in mental health, mastery scores, or marijuana or alcohol use. In contrast, paraprofessional-visited mothers displayed a greater sense of mastery and higher mental health scores than controls, and were less likely to have a subsequent miscarriage, yet there was no

difference in their rates or timings of subsequent pregnancies or in their use of marijuana or alcohol.

The Childhood Asthma Prevention Study reported that among mothers of four year old children who had been diagnosed with asthma, those who had received a home visiting intervention scored significantly higher on a measure of quality of life. The authors suggest that this may be due to increased coping skills and self-efficacy among these mothers (Klinnert et al., 2007).

The evidence on the impact of home visiting interventions on maternal health and wellbeing at forty-eight months is limited and the results from the available literature are mixed.

2.7.1 Maternal Health & Wellbeing Instruments

MATERNAL HEALTH

Mothers' current health status was assessed using a self-rated report of general health measured on a 5-point scale ranging from *excellent* to *poor*. This measure was dichotomised to create a binary indicator of good health if the participant rated her current health as good, *very good*, or *excellent*, rather than *fair* or *poor*. Participants were also asked how many times they visited the GP in the last 12 months (not including visits for their child), if they were currently pregnant, and if so, whether the pregnancy was planned. If they were not pregnant, they were asked whether they used birth control, and to identify what type from a list. Valid methods of birth control included: I take birth control pills at least sometimes, I take birth control pills regularly and I have my partner use condoms. Participants were also asked if they had been pregnant since the birth of the *PFL* child, and if so, what the outcome had been.

SUBSTANCE USE

Three binary indicators were used to assess whether participants smoked, drank alcohol or took drugs in the past 12 months. For yes responses to the smoking question, participants were asked how many cigarettes they smoked per day, and for a yes response to the alcohol question, participants were asked how often and how much they drank. A binary indicator was calculated indicating whether the participant consumed alcohol above the recommended level or not (that is, more than 14 units of alcohol per week on average). A binge drinking variable was created for participants who reported consuming more than 6 units in a sitting, at least once a week. Whether the participant changed her smoking habits was also calculated based on changes in reported smoking between thirty-six and forty-eight months. Positive numbers indicated a reduction in smoking, whereas negative numbers indicated that the participant was smoking more. The average change was reported.

PSYCHOLOGICAL WELLBEING

Maternal psychological wellbeing was assessed using the Edinburgh Postnatal Depression Scale (EPDS; Cox, et al., 1987) and the WHO-5 Index (WHO, 1998). Self-efficacy was measured using the Pearlin Self-Efficacy Scale (Pearlin & Schooler, 1978) and self-esteem was measured using the Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965).

EDINBURGH POSTNATAL DEPRESSION SCALE

The EPDS is a 10-item (α=0.89) measure designed to identify women who are at risk of depression. Mothers' responses to each question indicated how they had been feeling over the previous week on a 4-point scale. Reverse scoring was applied to some questions. The total score was created by summing each response, with higher scores indicating a greater likelihood of depression. Additionally, a binary variable was created with participants scoring above 10 considered to be at high risk of depression.

WHO-5 INDEX

The WHO-5 Index is a 5-item measure (α=0.90) of subjective wellbeing. Participants were presented with five statements relating to how they had been feeling over the previous fortnight. Participants were asked to select the option which was closest to how they had been feeling during this time on a 6-point scale

ranging from at *no time* to *all of the time*. A total score was calculated by summing all of the responses, giving a range from 0 to 25, with higher scores indicative of better quality of life and wellbeing. A binary variable was created for participants who scored below 13 indicating poor wellbeing.

PEARLIN SELF-EFFICACY SCALE

Maternal self-efficacy was measured using the 7-item mastery subscale from the Pearlin Self Efficacy Scale (Pearlin & Schooler, 1978) and six items of parental self-efficacy from the Abecedarian study (Borkowski, et al., 2001). Thus, participants were presented with 13 items related to how they feel about themselves, their life so far, and becoming a parent, and asked to rate how much they agree or disagree with each item on a scale ranging from zero meaning *strongly disagree* to four signifying *strongly agree*. These measures provided scores on two subdomains including 'mastery' (7 items, α=0.81) or the degree to which the mother feels she has control over things that happen to her and 'parental self-efficacy' (6 items, α=0.72) or the mother's belief that she is able to effectively parent her child/children, as well as a total 'maternal efficacy score' (13 items, α=0.85). The 'maternal efficacy score' was generated by summing the responses to each of the 13 individual items and dividing by 13 to get the average score. All subscales represent the average response to all items within that subscale and range from zero to four with higher scores indicating higher self-efficacy. The 'Pearlin cut-off score' was calculated using the mastery items from the Pearlin only. The cut-off represents the lowest 10th percentile of the whole sample (high treatment, low treatment and comparison group).

ROSENBERG SELF-ESTEEM SCALE

Maternal self-esteem was assessed using the Rosenberg Self Esteem Scale (RSE; Rosenberg, 1965), a 6-item (α=0.84) measure assessing maternal self-esteem on a continuous scale. Mothers were presented with statements about how they feel about themselves and were asked to rate how much they agree or disagree with each statement on a four point Likert scale ranging from zero meaning strongly agree to three representing strongly disagree. Scores were created by summing responses to all items, providing a range of zero to 18 with higher scores representing higher self-esteem.

2.7.2 Maternal Health & Wellbeing Results

Table 2.5 presents the results comparing the high and low treatment groups on the Maternal Health & Wellbeing domain.

MATERNAL PHYSICAL HEALTH & HEALTH BEHAVIOURS IN PAST 12 MONTHS

Of the two measures in this category, one was in the hypothesised direction and statistically significant. Ninety-two percent of participants in the high treatment group reported that they were in good health compared with other women, in comparison to 79% in the low treatment group (p<.05, d=0.36). The step-down test indicated that the joint effect of the measures in this category was statistically significant (p<.05), and this finding was driven by the significant result in the "good health compared with other women" measure.

MATERNAL MENTAL HEALTH

One of the three measures in this category was in the hypothesised direction, while the other two measures were in the non-hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of these measures was not statistically significant.

CURRENT SUBSTANCE USE

Three of the four measures in this category were in the hypothesised direction, and one was statistically significant. Fourteen percent of the high treatment group reported consuming more than 14 units of alcohol in the past week, compared with 25% of the low treatment group (p<.05, d=0.29). The step-down test showed that the joint effect of these measures was not statistically significant.

MATERNAL SELF-EFFICACY

Participants in the high and low treatment groups were equal on both measures in the maternal self-efficacy category, which runs counter to the hypothesis. The step-down test showed that the joint effect of these measures was not statistically significant.

NON STEP-DOWN MEASURES

Five of the nine non step-down measures were in the hypothesised direction, and one was statistically significant. Nineteen percent of participants in the high treatment group reported binge drinking compared with 32% of participants in the low treatment group ($p<.05$, $d=0.29$).

Table 2.5 - Results for High and Low Treatment Groups: Maternal Health & Wellbeing

	Variable	*N*	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size *d*
Maternal Physical Health & Health Behaviours in Past 12 Months										
	Good health compared with other women	147	(74/73)	0.92	(0.27)	0.79	(0.41)	$p<.05$	$p<.05$	0.36
*	No. of GP visits	145	(73/72)	4.21	(5.61)	3.40	(4.38)	ns	ns	0.16
Maternal Mental Health										
*	Edinburgh Postnatal Depression Score for past 7 days	147	(74/73)	6.95	(5.13)	7.33	(5.41)	ns	ns	0.07
	Rosenberg Self-Esteem Scale	147	(74/73)	13.01	(2.91)	13.07	(3.10)	ns	ns	0.02
	WHO-5 Percentage Score	147	(74/73)	61.24	(21.18)	63.45	(23.95)	ns	ns	0.10
Current Substance Use										
*	More than 14 units of alcohol consumed per week	147	(74/73)	0.14	(0.34)	0.25	(0.43)	$p<.05$	ns	0.29
*	Drank alcohol in past 12 months	147	(74/73)	0.88	(0.33)	0.90	(0.30)	ns	ns	0.08
*	Drug use in past 12 months	147	(74/73)	0.03	(0.16)	0.04	(0.20)	ns	ns	0.08
*	Currently a smoker	147	(74/73)	0.49	(0.50)	0.38	(0.49)	ns	ns	0.21
Maternal Self-Efficacy										
	Pearlin Efficacy Score	147	(74/73)	3.20	(0.57)	3.20	(0.56)	ns	ns	0.01
	Pearlin Mastery Score	147	(74/73)	2.97	(0.71)	2.97	(0.68)	ns	ns	0.00
Non Step-down Measures										
*	Binge drinking (> 6 units in any sitting at least once per week)	147	(74/73)	0.19	(0.39)	0.32	(0.47)	$p<.05$	-	0.29
*	Number of cigarettes per day	64	(36/28)	11.83	(5.25)	11.86	(5.21)	ns	-	0.00
	Reduction in smoking between 36 and 48 months (Number of cigarettes)	54	(30/24)	0.40	(4.66)	0.25	(3.60)	ns	-	0.16
*	Edinburgh Postnatal Depression cut-off (10)	147	(74/73)	0.32	(0.47)	0.34	(0.48)	ns	-	0.04
*	Below WHO-5 Score of 13	147	(74/73)	0.28	(0.45)	0.27	(0.45)	ns	-	0.02
	Pearlin Score	147	(74/73)	3.08	(0.58)	3.08	(0.56)	ns	-	0.24
*	Pearlin cut off	147	(74/73)	0.14	(0.34)	0.10	(0.30)	ns	-	0.12
	Been pregnant since birth of *PFL* Child	146	(74/72)	0.30	(0.46)	0.32	(0.47)	ns	-	0.05
	Currently using a valid form of birth control	106	(55/51)	0.49	(0.50)	0.37	(0.49)	ns	-	0.24
	New pregnancy planned	14	(6/8)	0.33	(0.52)	0.50	(0.53)	ns	-	0.34

Notes: '*N*' indicates the sample size. '*M*' indicates the mean. '*SD*' indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant

at the 1%, 5%, and 10% level respectively. 's~'indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.8 Maternal Social Support

Social support for mothers is an important resource for their physical and mental health (Beck, 2001; Berkman, Glass, Brissette, & Seeman, 2000; Kawachi & Berman, 2001; Webster et al., 2011). Social relationships can have both negative and positive effects on maternal stress levels, depression, anxiety and psychological wellbeing (Balaji et al., 2007; Kawachi & Berkam, 2001). Greater levels of maternal social support have also been associated with positive outcomes for children. For example, maternal social network size and quality have been shown to significantly predict higher intelligence test scores in forty-eight month old children (Melson, Ladd, & Hsu, 1993). Another study found that maternal stress and lack of social support were associated with lower intelligence scores in forty-two month old children, but that social support may act to reduce some of the negative effects on child development (Slykerman et al., 2005). While there are some interventions that directly aim to impact on maternal social support, for many programmes it is a secondary or mediating outcome not directly targeted by the programme (Hodnett & Roberts, 2007; Kearney & Deatrick, 2000).

SATISFACTION WITH FATHER'S INVOLVEMENT

Typically, researchers classify a partner's support as 'emotional' or 'practical' support (Dennis & Ross, 2006). Support from a partner and/or satisfaction with the relationship is one of the strongest predictors of maternal wellbeing in first time mothers (Crnic et al., 1983), and is associated with lower levels of stress and depression (Dennis & Ross, 2006; Eberhard–Gran et al., 2002; Hall et al., 1991). Fathers also play a pivotal role within a child's environment and their involvement is associated with a number of child outcomes including cognitive development, emotional wellbeing, and social abilities (Allen & Daly, 2007; Lamb, 2010). However, their influence and the amount of time fathers spend with their children is often meditated by barriers such as work schedules (Robinson & Codbey, 1997) and the extent to which the mother allows the father to be involved in the child's life (Palkovitz, Fagan, & Hull, 2013). For example, a mother who is satisfied with a child's father as a competent carer is more likely to be open to sharing custody with him following parental separation (Juby, Le Bourdais, & Mercil-Graffon, 2005).

IMPACT OF HOME VISITING INTERVENTIONS ON SOCIAL SUPPORT AT FORTY-EIGHT MONTHS

In a study investigating the Nurse Family Partnership programme, Olds et al. (1988) found that at forty-six months, the programme had no impact on the amount of help treatment families received from other family members. This is consistent with results from a Cochrane meta-analysis which found no evidence of intervention effectiveness in the area of social support (Barlow, 2003).

2.8.1 Maternal Social Support Instruments

FATHER/PARTNER SOCIAL SUPPORT

Mothers were asked questions relating to the father's level of involvement in his child's life, and her own level of satisfaction with that level of involvement. Mothers were asked to rate on a four-point scale the amount of support they felt they received from the child's father and/or partner. Responses were dichotomised into binary variables indicating whether or not the participant received *no/little/some support,* or *a lot of support*. If the mother was not in a relationship with the father, she was asked whether he paid child maintenance, and if so, whether this was paid regularly or not. A binary variable was used to assess whether or not the father was involved in the child's life. Participants were asked to rate their level of satisfaction based on 14 questions ($\alpha = .90$) assessing satisfaction in relation to helping with household chores, playing with the child, helping with transportation, helping with childcare, etc. Participants answered *very dissatisfied, somewhat dissatisfied, neither/neutral, somewhat satisfied, very satisfied,* or *he does not help in this way*. The responses to these questions were summed to create a scale representing mother's satisfaction with the father's involvement. This scale ranged from a minimum of 14 to a maximum of 90, with higher scores illustrating greater satisfaction.

SOCIAL SUPPORT

Mothers were asked to rate on a 4-point scale the amount of support they felt they received from their parents, close relatives, friends, and neighbours. Responses were categorised into *no/little/some support,* or *a lot of support*. The responses were used to generate four yes/no binary variables indicating whether or not the participant received a lot of support from her parents, relatives, friends, and neighbours. Participants were also asked how often they met with friends/relatives who did not live with them. A binary variable was created, indicating whether she met with them most days or less frequently.

VOTING BEHAVIOUR

Participants were asked whether they voted in the last general election and in the last local/European elections. Binary variables were calculated indicating whether participants reporting voting or not in each election.

2.8.2 Maternal Social Support Results

Table 2.6 presents the results comparing the high and low treatment groups on the social support domain.

PARTNER SOCIAL SUPPORT

One of the three measures in the Partner Social Support category was in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the three measures was not statistically significant.

SOCIAL SUPPORT

Three of the five measures in the Social Support category were in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the five measures was not statistically significant.

VOTING BEHAVIOUR

Both of the measures in the Voting Behaviour category were statistically significant in the hypothesised direction. Fifty-six percent of the high treatment group reported voting in the last local and European elections, compared with 41% of the low treatment group ($p<.05$, $d=0.30$). Similarly, 61% of the high treatment group reported voting in the last general election, compared with 49% of the low treatment group ($p<.10$, $d=0.23$). The step-down test showed that the joint effect of these measures was statistically significant.

NON STEP-DOWN MEASURES

Only one of the four non step-down measures was in the hypothesised direction and none were statistically significant.

Table 2.6 - Results for High and Low Treatment Groups: Social Support

Variable	*N*	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size *d*
Partner Social Support									
Support from baby's father	139	(69/70)	0.71	(0.46)	0.63	(0.49)	ns	ns	0.17
How often does father have contact with child (daily/not)	145	(73/73)	0.63	(0.49)	0.64	(0.48)	ns	ns	0.03
Support from partner	107	(53/54)	0.83	(0.38)	0.87	(0.34)	ns	ns	0.11
Social Support									
Support from relatives	144	(72/72)	0.38	(0.49)	0.28	(0.45)	ns	ns	0.21
Support from friends	147	(74/73)	0.23	(0.42)	0.15	(0.36)	ns	ns	0.20
Support from neighbours	140	(70/70)	0.09	(0.28)	0.04	(0.20)	ns	ns	0.18
Support from parent	137	(69/68)	0.59	(0.49)	0.62	(0.49)	ns	ns	0.05
Meet friends (most days/less)	147	(74/73)	0.64	(0.48)	0.68	(0.47)	ns	ns	0.11
Voting Behaviour									
Voted in last local elections and european elections	143	(73/70)	0.56	(0.50)	0.41	(0.50)	$p<.05$	$p<.10$	0.30
Voted in last general election	145	(74/71)	0.61	(0.49)	0.49	(0.50)	$p<.10$	$p<.10$	0.23
Non Step-down Measures									
Father is part of child's life	14	(6/8)	1.00	(0.00)	1.00	(0.00)	-	-	-
Child's father pays maintenance	46	(27/19)	0.74	(0.45)	0.58	(0.51)	ns	-	0.35
Satisfaction with partner's overall support	123	(64/59)	4.52	(1.13)	4.61	(0.87)	ns	-	0.09
Child maintenance is paid regularly	31	(20/11)	0.90	(0.31)	1.00	(0.00)	ns	-	0.41

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.9 Childcare

Aided by the introduction of a universal free preschool year (FPY) in 2010, most children in Ireland experience at least one year of preschool or centre-based care before starting formal schooling (Burke, Morris, & McGarrigle, 2012). The advantages of early childhood care and education before starting school are well documented, with the majority of studies reporting higher levels of school readiness for those who attend formal childcare (Burger, 2010). However, the interaction between the type, timing, and quality of childcare on child outcomes is complex (NICHD, 2004; Sylva et al., 2011). Although high quality centre-based childcare is predictive of healthy cognitive development (Sylva et al., 2011), children who experience early and extensive childcare also have a higher incidence of externalising behaviours (Loeb et al., 2007; National Institute of Child Health and Human Development, 2002).

Evidence from the UK-based Effective Provision of Pre-school Education (EPPE) project reports that childcare experience before starting school is globally beneficial for children; in particular, children from disadvantaged backgrounds benefit from good quality childcare (Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). Follow-up EPPE studies show sustained benefits of childcare attendance on cognitive outcomes at age seven (Sammons et al., 2004) and on academic achievement at age 16 (Sylva et al., 2014). Similarly, in a US sample, Belsky et al. (2007) reported sustained benefits on vocabulary ability in middle childhood for children who attended high quality early childcare centres. However, they also found an enduring effect on behaviour problems, whereby children who experienced a greater number of hours of centre-based care continued to have a higher incidence of behavioural problems in middle childhood (Belsky et al., 2007). A longitudinal Australian study found that the early skill advantages accrued through preschool or centre-based care at age four to five years had faded out by middle childhood (Claessens & Garrett, 2014).

Early childhood care and education is sometimes included as an intervention component of home visiting programmes (Howard & Brooks-Gunn, 2009), although childcare use is not usually reported as an outcome in evaluations of home visiting programmes. One evaluation of the Nurse Family Partnership found an unfavourable intervention effect on attendance at preschool, Head Start, or a licensed childcare centre at forty-eight months (Olds et al., 2004). However, there is an absence of comparable research from other home visiting programmes reporting the impact of such programmes on childcare uptake.

2.9.1 Childcare Instruments

Participants were asked if they have used any type of childcare for the *PFL* child, that is, if anyone besides themselves looked after the child for more than 10 hours per week. This was used to create a binary measure indicating whether the child was in any type of childcare. Those who indicated that they used childcare in the last 12 months were then asked to choose what type of childcare they mainly used from a list including the child's *grandparent, parent/friends/other relatives, nanny/child-minder,* or *nursery/crèche*. A binary variable was created indicating whether the participant used formal childcare (nursery/crèche) or not, and whether or not the child's grandparent provided childcare to them. Additionally, participants were asked how many hours per week their child was in childcare, whether they paid for this childcare, and if so how much, as well as what age their child was when he/she first started in this type of childcare. The cost of childcare on an hourly basis was calculated from this information. In addition, participants were asked how satisfied they were with this childcare. A variable was also created to represent whether or not a child attended a childcare centre that had received Síolta accreditation, a quality accreditation measure.

2.9.2 Childcare Results

Table 2.7 presents the results comparing the high and low treatment groups on the childcare domain.

CHILDCARE USE SAMPLE

Three of the seven measures in the Childcare Use category were in the hypothesised direction, however consistent with previous waves, none were statistically significant. The step-down test showed that the joint effect was not statistically significant. Of the participants who reported using childcare, a large majority of participants, 98%, used formal arrangements by forty-eight months, with no reported difference across the high and low treatment groups.

NON STEP-DOWN MEASURES

There was no difference between the two groups in relation to any type of childcare used.

Table 2.7 - Results for High and Low Treatment Groups: Childcare

	Variable	*N*	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size *d*
Childcare Use sample										
	Child attends Síolta accredited centre	117	(59/58)	0.71	(0.46)	0.64	(0.48)	ns	ns	0.16
	Hours per week in childcare	117	(59/58)	17.31	(7.49)	16.26	(6.42)	ns	ns	0.15
	Satisfaction with childcare	118	(59/59)	0.93	(0.25)	0.91	(0.28)	ns	ns	0.06
*	Uses grandmother care	119	(60/59)	0.02	(0.13)	0.02	(0.13)	ns	ns	0.00
	Uses formal childcare	119	(60/59)	0.98	(0.13)	0.98	(0.13)	ns	ns	0.00
	Age started childcare	117	(59/58)	30.02	(12.35)	31.78	(13.63)	ns	ns	0.14
	Childcare cost per hour	39	(21/18)	1.53	(0.81)	1.87	(1.61)	ns	ns	0.27
Non Step-down Measures										
	Uses any type of childcare	147	(74/73)	0.81	(0.39)	0.81	(0.40)	ns	-	0.01

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '*p*<.01', '*p*<.05' and '*p*<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.10 Household Factors & SES

Family social resources include factors such as parental education and intra-familial relations, while economic resources include wealth, occupational status, and dwelling conditions (Siddiqui et al., 2007). These types of family resources are considered to provide the most powerful explanation for differences in children's early development and wellbeing across societies (Siddiqui, Irwin, & Hertzman, 2007). Research has identified SES disparities in birth outcomes (Blumenshine, Egerter, Barclay, Cubbin, & Braveman, 2010), as well as in cognitive outcomes during the transition to school (Burkam et al., 2004). One large longitudinal study found that maternal education and family income were the most salient risk factors for cognitive delay at forty-eight months (Hillemeier, Morgan, Farkas, & Maczuga, 2011).

IMPACT OF HOME VISITING INTERVENTIONS ON HOUSEHOLD FACTORS AND SES AT FORTY-EIGHT MONTHS

Two studies report on household factors and SES between thirty-six and forty-eight months, and the results are mixed. In 1988, an evaluation of the Nurse Family Partnership programme by Olds et al. found that in contrast to those in the control group, nurse-visited mothers returned to school more rapidly and showed an increase in the number of months they were employed. In 2004, Olds et al. reporting on the Denver, Colorado trial, found that two years after the programme had finished those who received visits from paraprofessionals were less likely to be married or cohabiting than the control group and had greater participation in the workforce. However, they found no effects in the amount of welfare services each group received or in educational achievement at this time point.

2.10.1 Household Factors & SES Instruments

HOUSEHOLD COMPOSITION, LONE PARENT STATUS, AND SIBLINGS

Participants were asked several questions related to their household composition including how many people lived in the household, how many siblings the child had, and whether or not the child's grandparent lived in the household. Additionally, the participant reported her current relationship status. This information was used to generate two binary indicators, the first denoting whether or not the participant was currently in a relationship (*married, cohabitating,* or *boyfriend*) and the second denoting whether or not the participant was married. Furthermore, participants were asked if their current partner was the child's father and if this was the same partner they had when the child was thirty-six months old.

MATERNAL AND PATERNAL EMPLOYMENT

Several questions assessed the current work status of both the mother and the father. If there had been a change in work status since the thirty-six month interview, participants were asked to select their current work status from a list of options including *currently in paid work, in work but on leave, unemployed, student, looking after home/family, retired, not able to work due to disability/sickness, paid training,* or *unpaid training*. Responses to this question were used to create three binary variables, representing the proportion of mothers and fathers in paid work (which included paid training) versus not in paid work, the proportion of mothers and fathers currently unemployed, and the proportion of mothers who are currently looking after the family. Unemployed individuals were asked for how many months they had been without paid work. A binary variable denoting long term unemployment (greater than 12 months) was created. Participants also reported on whether they worked in full or part-time employment and the approximate annual income of both parents. Separate variables were created for the annual wage of part-time and full-time mothers.

FAMILY FINANCES

Participants' perceptions of financial difficulty were assessed by asking them to rate their level of satisfaction with their financial situation on a 5-point scale, ranging from *very dissatisfied* to *very satisfied.* Responses to this variable were used to generate a binary variable indicating whether the participants were satisfied with their financial situation (*very satisfied, sort of satisfied*) or not (*very dissatisfied, sort of dissatisfied, mixed feelings*). Participants were asked to rate how often they worried about their current financial situation on a 5-point scale, from *almost never* to *almost all of the time*. A binary variable was created indicating whether they worried about their financial situation (*almost all of the time, often*) or not (*almost never, once in a while, sometimes*). Participants were also asked how many people were being supported by their total household income and to predict how they thought their financial situation would change in the next twelve months, and a binary variable was created indicating whether they expected it to get better or worse.

Participants were asked whether or not they saved money on a regular basis, and were also asked for a detailed account of any social welfare payments currently received by any household members, from a list of 39 potential payments. Four binary variables were subsequently created; whether anyone in the household received any social welfare payments, whether anyone had a medical card, received one parent benefit or unemployment benefit. Participants also stated the household's weekly income from all sources by selecting from a scale where the lowest range was less than €50 and the highest was €1500 or more. As households differ in the number of people and composition, a variable representing the household equivalised weekly income was created. A weight of 1 was assigned to the first adult in the household, 0.66 to each subsequent adult (aged 14+ years) and 0.33 to each child (aged less than 14 years). The sum of the weights in each household gives the household's equivalised size – the size of the household in adult equivalents. The household equivalised weekly income is the reported household weekly income divided by the equivalised size.

MATERNAL EDUCATION

Participants were asked about their current participation in education. A binary variable was created to represent whether or not they were still in receipt of education.

DOMESTIC RISK

Participants were presented with a list of eight potential domestic risks and asked to indicate if any of these factors had been an issue for anyone in their family. These included separation, parenting problems, domestic violence, abuse, suicidal thoughts, mental health issues, addiction, and other risks. There was also the option of 'no risks', bringing the total number of possible responses to nine. A total number of domestic risks score was also calculated by summing the number of risks each participant indicated.

2.10.2 Household Factors & SES Results

Table 2.8 presents the results comparing the high and low treatment groups on the Household Factors and SES domain.

HOUSEHOLD FACTORS

Two of the five household factors were in the hypothesised direction, while three were in the non-hypothesised direction. There was one significant difference in the non-hypothesised direction: 27% of the high treatment group were residing with a grandparent compared to 18% of the low treatment group ($p<.10$, $d=22$). The step-down test showed that the joint effect of these factors was not significant.

MATERNAL EMPLOYMENT

Three of the four measures in the maternal employment category were in the hypothesised direction and one was in the non-hypothesised direction. None were significant, and the step-down test showed that the joint effect of this category was not statistically significant.

PATERNAL EMPLOYMENT

One of the three measures in the paternal employment category was in the hypothesised direction, while the other two were in the non-hypothesised direction. None were statistically significant, and the step-down test showed that the joint effect of this category was not statistically significant.

FINANCES

In the finances category, five of the 10 measures were in the hypothesised direction and five were in the non-hypothesised direction. Two of the non-hypothesised findings were significant: 84% of participants in the high treatment group were in possession of a medical card compared to 71% of the low treatment group ($p<.05$, $d=0.30$), and within high treatment households, 38% had a household member who was in receipt of unemployment benefit compared to 23% of low treatment households ($p<.05$, $d=0.32$). The step-down test showed that the joint effect of this category was not statistically significant.

DOMESTIC RISKS

Six of the ten measures in the domestic risks category were in the hypothesised direction, and two of these measures were statistically significant. Three percent of high treatment participants reported having mental health issues within their family compared to 12% of the low treatment group ($p<.05$; $d=0.37$). None of the high treatment group reported Other risks within their family compared to 3% of the low treatment group ($p<.10$, $d=0.24$). One of the findings in the non-hypothesised direction was significant: 4% of participants in the high treatment group reported issues concerning suicidal thoughts in their family compared to 0% of the low treatment group ($p<.10$, $d=0.29$). Overall, the step-down test showed that the joint effect of these 10 measures was significant ($p<.10$, $d=0.37$). This finding was driven by the first measure, mental health issues.

NON STEP-DOWN MEASURES

Five of the nine non step-down measures were in the hypothesised direction. None of the findings reached statistical significance.

Table 2.8 - Results for High and Low Treatment Groups: Household Factors & SES

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size d
Household Factors										
	Household size	147	(74/73)	4.72	(1.82)	4.55	(1.59)	ns	ns	0.10
	Married	147	(74/73)	0.23	(0.42)	0.21	(0.41)	ns	ns	0.06
*	Number of siblings	135	(66/69)	1.41	(1.14)	1.39	(1.25)	ns	ns	0.01
	Has a partner	147	(74/73)	0.73	(0.45)	0.74	(0.44)	ns	ns	0.02
*	Resides with grandparent	147	(74/73)	0.27	(0.45)	0.18	(0.39)	s~	ns	0.22
Maternal Employment										
	Looking after the family/home	133	(67/66)	0.24	(0.43)	0.18	(0.39)	ns	ns	0.14
*	Mother unemployed	132	(66/66)	0.24	(0.43)	0.30	(0.46)	ns	ns	0.14
	Mother in paid employment	132	(66/66)	0.41	(0.50)	0.39	(0.49)	ns	ns	0.03
	Mother improvement in work status (between 36 and 48 months)	129	(65/64)	0.06	(0.24)	0.09	(0.29)	ns	ns	0.12
Paternal Employment										
	Father improvement in work status (between 36 and 48 months)	79	(39/40)	0.13	(0.34)	0.05	(0.22)	ns	ns	0.28
*	Father unemployed	85	(45/40)	0.27	(0.45)	0.23	(0.42)	ns	ns	0.10
	Father in paid employment	88	(47/41)	0.62	(0.49)	0.71	(0.46)	ns	ns	0.19
Finances										
	Household financial situation over the next 12 months	147	(74/73)	0.86	(0.34)	0.82	(0.39)	ns	ns	0.12
	Household current financial situation compared to 12 months ago	147	(74/73)	0.72	(0.45)	0.67	(0.47)	ns	ns	0.10
	Saves regularly	147	(74/73)	0.47	(0.50)	0.42	(0.50)	ns	ns	0.10
	Equivalised weekly household income	141	(71/70)	221.76	(109.00)	214.85	(102.63)	ns	ns	0.07
*	In receipt of One Parent Family benefit	147	(74/73)	0.39	(0.49)	0.40	(0.49)	ns	ns	0.01
	Satisfied with household financial situation	147	(74/73)	0.35	(0.48)	0.38	(0.49)	ns	ns	0.07
*	Number of people supported by family income	147	(74/73)	4.70	(1.77)	4.53	(1.63)	ns	ns	0.10
*	Often worried about financial matters	147	(74/73)	0.46	(0.50)	0.38	(0.49)	ns	ns	0.15
*	Medical Card	147	(74/73)	0.84	(0.37)	0.71	(0.46)	s~	ns	0.30
*	Household Unemployment Benefit	147	(74/73)	0.38	(0.49)	0.23	(0.43)	s~	ns	0.32

Continued On Next Page...

Continued...

Variable	*N*	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size *d*
Domestic Risks									
* Mental health issues	147	(74/73)	0.03	(0.16)	0.12	(0.33)	$p<.05$	$p<.10$	0.37
* Other risks	147	(74/73)	0.00	(0.00)	0.03	(0.16)	$p<.10$	ns	0.24
No risks	147	(74/73)	0.69	(0.47)	0.59	(0.50)	ns	ns	0.21
* Abuse	147	(74/73)	0.00	(0.00)	0.01	(0.12)	ns	ns	0.17
* Separation	147	(74/73)	0.03	(0.16)	0.05	(0.23)	ns	ns	0.14
* Death/Bereavement	147	(74/73)	0.20	(0.40)	0.22	(0.42)	ns	ns	0.04
* Addiction	147	(74/73)	0.04	(0.20)	0.04	(0.20)	ns	ns	0.00
* Parenting problems	147	(74/73)	0.04	(0.20)	0.03	(0.16)	ns	ns	0.07
* Domestic violence	147	(74/73)	0.01	(0.12)	0.00	(0.00)	ns	ns	0.16
* Suicidal thoughts	147	(74/73)	0.04	(0.20)	0.00	(0.00)	s~	ns	0.29
Non Step-down Measures									
Father's annual wage	42	(20/22)	32,942	(35,958)	24,218	(12,627)	ns	-	0.34
* Mother long-term unemployed	131	(65/66)	0.11	(0.31)	0.15	(0.36)	ns	-	0.13
Partner is the child's father	108	(54/54)	0.89	(0.32)	0.85	(0.36)	ns	-	0.11
Mother still in education	147	(74/73)	0.15	(0.36)	0.12	(0.33)	ns	-	0.07
* Father long-term unemployed	84	(44/40)	0.11	(0.32)	0.13	(0.33)	ns	-	0.04
Same partner as 12 months ago	108	(54/54)	0.93	(0.26)	0.93	(0.26)	ns	-	0.00
Mother in part-time employment	52	(26/26)	0.65	(0.49)	0.69	(0.47)	ns	-	0.08
Mother's annual wage (full-time only)	17	(9/8)	22,356	(6855)	23,503	(7709)	ns	-	0.17
Mother's annual wage (part-time only)	34	(17/17)	11,223	(5962)	13,332	(9661)	ns	-	0.27

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.11 Main Results Summary

CHILD DEVELOPMENT

Children in the high and low treatment groups differed significantly on several of the child development domains.

- Children in the high treatment group were more likely to display developmentally appropriate fine motor skills than children in the low treatment group.
- Children in the high treatment group were also less likely to score above the cut-off for difficulties in fine motor development.
- None of the children in the high treatment group were rated on the CBCL as having externalising problems at the clinical level. High treatment children were also less likely to have iinternalising problems at the clinical level.
- Children in the high treatment group scored higher on the DP-3 scale for cognitive development, and were more likely to be above the cut-off for average development, than children in the low treatment group.

CHILD HEALTH

Children in the high and low treatment groups differed significantly on several of the child health domains.

- Children in the high treatment group were less likely to have been diagnosed with asthma than children in the low treatment group. However, counter to hypothesis, they were also more typically diagnosed later than those in the low treatment group.
- Children in the high treatment group were more likely to consume the recommended daily amount of vegetables than those in the low treatment group, and were less likely to be overweight according to their BMI.
- Regarding sleep, high treatment children reportedly slept for longer each day and had fewer reported sleep problems than their low treatment counterparts.
- High treatment children were more likely to be toilet trained than low treatment children. However counter to hypothesis, of those who had been toilet trained, low treatment children reportedly finished toilet training at a younger age when compared to their peers.

PARENTING

Mothers in the high and low treatment groups did not differ significantly across many of the parenting domains, including Parenting Daily Hassles, the Parenting Stress Index, parental perceptions of important school readiness traits, or maternal attitudes towards education. They differed significantly on two measured aspects of parenting:

- Mothers in the high treatment group reported fewer permissive parenting behaviours than those in the low treatment group.
- Mothers in the high treatment group reported that their children spent less time watching TV alone than those in the low treatment group.

HOME AND FAMILY ENVIRONMENT

High treatment group participants differed from low treatment participants on two of the four measures of the home environment:

- High treatment children were less likely to be exposed to cigarette smoke at home than low treatment children.
- High treatment families were less likely to be working with a social worker than low treatment families.

MATERNAL HEALTH AND WELLBEING

High treatment mothers differed from low treatment mothers on a small number of measures:

- High treatment mothers were more likely than low treatment mothers to report that they were in good health compared to other women.
- High treatment mothers were less likely to report than they consumed more than 14 units of alcohol per week, and were less likely to report binge drinking, than low treatment mothers.

SOCIAL SUPPORT

High and low treatment mothers differed in one area of social support: high treatment mothers were more likely to report having voted in the last local, European and general elections than low treatment mothers.

CHILDCARE

There was no difference between the high and low treatment groups on any of the childcare measures.

HOUSEHOLD FACTORS AND SES

High and low treatment groups differed on a number of measures of household factors and SES, sometimes in the non-hypothesised direction:

- In the non-hypothesised direction, high treatment mothers were more likely to report living with a grandparent, having a medical card, and being in receipt of household unemployment benefit than low treatment mothers. High treatment mothers were also more likely to report that someone in their household has an issue with suicidal thoughts than low treatment mothers.
- As hypothesised, high treatment mothers reported fewer mental health issues in their family and fewer risks in the "other" category than low treatment mothers.

SUMMARY

Overall, 191 outcome measures were assessed at forty-eight months. Of these one-tailed tests, 123 (64%) were in the hypothesised direction such that the high treatment group had better outcomes than the low treatment group, and 23 (12%) of these domains were statistically significant. These differences were found across all domains except childcare. Fifty-four (28%) of the measures were in the non-hypothesised direction, such that the low treatment group had better outcomes than the high treatment group, and six (3%) of these were statistically significant. These differences were found in the domains of child health and household factors and SES. In 7% of cases, there were no absolute differences between the high and low treatment groups. Of the 32 step-down categories tested, four were significant in the hypothesised direction, representing 13% of all categories, including child behaviour cut-off scores, maternal physical health and health behaviours in the past 12 months, voting, and domestic risks. None of the step-down categories were significant in the non-hypothesised direction.

Chapter Three

Implementation Analysis

Experimental evaluations of early childhood programmes are considered the optimal means of identifying whether a programme has a causal impact on the participating families. However, deviations from the programme protocol can compromise the evaluation and bias the results. The issues of attrition, engagement, and contamination in home visiting programmes, and the implications for evaluations of such programmes are discussed in detail in Chapter 4 of '*Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at Six Months'[4]. This chapter describes and analyses *PFL* implementation practices regarding attrition, engagement, misreporting of participant responses, and potential contamination between programme intake and when the *PFL* child was forty-eight months of age.

3.1 Participant Attrition up to Forty-Eight Months of Age

Attrition occurs when participants withdraw from a programme before its completion. It is important to investigate the extent of programme attrition from *PFL* as the existence of systematic attrition may break the key rationale underlying the randomisation process and lead to biased results. In a review of home visitation outcomes for children and parents, Gomby (2005) stated that 40% of families invited to participate in these programmes chose to defer or not take part in it, while for those who enrolled, between 20% and 80% exited the programme prematurely. Specific characteristics – such as mother's age, level of education, fluency in English, presence of partner, family size, and ethnicity – may predict the likelihood of participant attrition, however the results are highly varied and are sometimes conflicting (Daro & Harding, 1999; Holland, Christensen, Shone, Kearney, & Kitzman, 2014; Roggman, Cook, Peterson, & Raikes, 2008; Wagner, Spiker, Inman Linn, & Hernandez, 2003). In terms of the programme itself, attrition may increase if parental expectations are not met, or if they consider the content irrelevant (Holland et al., 2014; Roggman, et al., 2008). This section investigates the level and determinants of attrition in the *PFL* sample between baseline and the forty-eight month survey.

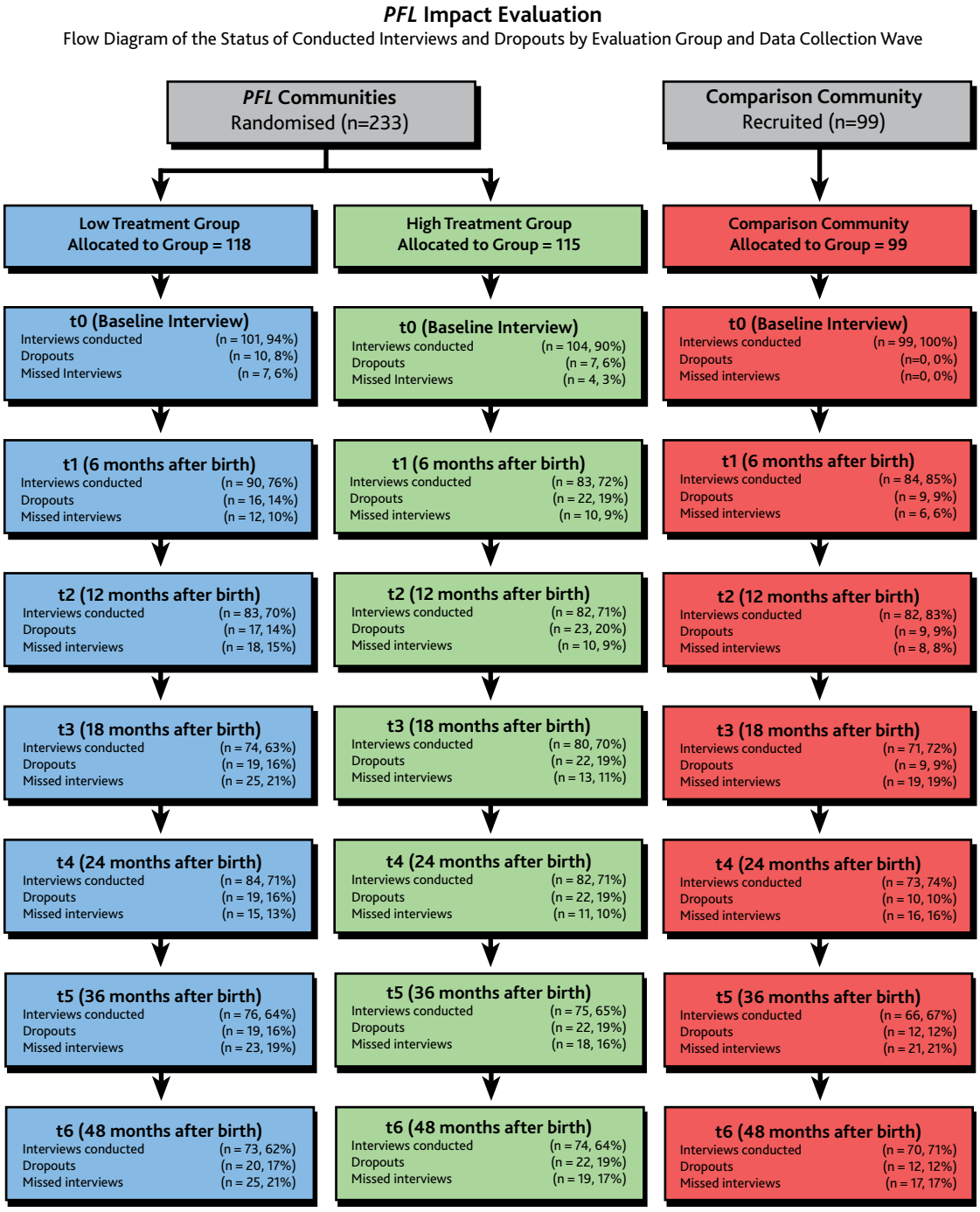
3.1.1 Attrition/Disengagement in *PFL*

The Consort Diagram (Figure 3.1) describes the progression of the participants between programme entry and forty-eight months. In total, 217 forty-eight month interviews (n*High* = 74; n*Low* = 73; n*LFP* = 70) were completed. These 217 participants represent 65.4% of the original sample recruited into the study (n*High* = 115; n*Low* = 118; n*LFP* = 99). The forty-eight month completion rate was very similar for the high (64%) and low (62%) treatment groups, and slightly higher for the comparison group (71%).

Dropout participants are defined as those who actively told the *PFL* programme staff or the evaluation team that they wanted to leave the programme. On average, 16% of the sample were classified as official 'dropouts' between baseline and forty-eight months, with the highest dropout rate experienced among the high treatment group at 19%, while the low treatment group experienced a dropout rate of 17%, and 12% of the comparison group dropped out before forty-eight months. The dropout rate between thirty-six and forty-eight months was minimal however. None of the high treatment group or the comparison group dropped out during this period, and one member of the low treatment group dropped out between thirty-six and forty-eight months.

In addition to those who dropped out, 18% of the sample did not complete a forty-eight month interview as either the interview could not be scheduled at a suitable time during the appropriate interview window, or the participants disengaged from the study. Disengaged participants are those who did not respond to repeated attempts by the evaluation team to be contacted or declined to be interviewed. The rates across the high and low treatment groups were 17% and 21% respectively, and the corresponding rate for the comparison group was 17%. The level of dropout since baseline is higher among the high treatment group and the level of disengagement is higher among the low treatment group, however the total level of attrition/disengagement is quite similar across both groups (*High* = 36%; *Low* = 38%) from randomisation to the forty-eight month interview, with the majority of attrition/disengagement occurring prior to the six month interview.

[4]Please see the following website under publications http://geary.ucd.ie/.



Note: Dropout participants include both voluntary and involuntary dropouts.

Figure 3.1 - Forty-Eight Month Consort Diagram

3.1.2 Analysis of Attrition/Disengagement Before Forty-Eight Months

It is important to examine whether the attrition and disengagement/missed interviews has led to systematic differences between the high and low treatment groups which may bias the outcome results. The analysis below compares the baseline characteristics of participants who completed a forty-eight month interview to those who did not complete a forty-eight month interview. Thus, the analysis of 'attritors' includes those who have officially dropped out of the programme between baseline and forty-eight months and

those who missed the forty-eight month interview during the appropriate time window and/or disengaged from the programme during this period. These baseline characteristics were chosen based on the literature presented in '*Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at Six Months'. The high treatment group, low treatment group, and comparison group are analysed separately.

Table 3.1 reports the baseline characteristics of the high treatment group by attrition status and tests for significant differences between the attrition/disengaged sample and the non-attrition/engaged sample based on each characteristic. It shows that of the 23 maternal characteristics examined, statistically significant differences were found for four baseline measures. Specifically, high treatment group mothers who did not complete a forty-eight month interview were younger, less likely to be employed, and had lower cognitive resources on average, than mothers who completed a forty-eight month interview. However, mothers who completed a forty-eight month interview were more likely to have drank alcohol during pregnancy.

Table 3.1 - Comparison of Baseline Characteristics between Attrition/Disengaged and Non-Attrition/ Engaged Sample: High Treatment Group

Variables	Attrition/Disengaged			Non-Attrition/Engaged			Individual Test
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*p*
Weeks in pregnancy at programme entry	31	22.10	8.11	73	21.37	7.78	ns
Mother's age	31	23.42	5.40	73	26.33	5.85	$p<.05$
Partnered	31	0.68	0.48	73	0.82	0.39	ns
Married	31	0.10	0.30	73	0.16	0.37	ns
Living with parent(s)	31	0.65	0.49	73	0.53	0.50	ns
First time mother	31	0.65	0.49	73	0.50	0.50	ns
Low education	31	0.39	0.50	73	0.32	0.47	ns
Mother employed	31	0.19	0.40	73	0.44	0.50	$p<.05$
Saves regularly	31	0.45	0.51	73	0.48	0.50	ns
Social housing	30	0.60	0.50	73	0.53	0.50	ns
Cognitive resources (WASI)	31	78.35	11.24	73	83.63	12.49	$p<.05$
Vulnerable attachment (VASQ)	31	18.81	3.70	73	18.00	3.79	ns
Self-efficacy (Pearlin)	31	2.68	0.70	73	2.81	0.60	ns
Self-esteem (Rosenberg)	31	12.71	2.91	73	12.86	2.62	ns
Knowledge of infant development (KIDI)	31	71.06	9.18	73	72.76	6.82	ns
Positive parenting attitudes (AAPI)	31	5.32	1.74	73	5.22	1.21	ns
Physical health condition	31	0.74	0.44	73	0.75	0.43	ns
Mental health condition	31	0.32	0.48	73	0.26	0.44	ns
Smoking during pregnancy	31	0.55	0.51	73	0.49	0.50	ns
Drinking during pregnancy	31	0.10	0.30	73	0.32	0.47	$p<.05$
Drug ever used	31	0.10	0.30	73	0.15	0.36	ns

Note: N= sample size, M=mean, *SD*=standard deviation. *p*-values were obtained from two-sided *t* tests based on permutation testing with 1000 replications. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

Table 3.2 reports the baseline characteristics of the low treatment group by attrition status and tests for significant differences between the attrition/disengaged sample and the non-attrition/engaged sample based on each characteristic. It shows that of the 23 maternal characteristics examined, three statistically significant differences were found. Specifically, low treatment mothers who did not complete their forty-eight month interview were more likely to be first time mothers, were more likely to have low levels of education, and were also younger.

Table 3.2 - Comparison of Baseline Characteristics between Attrition/Disengaged and Non-Attrition/Engaged Sample: Low Treatment Group

Variables	Attrition/Disengaged			Non-Attrition/Engaged			Individual Test
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*p*
Weeks in pregnancy at programme entry	29	20.59	7.53	72	21.64	6.73	ns
Mother's age	29	23.66	5.95	72	25.96	5.91	$p<.10$
Partnered	29	0.90	0.31	72	0.82	0.38	ns
Married	29	0.24	0.30	72	0.15	0.36	ns
Living with parent(s)	29	0.55	0.51	72	0.43	0.50	ns
First time mother	29	0.66	0.48	72	0.43	0.50	$p<.10$
Low education	29	0.55	0.51	72	0.33	0.47	$p<.10$
Mother employed	29	0.34	0.48	72	0.42	0.50	ns
Saves regularly	29	0.52	0.51	72	0.51	0.50	ns
Social housing	29	0.62	0.49	72	0.53	0.50	ns
Cognitive resources (WASI)	29	80.86	13.05	72	80.93	12.90	ns
Vulnerable attachment (VASQ)	29	17.97	3.84	72	17.76	4.06	ns
Self-efficacy (Pearlin)	29	2.85	0.54	72	2.89	0.63	ns
Self-esteem (Rosenberg)	29	12.76	2.63	72	12.79	2.96	ns
Knowledge of child development (KIDI)	29	67.83	7.94	72	70.62	8.20	ns
Positive parenting attitudes (AAPI)	29	4.97	1.60	72	5.18	1.35	ns
Physical health condition	29	0.55	0.51	72	0.65	0.48	ns
Mental health condition	29	0.14	0.35	72	0.28	0.45	ns
Smoking during pregnancy	29	0.48	0.51	72	0.47	0.50	ns
Drinking during pregnancy	29	0.24	0.44	72	0.28	0.45	ns
Drug ever used	29	0.24	0.44	72	0.11	0.32	ns

Note: N= sample size, M=mean, *SD*=standard deviation. *p*-values were obtained from two-sided *t* tests based on permutation testing with 1000 replications. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

Table 3.3 reports the baseline characteristics of the comparison group by attrition status and tests for significant differences between the attrition/disengaged sample and the non-attrition/engaged sample based on each characteristic. It shows that of the 23 maternal characteristics examined, one statistically significant difference was found. Specifically, comparison group mothers who did not complete their forty-eight month interview had lower levels of self-efficacy at baseline.

Table 3.3 - Comparison of Baseline Characteristics between Attrition/Disengaged and Non-Attrition/Engaged Sample: Comparison Group

Variables	Attrition/Disengaged			Non-Attrition/Engaged			Individual Test
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*p*
Weeks in pregnancy at programme entry	29	26.66	5.69	70	24.54	6.52	ns
Mother's age	29	26.21	5.51	70	27.73	6.46	ns
Partnered	29	0.76	0.44	70	0.91	0.28	ns
Married	29	0.21	0.41	70	0.17	0.38	ns
Living with parent(s)	29	0.38	0.49	70	0.30	0.46	ns
First time mother	29	0.38	0.49	70	0.43	0.50	ns
Low education	29	0.24	0.44	70	0.26	0.44	ns
Mother employed	28	0.32	0.48	69	0.48	0.50	ns
Saves regularly	29	0.52	0.51	70	0.58	0.50	ns
Social housing	29	0.46	0.51	70	0.41	0.50	ns
Cognitive resources (WASI)	29	85.24	16.44	70	88.89	12.90	ns
Vulnerable attachment (VASQ)	29	16.76	3.84	70	16.97	3.51	ns
Self-efficacy (Pearlin)	29	2.59	0.85	70	2.96	0.60	$p<.05$
Self-esteem (Rosenberg)	29	12.48	3.48	70	13.37	2.91	ns
Knowledge of child development (KIDI)	29	72.17	8.25	70	73.22	8.92	ns
Positive parenting attitudes (AAPI)	29	5.44	1.56	70	5.83	1.32	ns
Physical health condition	29	0.59	0.50	70	0.70	0.46	ns
Mental health condition	28	0.32	0.48	70	0.39	0.49	ns
Smoking during pregnancy	29	0.38	0.63	70	0.33	0.47	ns
Drinking during pregnancy	29	0.31	0.47	70	0.30	0.46	ns
Drug ever used	29	0.21	0.41	70	0.13	0.34	ns

Note: N= sample size, M=mean, *SD*=standard deviation. *p*-values were obtained from two-sided *t* tests based on permutation testing with 1000 replications. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

3.1.3 Inverse Probability Weighted Results

As the previous analysis highlighted some differences between participants who completed a forty-eight month interview and those who did not, this may be a potential source of bias regarding the main results presented in Chapter 2, especially as there was some evidence of different factors influencing attrition within the high and low treatment groups.

In order to account for any potential bias due to differential attrition, the robustness of the outcome analyses were tested using an inverse probability weighting (IPW) technique. This involved three steps. First, 47 permutation tests were estimated to more thoroughly examine the individual baseline factors associated with participation in the forty-eight month interview for the high and low treatment groups separately. Factors which were statistically significant (at the 10% level in a two-tailed test), had no missing data, and were not collinear with any other included factors were retained. For the high treatment group, six significant baseline characteristics were used: maternal age, maternal cognitive resources, maternal employment, drinking during pregnancy, Consideration of Future Consequences (CFC; Strathman, Gleicher, Boninger, & Edwards, 1994), and maternal satisfaction with neighborhood. For the low treatment group, eight significant baseline variables were retained: knowledge of infant development (as per the Knowledge of Infant Development Inventory KIDI; (MacPhee, 1981)), maternal openness (as per the Ten Item Personality Index (TIPI; Gosling, Rentfrow, & Swann, 2003)), maternal age, an indicator of low maternal education, maternal parity, maternal ethnicity, an indicator of whether the mother has ever had a mental health condition, and an indicator of whether the mother exercises regularly.

In particular, high treatment participants who conducted a forty-eight month survey were more likely to be employed, were older, had higher IQ, had higher consideration of future consequences, were more satisfied with their neighbourhood, yet drank more alcohol during pregnancy. Low treatment participants who conducted a forty-eight month survey were more likely to be older, were Irish, were non first time mothers, exercised more, were more open according to the TIPI, had higher level of education, had greater knowledge of child development, yet had more mental health problems.

In the second step, the significant variables relating to each group were included in separate logit models which were used to calculate the predicted probability of completing the forty-eight month interview for each participant. These logit models were conducted separately for the high and low treatment groups to allow for the possibility that differential attrition processes may exist.

In the third step, the permutation tests which generated Tables 2.1-2.8 were re-estimated applying the inverse of the predicted probability weights. Applying these weights ensured that a larger weight was given to participants who were underrepresented in the sample due to attrition i.e. those who completed the forty-eight month interview yet had baseline characteristics which were similar to those who dropped out. One participant who completed the forty-eight month interview, but did not complete the baseline assessment, was assigned the average IPW weight. A comparison of the IPW results and the unweighted results is contained in Table 3.4.

Table 3.4 - Summary of Main Unweighted & Inverse Probability Weighted (IPW) Results at Forty-Eight Months

PFL Low - *PFL* High	Proportion of Measures Significantly Different at Forty-Eight Months		Proportion of Measures Significantly Different at Forty-Eight Months	
	Unweighted Individual Tests	Unweighted Multiple Hypothesis Tests	IPW Individual Tests	IPW Multiple Hypothesis Tests
Child Development	19% (32)	17% (6)	56% (32)	67% (6)
Child Health	17% (35)	0% (5)	23% (35)	40% (5)
Parenting	6% (36)	0% (8)	8% (36)	0% (8)
Home Environment	50% (4)	~	25% (4)	~
Maternal Health & Wellbeing	14% (21)	25% (4)	19% (21)	50% (4)
Social Support	14% (14)	33% (3)	0% (14)	0% (3)
Childcare	0% (8)	0% (1)	0% (8)	0% (1)
Household Factors & SES	5% (41)	20% (5)	2% (41)	0% (5)
Total Statistically Different	12% (23/191)	13% (4/32)	18% (34/191)	25% (8/32)

In total, 18% of the individual tests were significantly different in the hypothesised direction when IPW was used, which is somewhat greater than the 12% of measures on which significant differences were identified when no weighting was applied. In addition, a substantially higher proportion of the multiple hypothesis tests were significant when IPW was used (25%) versus when no weighting was applied (13%). In general, more significant differences were found in the IPW results for *PFL*'s primary outcomes of child development and child health, while the remaining domains had mainly fewer significant findings. However, in both the individual tests and the multiple hypothesis tests the weighted results identified a higher proportion of significant differences than the unweighted results. Below, we present the IPW adjusted results for the two primary outcomes of *PFL* – child development and child health (see Table 3.5 and Table 3.6, respectively).

Table 3.5 - IPW Results for High and Low Treatment Groups: Child Development

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	IPW Individual Test p^1	IPW Step-down Test p^2
ASQ scores								
ASQ Fine Motor Score	145	(73/72)	47.40	(12.53)	43.68	(13.51)	p<.01	p<.10
ASQ Personal Social Score	145	(73/72)	51.98	(10.92)	49.86	(11.60)	p<.05	ns
ASQ Communication Score	145	(73/72)	53.84	(8.19)	53.06	(9.40)	ns	ns
ASQ Problem Solving Score	145	(73/72)	53.36	(9.05)	52.85	(9.34)	ns	ns
ASQ Gross Motor Score	145	(73/72)	32.67	(27.44)	32.71	(30.41)	ns	ns
* ASQ Social-Emotional Score	145	(73/72)	53.84	(8.56)	54.03	(8.71)	ns	ns
ASQ cut-off scores								
* ASQ Fine Motor cut-off	145	(73/72)	0.14	(0.35)	0.26	(0.44)	p<.01	p<.05
* ASQ Personal-Social cut-off	145	(73/72)	0.03	(0.16)	0.04	(0.20)	ns	ns
* ASQ Gross Motor cut-off	145	(73/72)	0.03	(0.16)	0.04	(0.20)	ns	ns
* ASQ Communication cut-off	145	(73/72)	0.05	(0.23)	0.07	(0.26)	ns	ns
* ASQ Social-Emotional cut-off	145	(73/72)	0.07	(0.25)	0.07	(0.26)	ns	ns
* ASQ Problem Solving cut-off	145	(73/72)	0.05	(0.23)	0.06	(0.23)	ns	ns
Strengths and Difficulties Questionnaire (SDQ)								
* SDQ Peer Problems	145	(73/72)	1.32	(1.42)	1.46	(1.52)	p<.10	ns
SDQ Pro-Social Behaviour	145	(73/72)	8.47	(1.60)	8.15	(1.85)	ns	ns
Child Behaviour Checklist (CBCL) domains								
* CBCL External Problems	144	(73/71)	7.27	(5.49)	8.63	(8.72)	p<.10	ns
* CBCL Total Score	144	(73/71)	21.78	(14.28)	24.63	(23.81)	p<.10	ns
* CBCL Internal Problems	144	(73/71)	6.38	(4.96)	6.86	(7.46)	ns	ns
Child Behaviour Checklist (CBCL) domains cut-off scores								
* CBCL External Problems cut-off	144	(73/71)	0.00	(0.00)	0.06	(0.23)	p<.01	p<.05
* CBCL Internal Problems cut-off	144	(73/71)	0.04	(0.20)	0.11	(0.32)	p<.05	p<.05
* CBCL Total Score cut-off	144	(73/71)	0.03	(0.16)	0.07	(0.26)	p<.05	p<.05
Child Behaviour Checklist (CBCL) subdomains								
* CBCL Aggressive Behaviour cut-off	144	(73/71)	0.00	(0.00)	0.06	(0.23)	p<.01	p<.05
* CBCL Other Problems	144	(73/71)	6.01	(4.21)	6.82	(6.70)	p<.10	ns
* CBCL Anxious/Depressed cut-off	144	(73/71)	0.00	(0.00)	0.03	(0.17)	ns	ns
* CBCL Sleep Problems cut-off	144	(73/71)	0.00	(0.00)	0.03	(0.17)	p<.05	ns
* CBCL Attention Problems cut-off	144	(73/71)	0.00	(0.00)	0.04	(0.23)	p<.05	p<.10
* CBCL Withdrawn cut-off	144	(73/71)	0.03	(0.16)	0.06	(0.23)	p<.10	ns
* CBCL Somatic Complaints cut-off	144	(73/71)	0.03	(0.16)	0.04	(0.20)	ns	ns
* CBCL Emotionally Reactive cut-off	144	(73/71)	0.00	(0.00)	0.03	(0.17)	p<.05	ns
Non Step-down Measures								
DP-3: Cognitive Development standardised score	145	(73/72)	108.34	(13.87)	102.94	(15.85)	p<.05	-
DP-3: Cognitive Development above average cut-off	145	(73/72)	0.34	(0.48)	0.19	(0.40)	p<.05	-
ASQ Standardised Total Score	145	(73/72)	100.43	(16.13)	97.78	(15.86)	p<.05	-
* Child receiving special services	145	(73/72)	0.19	(0.40)	0.19	(0.40)	ns	-

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [3] one-tailed (right-sided) *p* value from an inverse probability weighted individual permutation test with 100,000 replications. [4] one-tailed (right-sided) *p* value from an inverse probability weighted Step-down permutation test with 100,000 replications. All model specifications control for child's gender, and variables where the high and low treatment group differed at baseline. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. 'p<.01', 'p<.05' and 'p<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

In the area of child development, the IPW results found significant differences between the high and low treatment groups on 56% of the individual measures, which was substantially higher than the 19% identified in the unweighted analysis. Specifically, there were twelve significant findings identified in the IPW results that were not significantly different in the unweighted results. According to the weighted results, the high treatment group performed significantly better than the low treatment group on the ASQ personal-social scale, the ASQ fine motor scale and cut-off, the ASQ total score, the SDQ peer problems scale, the CBCL external problems and total score domains, the CBCL total, external and internal cut-off scores, the CBCL aggressive behaviour, other problems, sleep problems, attention problems, withdrawn, and emotionally reactive subdomains, and the DP-3 cognitive continuous and cut-off scores. Additionally, 67% of the step-down categories were statistically significant in the IPW results, compared to 17% in the unweighted results, with the ASQ scores, ASQ cut-off scores, the CBCL cut-off scores and CBCL subdomains surviving adjustment for multiple hypothesis testing.

Table 3.6 - IPW Results for High and Low Treatment Groups: Child Health

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	IPW Individual Test p^1	IPW Step-down Test p^2
Child Health in Last 12 months								
* Had skin problems	145	(73/72)	0.07	(0.25)	0.13	(0.33)	ns	ns
* Received asthma treatment	145	(73/72)	0.15	(0.36)	0.22	(0.42)	$p<.10$	ns
* No. of health problems taken to GP/health centre/casualty	145	(73/72)	1.18	(1.03)	1.29	(1.08)	ns	ns
* Had chest infection	145	(73/72)	0.25	(0.43)	0.29	(0.46)	ns	ns
Child had good health	145	(73/72)	0.95	(0.23)	0.93	(0.26)	ns	ns
* Had an ear infection	145	(73/72)	0.19	(0.40)	0.22	(0.42)	ns	ns
* Had an accident	145	(73/72)	0.10	(0.30)	0.13	(0.33)	ns	ns
* Stayed in hospital for at least one day	145	(73/72)	0.08	(0.28)	0.07	(0.26)	ns	ns
Long Term Child Health								
* Has a physical disability	145	(73/72)	0.00	(0.00)	0.00	(0.00)	ns	ns
* Diagnosed with asthma	143	(73/70)	0.12	(0.33)	0.26	(0.44)	$p<.05$	$p<.10$
* Has a chronic illness	145	(73/72)	0.07	(0.25)	0.06	(0.23)	ns	ns
Meeting Dietary Guidelines								
Vegetables	145	(73/72)	0.38	(0.49)	0.24	(0.43)	$p<.05$	ns
Dairy	145	(73/72)	0.70	(0.46)	0.61	(0.49)	ns	ns
Protein	145	(73/72)	0.34	(0.48)	0.28	(0.45)	ns	ns
Grains	145	(73/72)	0.63	(0.48)	0.60	(0.49)	ns	ns
Fruits	145	(73/72)	0.59	(0.50)	0.58	(0.50)	ns	ns
Sleep								
Usual amount of sleep each day (hours)	145	(73/72)	11.21	(1.21)	10.83	(1.26)	$p<.05$	ns
* Children's Sleep Habits Questionnaire	145	(73/72)	34.75	(8.45)	36.85	(9.42)	$p<.10$	ns
* Child does not have a regular weekend wake-up time	145	(73/72)	0.04	(0.20)	0.08	(0.28)	ns	ns
Child naps during the day	145	(73/72)	0.12	(0.33)	0.08	(0.28)	ns	ns
* Child does not have a regular weekend bedtime	144	(73/71)	0.08	(0.28)	0.11	(0.32)	ns	ns

Continued On Next Page

Variable			M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	IPW Individual Test p[3]	IPW Step-down Test p[4]
Sleep (Continued)								
* Child does not have a regular weekday wake-up time	145	(73/72)	0.01	(0.12)	0.03	(0.17)	ns	ns
* Child does not have a regular week night bedtime	145	(73/72)	0.07	(0.25)	0.06	(0.23)	ns	ns
Length of usual nap (minutes)	15	(9/6)	78.33	(40.93)	90.00	(32.86)	ns	ns
Toilet Training								
Is child toilet trained?	145	(73/72)	0.93	(0.25)	0.85	(0.36)	$p<.05$	$p<.10$
Age child was toilet trained	129	(68/61)	2.60	(0.63)	2.59	(0.52)	ns	ns
Finish toilet training earlier than other children	129	(68/61)	0.32	(0.47)	0.43	(0.50)	ns	ns
Non Step-down Measures								
* BMI Overweight	102	(52/50)	0.27	(0.45)	0.42	(0.50)	$p<.05$	-
* BMI Score	102	(52/50)	17.01	(2.42)	17.09	(1.83)	ns	-
* Child's current weight kg's	103	(53/50)	18.12	(2.70)	18.30	(2.84)	ns	-
Child's current height cm's	144	(73/71)	0.10	(0.30)	0.08	(0.28)	ns	-
Diet Quality Score	145	(73/72)	39.35	(11.13)	36.73	(10.84)	ns	-
Meeting Dietary Guidelines	145	(73/72)	0.10	(0.30)	0.08	(0.28)	ns	-
* Age diagnosed with asthma	25	(9/16)	2.51	(0.84)	1.71	(1.18)	$p<.05$	-
* Daily activities limited by asthma	27	(9/18)	0.22	(0.44)	0.22	(0.43)	ns	-

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [3] one-tailed (right-sided) *p* value from an inverse probability weighted individual permutation test with 100,000 replications. [4] one-tailed (right-sided) *p* value from an inverse probability weighted Step-down permutation test with 100,000 replications. All model specifications control for child's gender, and variables where the high and low treatment group differed at baseline. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

In the area of child health, the IPW results found significant differences between the high and low treatment groups on 23% of the individual measures, which was somewhat higher than the 17% identified in the unweighted analysis. Specifically, the IPW results yielded one significant child health difference which was not statistically different in the unweighted results. The IPW results indicated that the high treatment group were significantly less likely to have received asthma treatment, were less likely to be diagnosed with asthma, were more likely to meet nutritional guidelines for vegetable intake, slept for longer each day, had less sleep problems, were toilet trained, and were less likely to be overweight, compared to the low treatment group. Additionally, 40% of the step-down categories were statistically significant in the IPW results, compared to none in the unweighted results, with the Long Term Child Health and Toilet Training categories surviving adjustment for multiple hypothesis testing.

3.1.4 Summary of Other Domains

The IPW results for parenting (8% significant differences) were largely equivalent to the unweighted results (6% significant differences). There was one additional significant finding such that parents in the high treatment group were more likely to watch television with their child. There was no difference between the weighted and unweighted multiple hypothesis tests (0% significant differences) for parenting outcomes.

There were fewer differences between the IPW results (50% significant differences) and non-IPW results (25% significant differences) for the home environment domain, such that the unweighted results found one additional significant difference which was not found in the weighted results.

The IPW results were largely equivalent (19% significant differences) to the unweighted results (14% significant differences) for the maternal health and wellbeing domain. One additional significant finding was identified, such that high treatment mothers were more likely to be using a valid form of birth control compared to low treatment mothers. There was also one additional significant multiple hypothesis test in the IPW adjusted results for Current Substance Use, which was not identified in the unweighted results.

The IPW results found fewer significant findings (0% significant differences) compared to the unweighted results (14% significant differences) for the social support domain. The unweighted results indicated that more mothers in the high treatment group voted in the last general and local or European elections than in the low treatment group, while this result was not present in the weighted analysis. In addition, the unweighted results identified a significant multiple hypothesis test for the Voting subgroup, which was not replicated in the IPW results.

In the area of childcare, the results for the weighted procedure did not differ from the unweighted results.

The IPW results found fewer significant findings (5% significant differences) compared to the unweighted results (2% significant differences) for the household factors and SES domain. Additionally, while 20% of the step-down categories were significant in the unweighted analysis, none were significant in the IPW analysis.

3.1.5 Attrition Key Findings

The level of official attrition from *PFL* between baseline and forty-eight months was quite low at 16% across the whole sample. Importantly, the level of official attrition was minimal between the thirty-six and forty-eight month interview rounds, with no attrition experienced in the high treatment or comparison treatment groups and only 1% attrition in the low treatment group. Overall, official attrition between programme intake and forty-eight months was slightly higher among the high treatment group (19%) than among the low treatment group (17%) who were less intensively engaged in the *PFL* programme. As the high treatment group were more regularly in contact with the *PFL* programme staff they had more opportunities to officially inform the staff of their desire to dropout from the programme. Indeed, a greater proportion of the low treatment group (21%) was classified as disengaged or missed their forty-eight month interview when compared with the high treatment group (17%). Thus, it is possible that many of the participants who missed the forty-eight month survey represent participants who are less engaged with the programme and more inclined to dropout. Total non-completion (attrition & disengaged) at forty-eight months was very similar among the high (36%), and the low (38%) treatment groups, and lower among the comparison group (29%). Note that these figures represent the highest non-completion rates to date.

In order to test for non-random attrition, we compared the baseline characteristics of those who participated in the forty-eight month survey to those who did not. Overall, there is some evidence that there were systematic differences between these groups. In general, we found that more disadvantaged participants were more difficult to contact or had dropped out of the programme by forty-eight months. For example, in the low treatment group, those who did not participate in this survey were more likely to be first time mothers, were more likely to have low levels of education, and were younger than those who completed it. Mothers in the high treatment group who did not participate were also younger, less likely to be employed at baseline, and had lower levels of cognitive resources. In the comparison group, mothers who did not complete the forty-eight month survey had lower levels of self-efficacy at baseline.

In order to account for any potential bias which differential attrition may introduce, the outcome analyses contained in Chapter 2 were re-estimated using an IPW technique. In both the individual and multiple hypothesis tests there were more significant differences between the high and low treatment groups when the weighting was applied, as compared with the unweighted results. In particular, there were significantly more treatment effects found for the primary domains of child development and child health, while the other domains remained either largely unchanged or somewhat weaker. The difference between the IPW

and non-IPW findings for child development provides evidence of differential attrition at forty-eight months. This suggests that the mothers from the high or low treatment groups who did not participate in the forty-eight month interview had differential characteristics, leading to an under-estimation of the main treatment effects. Thus, when the IPW analysis corrected for this, there were more significant differences between the high and low treatment group children.

3.2 Participant Engagement up to Forty-Eight Months of Age

Engagement refers to the amount of treatment an individual receives during the programme, such as the duration of a prescribed activity or information session, or the frequency with which a participant meets with their mentor. Reviews of home visiting programmes report that, among families who have not dropped out, approximately half of all prescribed home visits are not received (Gomby et al., 1999; Rapoport & O'Brien-Strain, 2001). This is a significant issue as increased frequency of home visits is associated with better child outcomes (Kahn & Moore, 2010; Lyons-Ruth & Melnick, 2004; Nievar, Van Egeren, & Pollard, 2010; Sweet & Appelbaum, 2004). A number of individual, programme, and community factors have been identified as important predictors of engagement in home visiting programmes. This section investigates the level of participant engagement and the determinants of engagement in the *PFL* sample between baseline and the forty-eight month survey.

3.2.1 INSTRUMENTS

Information on participant engagement was gathered from two sources – the *PFL* database maintained by the *PFL* mentors and survey responses from participants at the forty-eight month interview.

MENTOR DATA

Participant engagement using the mentor data was measured in three ways: a) the number of home visits a participant received from entry into the programme until their child was forty-eight months old, b) the percentage of prescribed home visits delivered between intake and forty-eight months (calculated by dividing the number of visits delivered by the number of prescribed visits for this period), and c) the total duration in hours of all delivered home visits between intake and forty-eight months. As there were participants who were randomised into a treatment condition but never engaged with the programme, we examined these measures by restricting the sample to those who have received at least one home visit, although they may have subsequently dropped out of the study. Given that the mentors worked solely with those in the high treatment group, the analysis of engagement was restricted to participants in the high treatment group.

PARTICIPANT DATA

The frequency of meetings that a participant has with their mentor (high treatment group) or information officer (low treatment group) was measured using a single question in the forty-eight month survey which asked how often the participant meets with their mentor/information officer. Possible responses were once a week, two times a month, once a month, less than once a month, or other.

3.2.2 Participant Engagement from Mentor Records

Table 3.7 provides a summary of participant engagement in the *PFL* programme between programme entry and forty-eight months of age for the high treatment group. The analysis was disaggregated into the prenatal period, birth to six months, six to twelve months, twelve to eighteen months, eighteen to twenty-four months, twenty-four to thirty-six months, thirty-six to forty-eight months, and total engagement up to forty-eight months, and includes any participant who received at least one home visit in any period. Thus, the analysis includes those who may have dropped out of the programme before forty-eight months. The *PFL* manual initially set guidelines of weekly home visits during the pre- and postnatal period; however, the implementation team moved to fortnightly visits soon after the programme began as weekly visits

were not feasible to the majority of the *PFL* participants. Thus, the figures below are estimated based on prescribed fortnightly visits.

The prescribed number of prenatal home visits was dependent on when the participant joined the programme, thus, based on average entry into the programme, the prescribed number of home visits between programme entry and forty-eight months was 114 home visits. Table 3.7 shows that on average, participants in the high treatment group received 54.4 home visits between programme entry and forty-eight months. The minimum number of visits received was 1 and the maximum was 129. Figure 3.2 displays the variation in the total number of home visits delivered over the entire period. The average number of home visits in the pre-natal period was 6.2 and the average number of visits in each subsequent six month period up to twenty-four months was quite similar at 7.6, 6.9, 6.3, and 5.6. In the twelve month period between twenty-four and thirty-six months, the average number of visits delivered was 13.2. However, in the twelve month period between thirty-six and forty-eight months only 8.7 visits were delivered on average. This represents a large decline in average engagement within the high treatment group.

These figures were used to calculate the proportion of prescribed home visits actually delivered. Table 3.7 shows that, in total, based on a fortnightly prescribed visit, 48% of visits were delivered on average. The proportion was relatively similar in the earlier phases of the programme; however the current period represents the lowest proportion of visits delivered across the entire duration of the programme. This may be partly attributed to participant fatigue during the final year in the programme, as well as reports from mentors on reducing the amount of contact time with families in the final year of the programme to ensure a successful transition to programme exit.

Table 3.7 also reports the average and total duration of all home visits. These times were based on the amount of time the mentor spent with the participant during the home visit. On average, each visit was 58.3 minutes long, with the shortest visit lasting 5 minutes and the longest visit lasting 81 minutes. The duration of home visits was similar across the different time periods. On average, the high treatment group spent 55.4 hours participating in home visits. The minimum duration spent in home visits was 48 minutes and the maximum was 133 hours in total. Figure 3.3 displays the variation in the duration of home visits over the entire period.

Table 3.7 - Participant Engagement in Home Visits up to Forty-Eight Months of Age

	Prenatal – Birth	Birth - 6 Months	6 Months – 12 Months	12 Months – 18 Months	18 Months – 24 Months	24 Months – 36 Months	36 Months – 48 Months	Total
Prescribed no. of home visits (based on bi-monthly visits)	10	13	13	13	13	26	26	114
Delivered no. of home visits	6.2 (4.3) 0-21	7.6 (4.2) 0-19	6.9 (4.3) 0-17	6.3 (4.1) 0-21	5.6 (3.7) 0-17	13.1 (9.3) 0-39	8.7 (6.6) 0-24	54.4 (30.3) 1-129
% of prescribed home visits delivered (based on bi-monthly visits)	67.3 (45.5) 0-350	58.0 (32.0) 0-146	53.0 (33.2) 0-131	48.2 (31.8) 0-162	43.2 (28.3) 0-130	50.4 (35.7) 0-150	33.4 (25.3) 0-92	47.7 (26.3) 1-109
Mean duration of home visits (mins)	55.1 (17.6) 5-111	59.1 (11.9) 33-91	57.8 (12.5) 15-90	59.9 (11.0) 36-105	60.9 (11.4) 37-89.3	64.1 (14.8) 31-130	57.8 (13.4) 18-105	58.3 (11.2) 5-81
Total duration of home visits (hours)	5.8 (4.1) 0-18	7.6 (4.6) 0-19	6.8 (4.4) 0-18	6.3 (4.3) 0-19	5.8 (3.9) 0-14.3	14.5 (11.6) 0-60.8	8.6 (7.1) 0-26.8	55.4 (33.1) 0.8-133.3
N	96	96	96	96	96	96	96	

Note: The table presents the mean, standard deviation in parentheses, and the minimum and maximum values. These statistics were calculated for participants who received at least one home visit during the prenatal to forty-eight month period. However, for the mean duration, the sample size varies depending on the time period under examination as an average cannot be calculated for participants who received zero visits during the restricted time period.

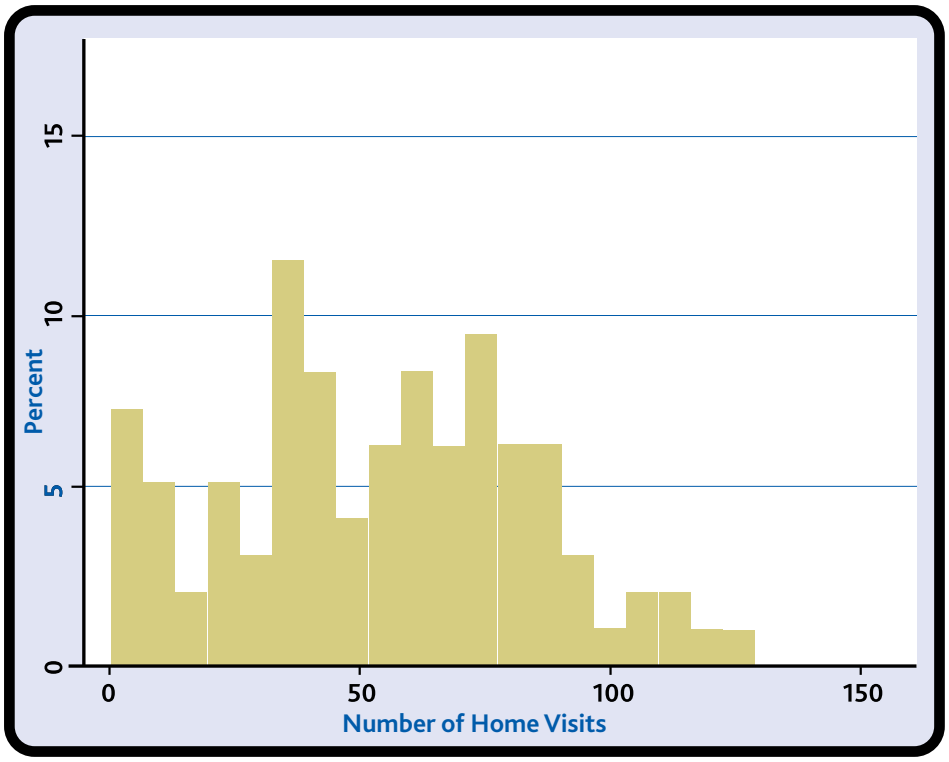


Figure 3.2 - Variation in Number of Home Visits from Entry to Forty-Eight Months

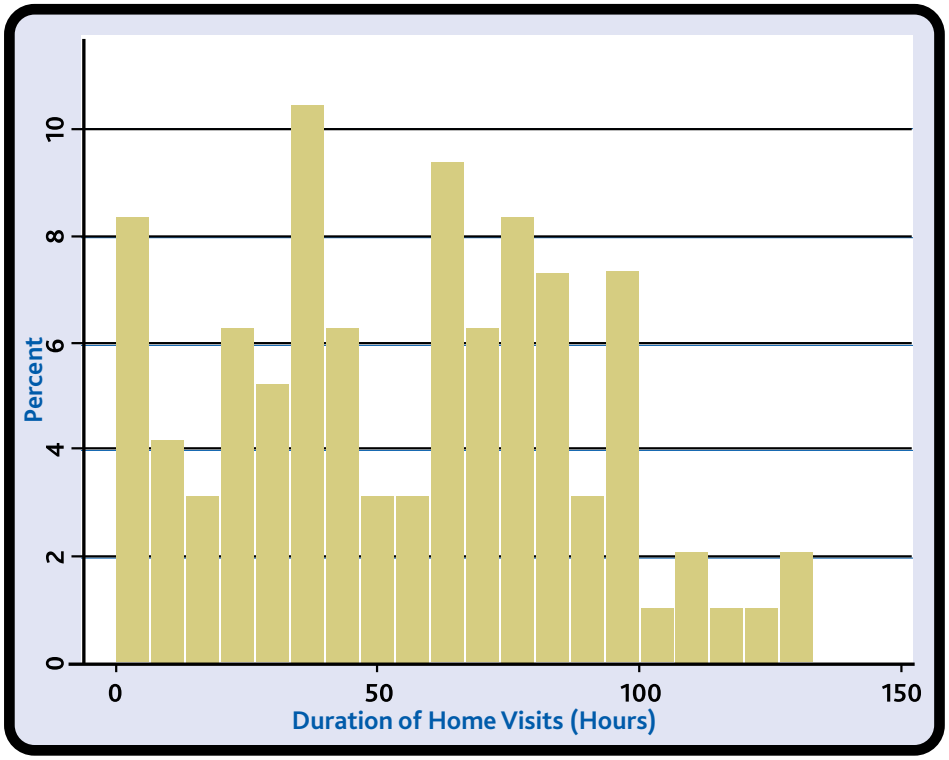


Figure 3.3 - Variation in Duration of Home Visits from Entry to Forty-Eight Months

3.2.3 Participant Engagement from Participant Interviews

HIGH TREATMENT GROUP

Based on participant responses to the forty-eight month interview, 1.5% of participants in the high treatment group reported meeting with their mentor once a week, 50.8% reported meeting twice a month, 34.3% reported meeting once a month, and 13.4% reported meeting their mentor less than once a month. Thus, the majority of participants reported meeting their mentor fortnightly. The *PFL* mentor database finds that zero participants received weekly visits, 2.1% received fortnightly visits, 51% received monthly visits, and 46.9% received home visits less than once a month on average. These differing engagement figures suggest that the participants may be over-reporting how often they meet with their mentor.

LOW TREATMENT GROUP

Based on participant responses to the forty-eight month interview, 0% of participants in the low treatment group reported meeting the Information Officer (IO) more than once a month, 17.7% reported meeting the IO once a month, and 82.4% reported meeting less than once a month. This corresponds to the *PFL* manual which states that the low treatment group should not receive any scheduled meetings. Rather, participants may schedule a meeting with the IO at their discretion.

3.2.4 Factors Associated with Engagement in Home Visiting

As described in Section 3.2.2, participants in the high treatment group were exposed to different degrees of treatment dosage and intensity as defined by the number of home visits they received and the length of contact time with mentors. In this section we examined the factors associated with participant engagement in the home visiting sessions between programme entry and forty-eight months. Specifically, we examined the relationships between participant engagement and a range of socio-demographic and maternal psychosocial factors collected at baseline. This allowed us to test whether the characteristics of the participants who engaged in more home visits were different from those who received less home visits.

Table 3.8 reports the relationship between maternal characteristics measured at baseline and the total number and duration of home visits which the high treatment group received according to the mentor database. It shows that only two maternal characteristics had a significant impact on the frequency and duration of home visits. Specifically, mothers with higher cognitive resources had more home visits since joining the programme and spent a longer amount of time in visits. Whereas mothers who smoked during their pregnancy had fewer home visits since joining the programme and spent a shorter amount of time in visits.

Table 3.8 - OLS Regression Model of Frequency & Duration of Home Visits Between Programme Entry and Forty-Eight Months

Dependent Variables	Frequency of Visits	Duration of Visits
	Prenatal - 48M	Prenatal - 48M
Weeks in pregnancy at programme entry	-0.63 (0.44)	-0.63 (0.46)
Mother's age	0.74 (0.93)	0.75 (0.99)
Partnered	8.23 (9.09)	11.69 (9.70)
Married	-12.38 (11.22)	-10.16 (11.96)
Living with parent(s)	3.54 (8.28)	5.32 (8.83)
First time mother	-0.44 (9.55)	1.73 (10.18)
Low education	3.82 (8.07)	1.63 (8.61)
Mother employed	7.07 (7.56)	8.65 (8.06)
Saves regularly	-8.92 (7.13)	-11.21 (7.60)
Social housing	7.03 (7.10)	7.07 (7.57)
Cognitive resources (WASI)	0.61* (0.33)	0.66* (0.35)
Mental well-being (WHO-5)	1.22 (0.78)	1.33 (0.83)
Vulnerable attachment (VASQ)	0.82 (1.01)	0.59 (1.07)
Self-efficacy (Pearlin)	2.32 (7.46)	4.44 (7.95)
Self-esteem (Rosenberg)	0.07 (1.72)	-0.06 (1.84)
Knowledge of child development (KIDI)	0.23 (0.49)	0.44 (0.52)
Positive parenting attitudes (AAPI)	-2.60 (3.27)	-3.66 (3.49)
Physical health condition	-0.48 (8.55)	-3.84 (9.12)
Mental health condition	3.40 (7.77)	4.34 (8.28)
Smoking during pregnancy	-17.30** (7.74)	-15.97* (8.26)
Drinking during pregnancy	2.10 (7.72)	0.93 (8.23)
Drug ever used	0.74 (9.27)	1.49 (9.89)
Constant	-44.61 (53.84)	-61.94 (57.40)
N	95	95

Note: Regression coefficients, standard errors and *p*-values obtained from an OLS regression. N=95 sample size. *** indicates that the test is statistically significant at the 1% level, ** at the 5% level, and * indicates statistical significance at the 10% level.

3.2.5 Engagement Key Findings

The analysis of participant engagement found that families in the high treatment group received an average of 54.4 home visits by the *PFL* mentors between programme entry and forty-eight months, representing 47.7% of prescribed home visits when based on fortnightly targets. Thus, slightly under half of all home visits were delivered between programme entry and forty-eight months, which translates into about one home visit per month on average. This is consistent with the majority of home visiting programmes which typically finds that over half of prescribed home visits are not delivered. The average number of home visits delivered during the first six months (7.6), the second six months (6.9), the third (6.3) and the fourth six months (5.8) were broadly similar, with a slight downward trend suggesting that a regular pattern of visits was established between the mentors and participants. The average number of home visits delivered between twenty-four and thirty-six months (13.2) was more than twice that of the third and fourth six-month periods, which reverses the pattern of reduced engagement. However, the average number of home visits between thirty-six and forty-eight months (8.7) fell substantially, such that the most recent period saw the lowest proportion of visits delivered by a wide margin. After spending over four years in the *PFL* programme, the participants may have felt that they had derived sufficient support from the mentors and

thus reduced their contact time. In addition, while not a documented element of the programme model, the mentors also discussed reducing the amount of contact time with some families in the final year of the programme to ensure a successful transition to programme exit. The average duration of home visits was in line with the *PFL* manual which recommended that each visit should last between 30 minutes and two hours. The average duration of home visits during the thirty-six to forty-eight month period was just under one hour, which was similar to the average duration experienced at earlier stages of the programme. Thus, the duration of home visits appears to have remained constant over time.

The results of the engagement analysis also suggest fidelity regarding the low treatment group, with the majority of participants reporting that they met the information officer less than once per month. This is in line with the original *PFL* model which states that the information officer is a resource which participants can avail of, if needed, and the information officer should not play the same role as a mentor.

The analysis regarding the relationship between the level of engagement and maternal characteristics between programme entry and forty-eight months indicated that relatively few individual participant characteristics were associated with engagement. Two factors were associated with both the frequency and total duration of home visits – maternal cognitive resources and smoking behaviour during pregnancy. Mothers who smoked during pregnancy had fewer visits and spent less total time in visits. The results also show that mothers with higher cognitive resources, as measured by the Weschler Abbreviated Scale of Intelligence (WASI) at three months, experienced more home visits and had a longer total duration of time spent in visits. The relationship between engagement and cognitive resources has received little in-depth attention in the literature. Cognitive resources are important, as time management skills, the ability to make and keep appointments, participant motivation and an understanding of the regular commitment home visitation entails can contribute to engagement levels (Baker, Piotrkowski, & Brooks-Gunn, 1999; Kitzman, Cole, Yoos, & Olds, 1997). That mothers in the *PFL* sample with higher cognitive resources participated in more home visits and had visits of a longer duration suggest that engagement may also be related to the mother's ability to understand the programme materials and recognise the need for the programme in their lives. By contrast, an evaluation of the Nurse Family Partnership found that the number of home visits decreased as the level of psychological resources increased, measured by intelligence (Shipley Scales of Adaptive Living), mental health, coping skills, self-efficacy, and active coping (Olds & Korfmacher, 1998).

Overall, we found little evidence to suggest that factors which are often identified as determinants of engagement in the literature were present in this sample. For example, factors such as age, marital status, employment status, and socio-emotional functioning were not associated with engagement in *PFL*. In addition, the level of engagement was not associated with socioeconomic factors, (i.e. education, employment) or parenting behaviour. These findings were consistent with the analysis of engagement reported in our previous reports, with very few individual characteristics associated with engagement at any time point.

3.3 Misreporting

Social desirability refers to an individual's tendency to adapt personal behaviour in line with the expectation that the shown behaviour is approved of by their social group. When completing a survey, there is a risk that participants may choose to provide answers that they feel are socially acceptable, rather than answering honestly (Mummendey, 1981). Social desirability has been associated with different personality traits (Stöber, 2001), whereby the extent of participants' desire to behave in a socially acceptable way may vary according to different personalities. It is important to track social desirability in research as the number of affirmative responses an individual provides may depend on the degree to which the affirmative answers are expected to be socially desirable (Cronbach, 1946, as cited in Mummendey, 1981; Edwards, 1953, 1957).

The measures used in the *PFL* interviews typically include a number of sensitive and personal questions, the answers to which may be perceived as potentially socially desirable/undesirable by participants. Therefore, it is important that the effects of social desirability are considered, as responses may be altered to make a favourable impression on the interviewer, which could lead to inaccurate results. The risk of a participant answering in a socially desirable way may depend on whether a response was provided directly to the interviewer, or by self-input without any interviewer involvement. *PFL* interviews typically take place in the family home, thus third parties (e.g. partner, parents, other adults) may be present during the interview. The influence of a third party has the potential to increase socially desirable responses (Casterline & Chidambaram, 1984; Taietz, 1962, as cited in Aquilino, 1997). Ehlers (1973) has suggested that these effects can be minimised and controlled for in a number of ways including non-transparent item construction, comparing specific responses at the item level, using control scales, or simply advising respondents to answer honestly.

3.3.1 Misreporting Instrument

The possibility of misreporting was tested in the forty-eight month interview using a "bogus question". Specifically, following a series of questions relating to their child's development, participants were asked if they had ever heard of the '*Ladder of Learning*'. This is not a term commonly associated with child development and participants would not have been exposed to this expression as a result of being in the programme. Therefore, if a higher proportion of high treatment participants reported that they had heard of the '*Ladder of Learning*', relative to the low treatment or comparison groups, we may have reason to believe that they were more likely to answer in a manner which presented a favourable image of themselves.

3.3.2 Misreporting Results

Table 3.9 presents the results relating to the proportion of the high treatment group, the low treatment group, and the comparison group who reported having heard of the phrase '*Ladder of Learning*'. The findings indicate that a similar proportion (7%) of the high and low treatment groups claimed to have heard of the *Ladder of Learning*. The comparison group was slightly less likely (4%) to report having heard of the phrase. None of the differences between the groups reached statistical significance.

Table 3.9 - Comparison of Misreporting across Groups

Have you heard the phrase ____ ?	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	*High – Low*	*High – LFP*	*Low – LFP*
							p	*p*	*p*
'Ladder of Learning'	0.068	(0.253)	0.068	(0.254)	0.043	(0.204)	ns	ns	ns
N	74		73		70				

Notes: *'M'* indicates the mean. *'SD'* indicates the standard deviation. *p* refers to a two-tailed *p* value from an individual permutation test with 1,000 replications. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

3.3.3 Misreporting Key Findings

The high and low treatment group participants were approximately equally likely to report having heard of the bogus phrase '*Ladder of Learning*'. This provides some evidence that relative to low treatment group participants, members of the high treatment group were no more likely to provide answers which they felt portrayed a better image of themselves as parents.

3.4 Contamination in *PFL*

Contamination occurs when individuals assigned to the control group either actively or passively receive all or part of the services designed for the treatment group (Cook & Campbell, 1979). Contamination may arise for multiple reasons including administrative error, deliberate subversion by programme staff, or an exchange of information between the treatment and control groups. While contamination may occur in any intervention or trial, it is much more prevalent in social or educational interventions aimed at behavioural change (Cook & Campbell, 1979), as the information is more readily transferable. Contamination is particularly undesirable in experimental evaluations as it may bias the results by reducing the mean differences between the treatment and control groups (Torgerson, 2001). Thus the reliability of the evaluation results, which are based on observations from a contaminated control group, may be questionable.

3.4.1 Measuring Contamination in *PFL*

The aim of this section is to measure potential contamination across the high and low treatment groups between programme intake and forty-eight months. Contamination may have occurred if the high treatment group engaged in cross-talk and shared materials with participants in the low treatment group. If substantial contamination occurred during this period it would impede the ability to identify programme effects for the forty-eight month outcomes. The potential for contamination in *PFL* is quite high as it is operating in a very small community with a population of <7,000 and participants were randomly assigned to two different treatment conditions at the individual level. Therefore, it is very likely that participants in the two treatment groups may be neighbours, friends, colleagues or even members of the same family. On the other hand, contamination between the high and low treatment groups may be low as *PFL* is a complex intervention which aims to change the behaviour of participants by building relationships between mentors and participants in the high treatment group. As it is often difficult to achieve behavioural change, even if contamination between the two treatment groups exists, it may not be enough to significantly affect the results. Both indirect and direct measures of contamination were used to provide an indication of whether contamination occurred during this period.

3.4.2 Indirect Measures of Contamination

Information to track contamination indirectly was collected from participants during the forty-eight month interview. Specifically, participants reported how many neighbours they knew who had a child the same age as their own child and also how many neighbours they knew who were taking part in the *PFL* study. The possible response options for all questions were *zero, 1-3, 4-6, 7-10, or 10+*. These questions were used to create *yes/no* binary variables. A binary variable was also used to indicate whether participants in the high and low treatment groups shared their *PFL* materials with anyone else. The results of this analysis are presented in Table 3.10 below.

Table 3.10 - Comparison of Indirect Contamination Responses across Groups

Variable	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	*High – Low* $p^{\dagger}$	*High – LFP* $p^{\dagger}$	*Low – LFP* $p^{\dagger}$
Knows neighbours with children the same age	0.82	(0.38)	0.83	(0.38)	0.87	(0.34)	ns	ns	ns
Knows neighbours taking part in *PFL*	0.58	(0.50)	0.46	(0.50)	0.11	(0.32)	ns	$p<.01$	$p<.01$
Shares *PFL* material with others	0.81	(0.39)	0.51	(0.50)	~	~	$p<.01$	~	~

Notes: '*M*' indicates the mean. '*SD*' indicates the standard deviation. † two-tailed *p* value from an individual permutation test with 1,000 replications. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

The first indirect measure of contamination shows that 82% of the high treatment group and 83% of the low treatment group know neighbours who have a child around the same age as their own child and there is no statistical difference across the two groups. The second indirect measure shows that 58% of the high treatment group and 46% the low treatment group know neighbours taking part in the *PFL* programme and there is no statistical difference across the two groups. A similar proportion of the comparison group know neighbours who have children of a similar age (87%), however, only 11% of the comparison group report knowing a neighbour who is involved in the study. This proportion was significantly lower than the equivalent figure for both the high and low treatment groups. Thus, there was an opportunity for contamination between the high and low treatment groups as participants across the two groups may have interacted with each other, but the likelihood of contamination from the treatment groups to the comparison group was much smaller.

The analysis also shows that a large proportion of both the high (81%) and low (51%) treatment groups shared their *PFL* materials with others, however a significantly greater proportion of the high treatment group reported sharing their materials with other parents in the community ($p<.01$). As the high treatment group received substantially more information and materials from the *PFL* programme than the low treatment group, this result was as expected and again suggests that there was a potential for contamination across the treatment groups, although we do not know who they shared their materials with.

3.4.3 Direct Measures of Contamination

In order to elicit a direct measure of contamination, a 'blue-dye' question was included in the forty-eight month interview. This question asked participants from the high and low treatment groups and the comparison group if they have heard of a particular parenting phrase, in this case *'The Feeling Wheel'*, and if they know what this phrase means. *'The Feeling Wheel'* is a circular chart, with cartoon faces showing different emotions. It was provided to participants in the high treatment group only and was discussed by the mentors during the home visits between thirty-six and forty-eight months.

This question may be used to measure contamination as if the participants in the low treatment group stated that they knew what the phrase meant and they correctly identified what it was, it suggests that they may have accessed material or information intended for the high treatment participants only. Specifically, if there was no statistical difference in the proportion of the low treatment group and the high treatment group reporting knowledge of the phrase, it suggests contamination did occur. A comparison of the responses of the high and low treatment groups to the comparison group, where contamination could not have occurred given the geographic distance, can then be used to estimate the magnitude of contamination. If contamination did not occur, a significantly higher proportion of the high treatment group should have stated that they were aware of this term or engaged in this behaviour compared to the low treatment group and the comparison group. If contamination between the high and low treatment groups was present, there should be a significant difference in the responses of the low treatment group and the comparison group - this difference represents the magnitude of contamination.

The 'blue-dye' question included in the forty-eight month survey had two possible responses – *yes, no,* which was used to generate a binary variable indicating whether the participants had heard of the phrase or not. It is possible, on social desirability grounds, that some participants who were not familiar with the phrase stated they were familiar with the phrase. In order to capture this, participants who stated that they had heard the phrase were asked a follow-up question: *"Can you tell me what The Feeling Wheel is"*. This was an open ended question which was subsequently recoded based on whether the respondent provided a valid response.

Table 3.11 shows that 27% of participants in the high treatment group reported having heard *'The Feeling Wheel'*, compared to 4% of both the low treatment and comparison groups. The difference between the high treatment group and the low treatment group was statistically significant ($p<.01$), as was the difference between the high treatment group and the comparison group ($p<.01$). The difference between the low treatment group and the comparison group was not statistically significant.

Table 3.11 - Comparison of Direct Contamination Responses across Groups

Have you heard the phrase ____ ?	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	*High – Low*	*High – LFP*	*Low – LFP*
							p	*p*	*p*
'The Feeling Wheel'	0.27	(0.45)	0.04	(0.20)	0.04	(0.20)	*p*<.01	*p*<.01	ns
N	74		73		70				

Notes: *'M'* indicates the mean. *'SD'* indicates the standard deviation. *p* refers to a two-tailed *p* value from an individual permutation test with 1,000 replications. 'ns' indicates the variable is not statistically significant. '*p*<.01', '*p*<.05' and '*p*<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

In order to provide a more accurate measure of contamination, participants who stated that they had heard of the parenting phrase, yet provided incorrect responses regarding how best to engage in this behaviour, were removed from the analysis. Six percent of the high treatment group, 1% of the low treatment group, and 1% of the comparison group gave an incorrect response. The results reporting the proportion of participants who accurately described *'The Feeling Wheel'* were then re-estimated.

Table 3.12 shows that once we excluded participants who stated that they had heard of the phrase, yet provided incorrect responses, the pattern of significant differences between the groups was maintained. Once again the difference between the high treatment group and the low treatment group was statistically significant (*p*<.01), as was the difference between the high treatment group and the comparison group (*p*<.01). As before, the difference between the low treatment group and the comparison group was not statistically significant

Table 3.12 - Corrected Comparison of Direct Contamination Responses across Groups

Have you heard the phrase ____ ?	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	*High – Low*	*High – LFP*	*Low – LFP*
(corrected)							*p*	*p*	*p*
'The Feeling Wheel'	0.21	(0.41)	0.03	(0.17)	0.03	(0.17)	*p*<.01	*p*<.01	ns
N	68		72		69				

Notes: *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p* value from an individual permutation test with 1,000 replications. 'ns' indicates the variable is not statistically significant. '*p*<.01', '*p*<.05' and '*p*<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

3.4.4 Contamination Key Findings

Overall, the contamination analysis revealed two findings. The indirect measures of contamination indicated that the potential for contamination in the *PFL* programme was high as participants in both the high and low treatment groups reported knowing multiple neighbours in the *PFL* programme with children of similar ages to their own. While there were no statistical differences across the high and low treatment group regarding the proportion of each group knowing neighbours with similarly aged children, the proportion reporting sharing their *PFL* materials was higher in the high treatment group. This result was as expected as the high treatment group received more materials from the programme than the low treatment group.

While the indirect measures only provide an indication of the likelihood of contamination, they cannot be used to directly determine whether contamination occurred. The 'blue-dye' question was therefore used to elicit a direct measure of contamination. The results indicated that once incorrect responses had been removed, the high treatment group reported a greater knowledge of the phrase than the low treatment group and the comparison group. This suggests that contamination may not have occurred between the high and low treatment groups. The fact that a similar proportion of the low treatment and comparison groups reported knowledge of the phrase is further evidence of the absence of major contamination.

Chapter Four

Dynamic Results and Comparison Group Summary

This chapter presents summaries of the dynamic analysis which examined changes in child and parent outcomes between baseline and forty-eight months for the high and low treatment groups. It also summarises the results comparing the forty-eight month outcomes of the low treatment group to the comparison group. The purpose of these analyses is to explore different aspects of the data not captured in the main analysis. .

4.1 Dynamic Analysis

A number of standardised instruments used to evaluate the *PFL* programme were collected at multiple time points. This allows us to compare the responses for the same participants over time in order to track changes in child and parent outcomes. It also allows us to examine changes in outcomes across the high and low treatment groups. This process of tracking change over time is referred to as dynamic analysis. Table 4.1 lists the instruments collected at multiple time points between baseline (BL) and the forty-eight month data collection point.

Table 4.1 - Instruments included in the Dynamic Analysis

	Measure	Baseline	6 Month Interview	12 Month Interview	18 Month Interview	24 Month Interview	36 Month Interview	48 Month Interview
Child Development	Ages & Stages Questionnaire (ASQ)		X	X	X	X	X	X
	Ages & Stages Questionnaire (ASQ) Social/Emotional		X	X	X	X	X	X
	Developmental Profile 3 (DP-3)			X	X	X	X	X
	Child Behaviour Check-List (CBCL)					X	X	X
Maternal Health & Wellbeing	Maternal Psychological Wellbeing (WHO-5)	X	X	X			X	X
	Edinburgh Postnatal Depression Scale (EPDS)		X		X	X	X	X
	Pearlin Self-Efficacy scale	X		X				X
	Rosenberg Self-Esteem Scale	X			X	X		X
Parenting	Parenting Daily Hassles Scale				X		X	X
	Parenting Styles and Dimensions Questionnaire (PSDQ)						X	X
	Parenting Stress Index (PSI)		X			X		X
Home Environment	Framingham Safety Survey (FSS)		X		X			X

Generalized Estimating Equations (GEE) were used to evaluate changes in parent and child outcomes over time and to determine whether the programme's impact increased or diminished from wave to wave. GEE was put forward by Zeger and Liang (1986) and allows for the analysis of data when there is an unknown correlation structure between observations. It is commonly applied to longitudinal data where multiple observations are taken from the same participant over time and as a result should not be considered as independent. Failure to account for the correlation within participants can result in an incorrect estimation of regression coefficients' variances. This in turn could lead to incorrect conclusions regarding the research question of interest. For the following dynamic analysis, no predetermined correlation structure was assumed.

The GEE approach uses the generalised linear model to estimate regression parameters which are more efficient relative to ordinary least squares regression, as it accounts for the within subject correlation of responses on dependent variables (Ballinger, 2004). GEEs can also be applied regardless of whether the dependent variable is continuous, binary, or in the form of counts (Hanley et al., 2003). A further advantage of GEE is that it does not require a balanced dataset. As the sample size in the current evaluation changes over time, this method allows us to utilise all data from each time point rather than restricting the analysis to participants who were assessed at each time point. Each GEE model has the repeated outcome of interest as its dependent variable. The independent variables are a set of time dummy variables, a treatment indicator, and interaction variables between time and treatment.

4.1.1 Dynamic Analysis Results

In total, 50 individual dynamic tests were conducted. The goal of the dynamic analyses was to determine whether the magnitude of treatment effects identified on standardised instruments at forty-eight months differed significantly from the size of treatment effects arising from previous waves on the same measures. Overall, only two (4%) of the 50 measures upon which dynamic analyses were conducted yielded significant findings relating to changes in the size of treatment effects at forty-eight months relative to previous waves. Both of these findings, which were in the domains of child development and parenting, are presented below.

CHILD DEVELOPMENT

The comparison of the standardised DP-3 scores found a statistically significant difference between the change in results of the high and low treatment groups from thirty-six to forty-eight months. Both groups experienced a decrease in their cognitive scores between thirty-six and forty-eight months. While the reduction was experienced by both groups, the low treatment group experienced a slightly stronger decline than the high treatment group.

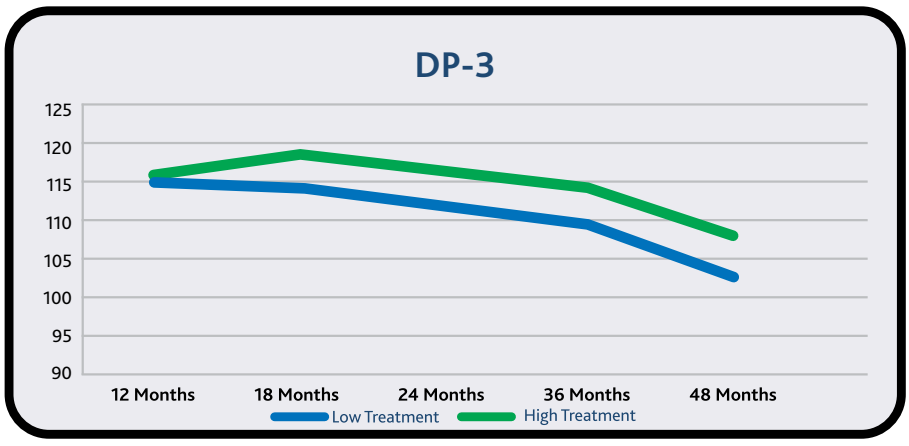


Figure 4.1 - High and Low Treatment DP-3 Scores over time

PARENTING

There was a significant change over time between the high and low treatment groups on the Tasks subscale of the Parenting Daily Hassles Scale. While at thirty-six months the low treatment group exhibited a higher score than the high treatment group, indicating that the low treatment parents had more difficulty with tasks at this time, this trend had reversed by forty-eight months, and the change in score over time was significantly larger for those in the high treatment group.

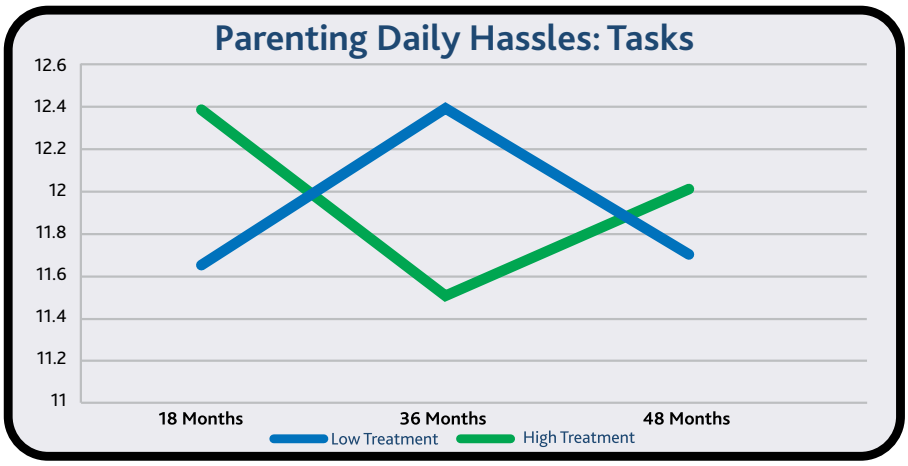


Figure 4.2 - Parenting Daily Hassles Scale Tasks Scores over time

4.2 Low Treatment and Comparison Group Analysis

This section presents the results comparing the forty-eight month outcomes of the low treatment group to those of the external, no treatment comparison group. The purpose of this comparison group was to have an external sample which was not at risk of contamination from the high treatment group. Therefore, if the low treatment group outperformed the comparison group, it may be an indication that the low treatment group was contaminated by the high treatment group and thus is not a viable comparison group for determining the impact of the *PFL* programme. In addition, comparing the low treatment group to the comparison group allows us to measure the impact, if any, of the low treatment supports. Therefore, a finding that the low treatment group outperformed the comparison group at forty-eight months suggests that either the low treatment group may have received some of the high treatment supports, or that the low treatment supports were effective.

4.2.1 Hypothesis

We hypothesised that there would be a limited number of statistically significant differences between the *PFL* low treatment group and the LFP comparison group. Furthermore, we hypothesised that the low treatment group would outperform the comparison group on some domains, while the comparison group would outperform the low treatment group on others.

4.2.2 Key Findings

In total, 192 measures were included in the analysis, of which positive significant differences between the low treatment group and the comparison group were found on 14 (7%) of the individual measures and 15 (8%) were statistically significant in the opposite direction, such that the comparison group outperformed the low treatment group. The finding that the comparison group outperformed the low treatment group on a slightly higher proportion of measures suggests that contamination was not an issue, and also suggests that the low treatment supports had minimal effect. Below, we summarises the results for each domain.

CHILD DEVELOPMENT

Children in the low treatment group and the comparison group did not differ on many of the domains of child development. There were no differences between the groups in terms of the ASQ, the SDQ, or the CBCL. However, the two groups differed significantly on one domain of child development, with the comparison group outperforming the low treatment group:

- Children in the comparison group had higher scores on the DP-3 binary measure than children in the low treatment group, suggesting that they were more likely to score above average in terms of cognitive development.

CHILD HEALTH

Children across both groups differed significantly in a number of domains of child health, with the comparison group generally outperforming the low treatment group:

- Children in the low treatment group were less likely to have suffered from a chest infection in the last year than children in the comparison group. However, they were more likely to have been diagnosed with asthma.
- A higher proportion of children in the comparison group met the daily requirements for consumption of dairy products than low treatment group children.
- Children in the comparison group were less likely to be overweight, as measured by their BMI, than those in the low treatment group.

PARENTING

The low treatment group and the comparison group did not differ on many domains of parenting, including the PDH scale, the PSI, and their views of the importance of different school readiness traits. There were statistically significant differences between the low treatment group and the comparison group in two domains of parenting:

- Mothers in the comparison group spent longer watching television with their children than mothers in the low treatment group.
- Mothers in the low treatment group had lower scores on the PSDQ subdomain of hostility, indicating that they were less likely to engage in hostile parenting behaviours.

HOME ENVIRONMENT AND SAFETY

There were no differences across the two groups on any of the measures of the home environment and safety. Mothers in both groups were equally likely to have a social worker working with the family, to provide a safe environment for the child, and to provide an enriching home learning environment for the child.

MATERNAL HEALTH AND WELLBEING

The low treatment group and the comparison group differed on a number of domains of maternal health and wellbeing:

- Mothers in the low treatment group reported fewer GP visits in the previous year than those in the comparison group.
- Mothers in the low treatment group scored higher in terms of self-efficacy as measured by the Pearlin self-efficacy scale than mothers in the comparison group.
- Mothers in the comparison group were less likely to be at risk of depression as measured by the Edinburgh Postnatal Depression Scale.
- Mothers in the comparison group were more likely to report that they used a valid form of birth control than those in the low treatment group.

MATERNAL SOCIAL SUPPORT

There were a number of differences across the two groups in terms of maternal social support:

- Mothers in the low treatment group were more likely to have a partner who was involved in the child's life. However, in the comparison group, mothers were more likely to report that the child's biological father was involved in the child's life.
- Mothers in the low treatment group were more likely to report that the child's father made regular maintenance payments than those in the comparison group.

CHILDCARE

The low treatment group did not differ significantly from the comparison group in terms of the number of hours per week that the child was in childcare or the type of childcare used. However, there was one statistically significant difference between the two groups:

- Children in the comparison were, on average, older when starting childcare than children in the low treatment group.

HOUSEHOLD FACTORS AND SES

There were no differences between the two groups in relation to worries about finances and paternal employment. However, there were a number of differences across the groups in the other domains which measure household factors and SES:

- Mothers in the low treatment group were more likely to be in a relationship or to be married than those in the comparison group.
- Rates of employment were more favourable in the comparison group, with mothers more likely to report that they were in paid employment. More mothers in the low treatment group reported being in part-time employment than those in the comparison group, although they were also more likely to report that they had been unemployed for more than one year. Mothers in the low treatment group were less likely to report that they were in receipt of unemployment benefit.
- Annual wages for full-time working mothers were higher for those in the comparison group. This group were also more likely to be optimistic that their household financial situation would improve over the coming year than those in the low treatment group.
- In relation to the presence of domestic risks, fewer mothers in the low treatment group reported members of the family had issues with suicidal thoughts, abuse or addiction, while they were more likely to report experiencing bereavement in the family within the last year.

Overall, the results of the differences between the low treatment group and the comparison group were mixed. This finding supports the study design as it suggests that the low treatment group was not performing systematically better than the comparison group across most domains. The low treatment group outperformed the comparison group on 7% of measures while the comparison group outperformed the low treatment group on 8% of measures, with the majority of the differences being clustered around household factors and SES outcomes. The lack of a coherent pattern suggests that the low treatment group did not receive the services and supports designed for the high treatment group.

Chapter Five

Testing for Differential Programme Effects

5.1 Introduction

There is some evidence to suggest that certain groups of participants may benefit more from early intervention programmes than others (e.g. Heckman, Malofeeva, Pinto, & Savelyev, 2010). It is possible that the main results for child development reported in Chapter 2 may mask treatment effects for particular types of participants. In order to investigate differential effects of the *PFL* programme on child development, interaction and subgroup analyses were conducted for categories commonly found to be of relevance to early child and family interventions - child gender, maternal parity status, and maternal cognitive resources. A further subgroup analysis investigated differences by Triple P attendance.

5.2 Method

The interaction analyses for gender, maternal parity status, and maternal cognitive resources were conducted using multiple linear regression. These models included three variables indicating the treatment status of the participant, the category of interest (e.g. gender), and an interaction between treatment status and the category of interest. Subgroup analyses were then conducted using the permutation method described in Chapter 2 to explore any statistically significant interaction effects in more detail. This method was particularly advantageous for the subgroup analysis as the sample sizes used were small. The subgroup analysis tested treatment effects for all significant interactions for each subgroup: child gender (boys and girls), maternal parity status (primiparous and multiparous), and maternal cognitive resources (high cognitive resources and low cognitive resources).

Permutation tests were used to investigate differential effects by Triple P attendance. As Triple P was only offered to high treatment participants, the first stage of the analysis compared only Triple P participants to non-Triple P participants within the high treatment group. For outcomes which were statistically significant in this analysis, further analyses were carried out by comparing the high treatment non-Triple P participants to the low treatment group and by comparing the high treatment Triple P participants to the low treatment group.

Below we provide a brief literature review on why differential programme effects may be expected within each subgroup and a presentation of the results for each interaction and subgroup analysis. Findings related to the primary domain of child development are reported for gender, maternal parity status, and maternal cognitive resources. Findings related to parenting are also reported in the Triple P analysis.

5.3 Gender

Gender differences emerge early in childhood and are amplified as children grow (Campbell & Eaton, 1999). Children become either 'masculine' or 'feminine' at an early age and by the time they are 4 or 5 years old generally exhibit well-established gender-based preferences for toys and activities (Bem, 1983). In terms of child outcomes, meta-analytic evidence suggests that girls and boys display differential developmental patterns across multiple domains including the prevalence and expression of behaviour problems (Chaplin, 2015); language use (Leaper & Smith, 2004); and aggression (Archer, 2004). Although some gender differences are biologically based (Cho et al., 2010; Elsmén et al., 2004; Feldman et al., 2009; Hintz et al., 2006), debate remains about how other differences emerge as children grow. A number of theories have been proposed to explain these differences, including social learning theory, cognitive developmental theory, and gender schema theory. The most commonly utilised theory, social learning theory, suggests that children are rewarded for engaging in gender appropriate behaviours and that these skills and traits are reinforced in each gender (Field et al, 1980; Halpern, 1997). For example, boys may be rewarded for performing well on visio-spatial tasks and are expected to be more independent, while girls are expected to be more verbal and are typically praised for language skills (Clearfield & Nelson, 2006; Reinisch & Sanders, 1992 as cited in Cho et al., 2010). Evidence also suggests that these systematic gender

differences in interaction are apparent from as early as six months postpartum, with children's gender-based behavioural reactions shaping interactions (Weinberg et al., 1999).

As there are gender differences in children's development and gender differences in parents' interaction patterns with their children, it is conceivable that interventions may have gender-based differential effects on outcomes (Webster-Stratton, 1996). Indeed, based on this assumption, a re-evaluation of three large scale studies demonstrated that centre-based early intervention programmes significantly improved the later life outcomes for females, particularly in terms of academic achievement, while treatment effects for males were modest or non-existent (Anderson, 2008). An evaluation of the Perry Preschool Programme found the programme had a positive impact on academic motivation, primarily among girls. In addition, while both genders experienced boosted long-term achievement test scores, this improvement was stronger for girls than boys (Heckman et al., 2013). Furthermore, a meta-analysis of four large experimental studies of both home and centre-based programmes reported greater impacts on achievement test scores for girls than boys (Barnett, 1995). Evidence also suggests that gender differences may interact with socioeconomic disadvantage, with girls from chronically poor backgrounds demonstrating the greatest increase in cognitive skills and decrease in suspension, expulsion, and grade repetition as a result of engagement in a Head Start programme, while participation did not have a significant impact on boys' outcomes (Joo, 2010).

In the analyses of home visiting interventions, child gender is often included as a control variable (Ramey & Ramey, 1994; Wagner & Clayton, 1999), but few studies specifically test for differential gender effects. There are however some exceptions. For example, Sidora-Arcoleo et al. (2010) found that the effects of the Nurse Family Partnership programme at age two were concentrated among females. Furthermore, Eckenrode et al. (2010) found that the Nurse Family Partnership programme's long-term beneficial effects at age 19 were driven by the substantial impact of the programme on girls, with few programme effects observed for boys. This reflects results from a German programme, Pro Kind, which found that the programme was more beneficial for girls than for boys in relation to their mental development in the first year (Jungmann et al., 2015).

To date, there has been little research to inform whether home visiting interventions affect children differently according to their gender, or indeed, how this process might occur. Sidora-Arcoleo et al. (2010) posit that the inclusion of gender analysis is critical in the evaluation of home visiting programmes, however, given the sparsity of literature in the area it is difficult to hypothesise the likely gender effects. Based on findings from the limited pool of available literature, we hypothesised that there would be significant differences in outcomes according to gender with more effects for girls than boys.

5.3.1 Gender Results

In order to investigate differential treatment effects by gender, interaction analyses were conducted. The results presented in Table 5.1 indicate that very few differences arose in how the programme impacted upon boys and girls. The interaction of gender with treatment status was significant for only one of the 32 outcomes analysed. This one significant result suggests that the programme had differential gender effects in respect to the children's pro-social behaviour. To explore this result in more detail, Table 5.2 reports the results of a subgroup analysis examining the outcomes of high and low treatment group boys (44% of the sample) and the outcomes of high and low treatment group girls (56% of the sample). The results indicate that the differential effect was driven by boys, such that boys in the high treatment group were significantly more likely to display pro-social behaviour than boys in the low treatment group ($p<.05$, $d=0.55$). Thus, overall the results suggest that the programme largely did not affect children differently depending on their gender.

Table 5.1 - Results for Linear Regression Analysis of Gender on Child Development Outcomes

Variable	N	Treatment		Gender		Treatment *Gender	
ASQ scores							
ASQ Communication Score	147	1.25	(2.16)	4.41**	(2.07)	0.18	(2.87)
ASQ Gross Motor Score	147	-0.81	(2.17)	0.82	(2.08)	1.60	(2.88)
ASQ Fine Motor Score	147	6.08**	(3.02)	11.67***	(2.90)	-2.21	(4.01)
ASQ Problem Solving Score	147	-0.40	(2.28)	2.32	(2.19)	2.40	(3.03)
ASQ Personal-Social Score	147	4.15	(2.78)	5.95**	(2.66)	-2.57	(3.69)
* ASQ Social-Emotional Score	147	-4.56	(7.34)	-4.28	(7.03)	9.01	(9.75)
ASQ cut-off scores							
* ASQ Communication cut-off	147	-0.04	(0.06)	-0.12**	(0.05)	0.01	(0.07)
* ASQ Gross Motor cut-off	147	0.01	(0.04)	0.00	(0.04)	-0.06	(0.06)
* ASQ Fine Motor cut-off	147	-0.23**	(0.09)	-0.35***	(0.09)	0.13	(0.12)
* ASQ Problem Solving cut-off	147	0.00	(0.05)	-0.03	(0.05)	-0.02	(0.07)
* ASQ Personal-Social cut-off	147	-0.02	(0.04)	-0.05	(0.04)	-0.00	(0.06)
* ASQ Social-Emotional cut-off	147	-0.02	(0.06)	-0.00	(0.06)	0.03	(0.08)
Child Behaviour Checklist (CBCL) domains							
* CBCL Internal Problems	146	-0.20	(1.60)	0.01	(1.54)	-0.31	(2.14)
* CBCL External Problems	146	-0.72	(1.85)	0.07	(1.78)	-0.91	(2.46)
* CBCL Total Score	146	-3.46	(4.98)	-1.79	(4.79)	1.38	(6.62)
Child Behaviour Checklist (CBCL) domains cut-off scores							
* CBCL Internal Problems cut-off	146	-0.05	(0.06)	0.00	(0.06)	-0.02	(0.08)
* CBCL External Problems cut-off	146	-0.07*	(0.04)	-0.03	(0.04)	0.03	(0.05)
* CBCL Total Score cut-off	146	-0.05	(0.05)	-0.06	(0.05)	0.01	(0.07)
Child Behaviour Checklist (CBCL) subdomains							
* CBCL Emotionally Reactive	146	0.03	(0.53)	0.01	(0.51)	-0.09	(0.71)
* CBCL Anxious/Depressed	146	-0.20	(0.52)	0.25	(0.50)	-0.04	(0.70)
* CBCL Somatic Complaints	146	-0.32	(0.49)	-0.16	(0.47)	0.56	(0.65)
* CBCL Withdrawn	146	0.29	(0.47)	-0.09	(0.45)	-0.74	(0.62)
* CBCL Sleep Problems Behaviour	146	-0.48	(0.57)	-0.21	(0.54)	0.48	(0.75)
* CBCL Attention Problems	146	0.06	(0.46)	-0.20	(0.44)	-0.36	(0.61)
* CBCL Aggressive Behaviour	146	-0.79	(1.52)	0.28	(1.46)	-0.55	(2.02)
* CBCL Other Problems	146	-2.05	(1.40)	-1.66	(1.35)	2.12	(1.87)
Strengths and Difficulties Questionnaire (SDQ)							
* SDQ Peer Problems	147	0.32	(0.37)	0.07	(0.35)	-0.80	(0.49)
SDQ Pro-Social Behaviour	147	0.89**	(0.43)	0.75*	(0.41)	-0.99*	(0.57)
Other Measures							
ASQ Standardised Total Score	147	3.75	(3.84)	10.07***	(3.68)	0.26	(5.10)
DP-3: Cognitive Development standardised score	147	4.43	(3.74)	2.50	(3.58)	2.43	(4.96)
DP-3: Cognitive Development above average cut-off	147	0.15	(0.11)	0.12	(0.10)	0.00	(0.14)
* Child receiving special services	147	-0.02	(0.09)	-0.16*	(0.09)	0.00	(0.13)

Notes: Standard errors in brackets. ***$p<.01$, **$p<.05$, *$p<.10$

Table 5.2 - Subgroup Analysis Results for High and Low Treatment Group by Gender

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test $p^{\dagger}$	Effect Size *d*
SDQ Pro-Social Behaviour								
Boys	64	(37/27)	8.59	(1.48)	7.70	(1.88)	$p<.05$	0.55
Girls	83	(37/46)	8.35	(1.70)	8.46	(1.80)	ns	0.06

Notes: '*N*' indicates the sample size. '*M*' indicates the mean. '*SD*' indicates the standard deviation. † one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test.

5.4 Parity

Becoming a mother can be a challenge, with many first time or primiparous mothers reporting stress, anxiety, conflict, and tension in the transition into motherhood and the time immediately after birth (Raynor & England, 2011; Razurel et al., 2011). Yet for some women, new motherhood can be a welcome experience, perceived as a developmental milestone (Cronin, 2003). A study on the experiences of young Irish first time mothers identified that women were generally unprepared for birth and becoming a mother (Cronin, 2003), even in cases where they had attended prenatal classes and spoken to healthcare professionals. Some mothers reported being stressed when they returned home and struggled to adjust to the responsibilities of caring for a baby, while others experienced depression and loneliness. Darvill et al. (2010) described how first time mothers typically undergo a significant transition in early pregnancy, but may experience difficult periods before and after birth, such as diminished self-concept and lack of energy; however they regain a sense of control over the situation as their confidence increases with their continued experience of infant care. Darvill et al. (2010) also posited that such mothers lack the supports to help them overcome difficulties, citing the importance of support from other mothers and health professionals. Similarly, Sayil, et al. (2006) suggested that first time mothers may benefit from training and information, focusing specifically on adapting to motherhood, infant care, maternal self-perception and anxiety reduction, particularly in lower socioeconomic groups.

The psychological impact of first time motherhood can be determined by demographic, environmental, and individual factors which influence maternal experiences and wellbeing (Leahy-Warren et al., 2011; Sayil, et al. 2006). Sayil, et al. (2006) found that first time mothers with lower income, self-esteem, and self–efficacy were at increased risk of prenatal anxiety. Furthermore, maternal postnatal depression was significantly associated with unplanned pregnancy and higher prenatal anxiety, though mothers who had more perceived support from their partners exhibited fewer depressive symptoms. Leahy-Warren et al. (2011) concluded that maternal self-efficacy in first time mothers can be significantly enhanced by family support and other sources of informal support such as friends, suggesting that social support contributes significantly to the wellbeing of first time mothers six months post-delivery, and can in fact be a preventive factor for postnatal depression, as it helps to reduce maternal anxiety, stress, and fatigue.

Psychological distress in pregnancy can negatively affect mother and baby relationships and later outcomes (Furber et al., 2009). Several studies have examined differences in child outcomes by parity status. Birth order has a strong effect on a number of child outcomes, with first-born children typically outperforming those who are born subsequently (Price, 2008). There is also a strong link between the availability of maternal resources and child outcomes. Maternal time spent with children is crucial to improving children's educational outcomes (Cho, 2011), yet the amount of time available to a mother to spend with her child may diminish as she has subsequent children. Accordingly, there is a negative relationship between birth order and maternal quality time, with first time mothers typically spending 25-30 minutes more quality time with a first-born than a subsequent child (Price, 2008). First time mothers typically pay more attention to their children, spending more time talking to and interacting with them, than those with more children (Monfardini & See, 2012). Murray et al. (2008) examined the impact of maternal depression and adversity on mother-child interactions among first time mothers, and found that disturbances in mother-child interactions predicted poorer infant cognitive outcomes at eighteen months. Goldstein (2004) found

that primiparous women who experienced what he described as maternity 'blues' avoided all types of physical contact with their babies, whereas multiparous women with the same difficulty provided firm touch and holding. This is similar to the findings identified in women with postnatal depression (Murray et al., 1996). Maternal touch is considered an essential part of the mother-child relationship and positively impacts child developmental outcomes (Goldstein, 2004).

As the majority of home visiting programmes target primiparous mothers in order to prevent harmful parenting behaviours from developing (Howard & Brooks-Gunn, 2009), there is limited evidence on whether such programmes may have different effects on families depending on parity. *PFL* is thus uniquely placed to address this issue and provide some initial evidence on whether home visiting programmes may be beneficial to non-first time mothers.

There is some evidence that primiparous mothers may receive more benefits from home visiting programmes than multiparous mothers (Monsen, 2006). Healthy Families America provides home-based support to disadvantaged families, including both primiparous and multiparous parents, from prenatal/birth to 3 - 5 years (Howard & Brooks-Gunn, 2009; Rodriguez, et al., 2010). DuMont et al. (2008) evaluated the impact of the Healthy Families New York programme on parenting behaviours in the first two years of life, reporting that first time mothers in the intervention group were less likely to engage in minor physical aggression and harsh parenting than those in the control group. These results were subsequently replicated in a follow-up study at 3 years which showed that young, first time mothers were less likely to engage in harsh parenting during structured interactive tasks than those in the control group (DuMont, et al., 2010). Differential parity results have also been observed in several lesser-known home visiting programmes. Research on the Dutch SCRIPT study, which investigated the effectiveness of an early intervention home visiting programme on externalising problems, found that primiparous mothers in the intervention group displayed a significant increase in their use of positive discipline strategies compared to primiparous mothers in the control group, however this effect was not observed for multiparous mothers (Stolk et al., 2008). Armstrong et al. (1999) carried out a randomised control trial evaluation of Family Care, an Australian home visiting programme delivered by nurses in families where the child was at risk of poor health and developmental outcomes. At six weeks postpartum, primiparous women in the intervention group showed a significant reduction in postnatal depression, which was not observed for multiparous women. In addition, some studies find evidence which suggests that first time mothers are less likely to drop out of home visiting interventions than other participants (DuMont et al., 2008). Based on the literature, we hypothesised that the *PFL* programme would have a slightly stronger impact on primiparous women than on multiparous women.

5.4.1 Parity Results

Interaction analysis was used to explore differential treatment effects by maternal parity status. These results are presented in Table 5.3. The interaction of parity and treatment status was significant in 19% (6/32) of the child development outcomes. Thus there was some evidence that the programme worked differently for the children of first time and non-first time mothers in respect to the ASQ personal-social score, the CBCL internal problems score, the CBCL measures of withdrawn behaviour and somatic complaints, the SDQ peer problems score, and the ASQ standardised total score.

Table 5.4 presents separate subgroup analyses for first time mothers (46% of the sample) and non-first time mothers (54% of the sample) to further explore the nature of these differential effects. The results indicate that the programme had more effects on children of first time mothers than children of multiparous mothers. First born children in the high treatment group performed better than those in the low treatment group on 83% of the six measures analysed, while the children of multiparous women in the high treatment group did not score significantly higher than those in the low treatment group on any of the measures. First born children in the high treatment group exhibited higher levels of personal social skills, as measured by the ASQ ($p<.05$, $d=0.46$). In terms of their behaviour, as indicated by the CBCL, they were less likely to experience peer problems ($p<.05$, $d=0.49$). Within the CBCL subdomains, they were less likely to be withdrawn ($p<.10$, $d=0.37$) or have somatic complaints ($p<.10$, $d=0.36$). The children of primiparous women also exhibited significant treatment effects on the ASQ standardised score ($p<.05$, $d=0.44$).

Table 5.3 - Results for Linear Regression Analysis of Parity on Child Development Outcomes

Variable	N	Treatment		Parity		Treatment *Parity	
ASQ Scores							
ASQ Communication Score	145	-1.30	(1.99)	-2.53	(2.09)	4.55	(2.93)
ASQ Gross Motor Score	145	-1.29	(1.95)	-3.39	(2.04)	2.67	(2.87)
ASQ Fine Motor Score	145	1.32	(2.94)	0.33	(3.09)	4.80	(4.33)
ASQ Problem Solving Score	145	-0.84	(2.09)	-1.31	(2.19)	2.91	(3.08)
ASQ Personal-Social Score	145	-0.45	(2.48)	-7.97***	(2.61)	6.24*	(3.66)
* ASQ Social-Emotional Score	145	-0.46	(6.58)	4.59	(6.91)	0.27	(9.70)
ASQ cut-off scores							
* ASQ Communication cut-off	145	-0.01	(0.05)	-0.00	(0.05)	0.01	(0.08)
* ASQ Gross Motor cut-off	145	0.02	(0.04)	0.09**	(0.04)	-0.09	(0.06)
* ASQ Fine Motor cut-off	145	-0.05	(0.09)	0.04	(0.09)	-0.15	(0.13)
* ASQ Problem Solving cut-off	145	0.00	(0.05)	0.01	(0.05)	-0.01	(0.07)
* ASQ Personal-Social cut-off	145	0.02	(0.04)	0.09**	(0.04)	-0.09	(0.06)
* ASQ Social-Emotional cut-off	145	-0.02	(0.05)	0.04	(0.06)	0.03	(0.08)
Child Behaviour Checklist (CBCL) domains							
* CBCL Internal Problems	144	1.14	(1.41)	3.59**	(1.49)	-3.80*	(2.09)
* CBCL External Problems	144	-0.94	(1.63)	3.17*	(1.72)	-1.30	(2.41)
* CBCL Total Score	144	0.11	(4.39)	9.58**	(4.66)	-7.39	(6.50)
Child Behaviour Checklist (CBCL) domains cut-off scores							
* CBCL Internal Problems cut-off	144	-0.04	(0.06)	0.09	(0.06)	-0.06	(0.08)
* CBCL External Problems cut-off	144	-0.02	(0.03)	0.07*	(0.03)	-0.07	(0.05)
* CBCL Total Score cut-off	144	0.00	(0.04)	0.10**	(0.05)	-0.10	(0.07)
Child Behaviour Checklist (CBCL) subdomains							
* CBCL Emotionally Reactive	144	0.21	(0.47)	0.62	(0.50)	-0.62	(0.70)
* CBCL Anxious/Depressed	144	-0.04	(0.47)	0.55	(0.49)	-0.58	(0.69)
* CBCL Somatic Complaints	144	0.51	(0.42)	1.52***	(0.45)	-1.34**	(0.63)
* CBCL Withdrawn	144	0.45	(0.42)	0.89**	(0.44)	-1.24**	(0.62)
* CBCL Sleep Problems Behaviour	144	-0.27	(0.51)	0.30	(0.54)	0.08	(0.75)
* CBCL Attention Problems	144	0.23	(0.41)	0.96**	(0.43)	-0.88	(0.61)
* CBCL Aggressive Behaviour	144	-1.17	(1.34)	2.21	(1.42)	-0.41	(1.98)
* CBCL Other Problems	144	0.19	(1.25)	2.51*	(1.33)	-2.37	(1.86)
Strengths and Difficulties Questionnaire (SDQ)							
* SDQ Peer Problems	145	0.316	(0.33)	0.66*	(0.34)	-1.01**	(0.48)
SDQ Pro-Social Behaviour	145	0.032	(0.39)	-0.32	(0.41)	0.61	(0.57)
Other Measures							
ASQ Standardised Total Score	145	-1.46	(3.60)	-6.83*	(3.78)	9.20*	(5.31)
DP-3: Cognitive Development standardised score	145	2.65	(3.37)	-1.31	(3.54)	5.74	(4.97)
DP-3: Cognitive Development above average cut-off	145	0.12	(0.10)	0.05	(0.10)	0.03	(0.14)
* Child receiving special services	145	-0.00	(0.09)	-0.00	(0.09)	0.00	(0.13)

Notes: Standard errors in brackets. ***$p<.01$, **$p<.05$, *$p<.10$

Table 5.4 - Subgroup Analysis Results for High and Low Treatment Group by Maternal Parity Status

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Effect Size d
	ASQ Personal Social Score								
	Primiparous	67	(36/31)	51.11	(12.08)	45.32	(13.66)	$p<.05$	0.46
	Multiparous	78	(37/41)	52.84	(9.76)	53.29	(8.41)	ns	0.05
*	SDQ Peer Problems								
	Primiparous	67	(36/31)	1.14	(1.25)	1.84	(1.66)	$p<.05$	0.49
	Multiparous	78	(37/41)	1.49	(1.57)	1.17	(1.36)	ns	0.22
*	CBCL Internal Problems								
	Primiparous	66	(36/30)	6.28	(4.93)	8.93	(9.63)	ns	0.36
	Multiparous	78	(37/41)	6.49	(5.06)	5.34	(4.94)	ns	0.23
*	CBCL Withdrawn								
	Primiparous	66	(36/30)	-1.08	(1.44)	-1.87	(2.76)	$p<.10$	0.37
	Multiparous	78	(37/41)	-1.43	(1.99)	-0.98	(1.13)	ns	0.99
*	CBCL Somatic Complaints								
	Primiparous	66	(36/30)	-1.47	(1.8)	-2.3	(2.87)	$p<.10$	0.36
	Multiparous	78	(37/41)	-1.30	(1.68)	-0.78	(1.08)	ns	0.94
	ASQ Standardised Total Score								
	Primiparous	67	(36/31)	101.63	(16.62)	93.88	(18.84)	$p<.05$	0.44
	Multiparous	78	(37/41)	99.26	(15.78)	100.72	(12.65)	ns	0.10

Notes: '*N*' indicates the sample size. '*M*' indicates the mean. '*SD*' indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test.

5.5 Cognitive Resources

Across multiple studies, maternal cognitive functioning has consistently been identified as a powerful predictor of child cognitive functioning including outcomes in maths, reading, and language (Black et al., 2007; Cornelius et al., 2009; Crane, 1996; Longstreth et al., 1981; Sommer et al., 2000; Tong et al., 2007). Maternal cognitive resources contribute directly to child functioning through genetic transmission (Haworth et al., 2010), but may also have an indirect effect on the wider area of child development through parenting behaviour. For example, mothers with higher cognitive resources are more likely to breastfeed (Der, Batty, & Deary, 2006), less likely to smoke (Kubicka et al., 2001), provide better quality home environments (Burchinal et al., 1997), are more satisfied in their parenting role (Bornstein et al., 2003), and their children have better diets (Wachs & McCabe, 2001). It also seems that these non-genetic influences on cognition may impact differentially depending on family SES. Turkheimer et al. (2003) report that about 60% of the variance in cognitive resources in low SES families is attributable to the environment, with limited genetic impact. Conversely, in high SES families, the majority of the variance is attributable to genetics and the environment has little impact.

Early intervention may provide a means to mitigate the impact of poor maternal cognitive resources on children, yet relatively little is known about the benefits of home visiting programmes for parents with different abilities. A series of studies evaluating the Nurse Family Partnership programme identified differentially favourable outcomes for children of mothers with low cognitive resources (Olds, 2002; Olds et al., 2007; Olds, Kitzman et al., 2004; Olds, Robinson et al., 2004). These studies assessed cognitive resources as part of a composite variable which included limited cognitive functioning, mental health, and sense of control. Results indicated that, at twenty-four months, the children of mothers with low cognitive

resources who received the intervention were more responsive and communicative and had better physical health than those in the control group (Olds, 2002). Moreover, their children had higher gains in language, executive functioning, and behavioural adaptation at forty-eight months (Olds, Robinson et al., 2004). The children also demonstrated lower levels of aggression at age six (Olds, Kitzman et al., 2004), and better academic ability at ages six, nine, and 12 years than their control group counterparts (Kitzman et al., 2010; Olds, Kitzman et al., 2004; Olds et al., 2007). This research suggests that home visiting interventions may be particularly beneficial for children whose parents have low cognitive resources. Based on the literature, we hypothesised that the *PFL* programme would also be of particular benefit to children of parents with lower cognitive resources.

5.5.1 Cognitive Resources Results

To investigate whether the programme impacted mothers with relatively high cognitive resources and mothers with relatively low cognitive resources differently, an interaction and subgroup analysis by cognition was conducted for the child development domain.

To gain an index of maternal cognition, the Wechsler Abbreviated Scale of Intelligence (WASI) cognitive assessment was administered to all mothers participating in the study when their baby was approximately three months old. The WASI is a short, four-subset version of the Wechsler Adult Intelligence Scale (WAIS) which focuses on such domains as vocabulary, similarities of constructs, block design, and matrix reasoning. The assessment was administered by a trained assessor and took approximately 45 minutes to complete. The WASI provides standardised measures of verbal, performance, and a full scale measure of cognitive functioning. To conduct the subgroup analysis, the full scale measure was dichotomised to create an indicator that represented mothers with relative higher cognitive resources and mothers with relatively lower cognitive resources. The dichotomisation was based on scoring above (47% of the sample) or below (53% of the sample) the median score (84 points) within the sample.

The findings for the interaction analyses, exploring differential treatment effects by maternal cognitive resources, are presented in Table 5.5. The interaction of maternal cognitive resources and treatment status was significant for 22% (7/32) of the measures, providing some evidence that the programme impacted upon the children of mothers of higher and lower cognitive resources in different ways. Differential treatment effects were evident regarding the children's personal-social and gross motor skills, their CBCL domain cut-off scores, and with respect to withdrawn behaviour.

Subgroup analysis further investigating these differential effects is presented in Table 5.6. Overall, the results indicate that the *PFL* programme is of particular benefit to children of mothers with lower cognitive resources. These children scored significantly higher than those in the low treatment group on 86% (6/7) of the individual permutation tests, while the children of mothers with higher cognitive resources in the high treatment group did not score significantly higher than those in the low treatment group on any of the measures. Children of low cognitive resource mothers in the high treatment group were less likely to be within the range for internalising problems ($p<.05$, $d=0.56$), externalising problems ($p<.05$, $d=0.46$), and total scores ($p<.10$, $d=0.37$), as measured by the CBCL. The step-down tests indicated that the joint effect of the CBCL domains cut-off were significant and driven by all three measures. The children of low cognitive resource mothers were significantly more likely to display better personal-social skills ($p<.05$, $d=0.50$) and gross motor skills ($p<.10$, $d=0.37$), and less likely to fall within the cut-off for poor gross motor skills as measured by the ASQ ($p<.10$, $d=0.39$). In contrast, the children of higher cognitive resource mothers in the high treatment group were significantly less likely to score higher than their low treatment counterparts on the gross motor skills outcome. Thus, the children of mothers with lower cognitive resources appear to derive greater benefits from the programme.

Table 5.5 - Results for Linear Regression Analysis of Maternal Cognitive Resources on Child Development Outcomes.

Variable	N	Treatment		Cognitive Resources		Treatment *Cognitive	
ASQ scores							
ASQ Communication Score	147	0.46	(2.00)	0.23	(2.08)	0.57	(2.92)
ASQ Gross Motor Score	147	3.18*	(1.91)	4.76**	(1.99)	-7.20**	(2.79)
ASQ Fine Motor Score	147	2.88	(2.94)	2.07	(3.06)	0.90	(4.30)
ASQ Problem Solving Score	147	1.21	(2.09)	1.15	(2.17)	-1.56	(3.05)
ASQ Personal-Social Score	147	5.09**	(2.52)	3.30	(2.62)	-6.41*	(3.69)
* ASQ Social-Emotional Score	147	1.12	(6.39)	-14.7**	(6.64)	0.58	(9.33)
ASQ cut-off scores							
* ASQ Communication cut-off	147	-0.01	(0.05)	-0.01	(0.05)	0.01	(0.08)
* ASQ Gross Motor cut-off	147	-0.07*	(0.04)	-0.07*	(0.04)	0.12**	(0.06)
* ASQ Fine Motor cut-off	147	-0.18**	(0.09)	-0.13	(0.09)	0.13	(0.13)
* ASQ Problem Solving cut-off	147	-0.04	(0.05)	-0.04	(0.05)	0.09	(0.07)
* ASQ Personal-Social cut-off	147	-0.04	(0.04)	-0.01	(0.04)	0.07	(0.06)
* ASQ Social-Emotional cut-off	147	-0.01	(0.05)	-0.12**	(0.05)	0.04	(0.08)
Child Behaviour Checklist (CBCL) domains							
* CBCL Internal Problems	146	-1.77	(1.40)	-4.19***	(1.46)	3.30	(2.05)
* CBCL External Problems	146	-1.42	(1.66)	-1.65	(1.73)	0.65	(2.42)
* CBCL Total Score	146	-5.21	(4.41)	-9.63**	(4.59)	6.41	(6.43)
Child Behaviour Checklist (CBCL) domains cut-off scores							
* CBCL Internal Problems cut-off	146	-0.17***	(0.05)	-0.20***	(0.06)	0.22***	(0.08)
* CBCL External Problems cut-off	146	-0.10***	(0.03)	-0.10***	(0.03)	0.10*	(0.05)
* CBCL Total Score cut-off	146	-0.09**	(0.04)	-0.12**	(0.05)	0.12*	(0.07)
Child Behaviour Checklist (CBCL) subdomains							
* CBCL Emotionally Reactive	146	-0.31	(0.47)	-1.01**	(0.49)	0.72	(0.69)
* CBCL Anxious/Depressed	146	-0.63	(0.45)	-1.63***	(0.47)	0.92	(0.66)
* CBCL Somatic Complaints	146	-0.29	(0.44)	-0.75	(0.45)	0.61	(0.64)
* CBCL Withdrawn	146	-0.53	(0.42)	-0.79*	(0.44)	1.03*	(0.62)
* CBCL Sleep Problems Behaviour	146	-0.44	(0.50)	-1.07**	(0.52)	0.58	(0.73)
* CBCL Attention Problems	146	-0.36	(0.41)	-0.71	(0.43)	0.63	(0.60)
* CBCL Aggressive Behaviour	146	-1.06	(1.37)	-0.94	(1.42)	0.02	(1.99)
* CBCL Other Problems	146	-1.56	(1.25)	-2.70**	(1.30)	1.86	(1.82)
Strengths and Difficulties Questionnaire (SDQ)							
* SDQ Peer Problems	147	-0.30	(0.33)	-0.47	(0.35)	0.49	(0.49)
SDQ Pro-Social Behaviour	147	0.11	(0.39)	0.23	(0.40)	0.32	(0.57)
Other Measures							
ASQ Standardised Total Score	147	5.55	(3.61)	5.21	(3.75)	-6.58	(5.27)
DP-3: Cognitive Development standardised score	147	6.45*	(3.30)	7.84**	(3.43)	-3.22	(4.82)
DP-3: Cognitive Development above average cut-off	147	0.12	(0.09)	0.10	(0.10)	0.03	(0.14)
* Child receiving special services	147	-0.00	(0.090)	0.04	(0.093)	0.00	(0.13)

Notes: Standard errors in brackets. ***p<.01, **p<.05, *p<.10

Table 5.6 - Subgroup Analysis Results for High and Low Treatment Group by Maternal Cognitive Resources

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step -down Test p^2	Effect Size *d*
	ASQ Personal-Social Score									
	Lower Cognitive Resources	78	(37/41)	53.51	(7.81)	48.41	(12.27)	p<.05	-	0.50
	Higher Cognitive Resources	69	(37/32)	50.41	(13.14)	51.72	(10.37)	ns	-	0.11
	ASQ Gross Motor Score									
	Lower Cognitive Resources	78	(37/41)	55.14	(6.82)	51.95	(9.93)	p<.10	-	0.37
	Higher Cognitive Resources	69	(37/32)	52.70	(9.90)	56.72	(5.77)	s~	-	0.49
*	ASQ Gross Motor cut-off									
	Lower Cognitive Resources	78	(37/41)	0.00	(0.00)	0.07	(0.26)	p<.10	-	0.39
	Higher Cognitive Resources	69	(37/32)	0.05	(0.23)	0.00	(0.00)	ns	-	0.33
	Lower Cognitive Resources									
*	CBCL Internal Problems cut-off	77	(37/40)	0.03	(0.16)	0.20	(0.41)	p<.05	p<.05	0.56
*	CBCL External Problems cut-off	77	(37/40)	0.00	(0.00)	0.10	(0.30)	p<.05	p<.05	0.46
*	CBCL Total Score cut-off	77	(37/40)	0.03	(0.16)	0.13	(0.33)	p<.10	p<.10	0.37
	Higher Cognitive Resources									
*	CBCL External Problems cut-off	69	(37/32)	0.00	(0.00)	0.00	(0.00)	-	-	-
*	CBCL Total Score cut-off	69	(37/32)	0.03	(0.16)	0.00	(0.00)	ns	ns	0.23
*	CBCL Internal Problems cut-off	69	(37/32)	0.05	(0.23)	0.00	(0.00)	ns	ns	0.33
*	CBCL Withdrawn									
	Lower Cognitive Resources	77	(37/40)	1.16	(1.46)	1.70	(2.38)	ns	-	0.27
	Higher Cognitive Resources	69	(37/32)	1.41	(1.99)	0.91	(1.33)	ns	-	0.29

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. 'p<.01', 'p<.05' and 'p<.10' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

5.6 Triple P

The Triple P Positive Parenting Programme promotes healthy parenting practices and positive attachment relationships between parents and children (Sanders et al., 2003). The programme was offered to all high treatment group participants when the *PFL* child was at least two years old. Four different types of Triple P were offered to participants, covering three of the four official Triple P levels: Selected Triple P (seminar series) (Level 2), Triple P Discussion Groups (Level 3), Primary Care (Level 3), and Group Triple P (Level 4).

Triple P has been extensively evaluated internationally using diverse experimental and non-experimental designs (Leung et al., 2003; Martin & Sanders, 2003; Nowak & Heinrichjs, 2008; Prinz et al., 2009; Sanders et al., 2014; UN Office on Drugs & Crime, 2009). These evaluations have highlighted both short-term and long-term positive impacts on children's social, emotional and behavioural outcomes, and on parenting practices, parenting satisfaction and efficacy, parental adjustment, and parental relationships (Sanders et al., 2014). Based on the literature, we hypothesised that the programme would have a stronger impact on parenting outcomes and child development outcomes for those who took part in Triple P than those who did not.

5.6.1 Triple P Results

To evaluate the impact of the Triple P programme on child development and parenting outcomes, separate analyses were conducted for participants who took part in Triple P and those who did not. Note that, consistent with the programme design, none of the low treatment group participated in the Triple P programme, therefore the analyses below compared those in the high treatment group who took part in Triple P (62% of the forty-eight month sample) to those in the high treatment group who did not take part in Triple P (38% of the forty-eight month sample). To further explore the significant differences that arose in this analysis, those in the high treatment group who participated in Triple P were first compared to the low treatment group, then, those in the high treatment group who did not participate in Triple P were compared to the low treatment group. As not all eligible parents in the high treatment group choose to participate in Triple P, these analyses are inherently limited by any potential selection of the high treatment group into the Triple P programme.

CHILD DEVELOPMENT

The results in Table 5.7 present the impact of Triple P within the high treatment group only. The findings indicate that parental participation in Triple P had a limited positive effect on children's developmental outcomes. Significant differences were evident for two outcomes only, both concerning the children's fine motor skills. To further investigate these results, Table 5.8 reports the results of separate subgroup analysis. The children of Triple P participants in the high treatment group scored significantly higher than the children from the low treatment group on the ASQ measure of fine motor skills ($p<.10$, $d=0.56$) and were less likely to be within the cut-off range for poor fine motor skills ($p<.10$, $d=0.52$). In contrast, children from the high treatment group whose parents did not participate in Triple P did not outperform their low treatment counterparts on either measure. As such, there is limited evidence that the Triple P programme had a positive impact on child development outcomes.

Table 5.7 - Results for Triple P Participants and Non-Participants in the High Treatment Group: Child Development

	Variable	N	($n_{TRIPLEP}$/ $n_{NONTRIPLEP}$)	$M_{TRIPLE\ P}$	(SD)	$M_{NON\ TRIPLE\ P}$	(SD)	Individual Test p^1	Step -down Test p^2	Effect Size d
ASQ scores										
	ASQ Fine Motor Score	74	(46/28)	50.54	(9.62)	41.96	(14.80)	$p<.01$	$p<.01$	0.74
	ASQ Communication Score	74	(46/28)	54.35	(7.93)	53.21	(8.63)	ns	ns	0.14
	ASQ Gross Motor Score	74	(46/28)	54.24	(8.88)	53.39	(8.06)	ns	ns	0.10
	ASQ Problem Solving Score	74	(46/28)	53.70	(9.74)	53.04	(7.86)	ns	ns	0.07
*	ASQ Social-Emotional Score	74	(46/28)	-33.80	(28.58)	-32.14	(26.37)	ns	ns	0.06
	ASQ Personal-Social Score	74	(46/28)	51.63	(11.50)	52.50	(9.86)	ns	ns	0.08
ASQ cut-off scores										
*	ASQ Fine Motor cut-off	74	(46/28)	-0.07	(0.25)	-0.25	(0.44)	$p<.05$	$p<.10$	0.56
*	ASQ Communication cut-off	74	(46/28)	-0.04	(0.21)	-0.07	(0.26)	ns	ns	0.12
*	ASQ Personal-Social cut-off	74	(46/28)	-0.02	(0.15)	-0.04	(0.19)	ns	ns	0.09
*	ASQ Social-Emotional cut-off	74	(46/28)	-0.07	(0.25)	-0.07	(0.26)	ns	ns	0.02
*	ASQ Problem Solving cut-off	74	(46/28)	-0.07	(0.25)	-0.04	(0.19)	ns	ns	0.13
*	ASQ Gross Motor cut-off	74	(46/28)	-0.04	(0.21)	0.00	(0.00)	ns	ns	0.27
Strengths and Difficulties Questionnaire (SDQ)										
*	SDQ Peer Problems	74	(46/28)	-1.22	(1.36)	-1.61	(1.64)	ns	ns	0.27
	SDQ Pro-Social Behaviour	74	(46/28)	8.54	(1.56)	8.36	(1.66)	ns	ns	0.12
Child Behaviour Checklist (CBCL) domains										
*	CBCL Total Score	74	(46/28)	-20.98	(13.51)	-24.07	(16.12)	ns	ns	0.22
*	CBCL Internal Problems	74	(46/28)	-6.13	(4.70)	-7.11	(5.56)	ns	ns	0.20
*	CBCL External Problems	74	(46/28)	-7.15	(5.75)	-7.93	(5.53)	ns	ns	0.14
Child Behaviour Checklist (CBCL) domains cut-off scores										
*	CBCL External Problems cut-off	74	(46/28)	0.00	(0.00)	0.00	(0.00)	-	-	-
*	CBCL Total Score cut-off	74	(46/28)	-0.02	(0.15)	-0.04	(0.19)	ns	ns	0.09
*	CBCL Internal Problems cut-off	74	(46/28)	-0.04	(0.21)	-0.04	(0.19)	ns	-	0.04
Child Behaviour Checklist (CBCL) sub-domains										
*	CBCL Attention Problems	74	(46/28)	-1.59	(1.42)	-2.00	(1.52)	ns	ns	0.29
*	CBCL Other Problems	74	(46/28)	-5.63	(4.03)	-6.79	(4.047)	ns	ns	0.28
*	CBCL Withdrawn	74	(46/28)	-1.13	(1.74)	-1.54	(1.73)	ns	ns	0.24
*	CBCL Emotionally Reactive	74	(46/28)	-1.13	(1.75)	-1.53	(1.73)	ns	ns	0.21
*	CBCL Anxious/Depressed	74	(46/28)	-1.72	(1.73)	-2.11	(2.33)	ns	ns	0.20
*	CBCL Sleep Problems Behaviour	74	(46/28)	-2.07	(1.78)	-2.25	(2.27)	ns	ns	0.09
*	CBCL Aggressive Behaviour	74	(46/28)	-5.57	(4.83)	-5.92	(4.74)	ns	ns	0.08
*	CBCL Somatic Complaints	74	(46/28)	-1.46	(1.93)	-1.29	(1.33)	ns	ns	0.10
Non Step-down Measures										
	ASQ Standardised Total Score	74	(46/28)	101.95	(16.78)	98.15	(14.73)	ns	-	0.24
	DP-3: Cognitive Development standardised score	74	(46/28)	108.87	(13.34)	107.11	(14.78)	ns	-	0.13
	DP-3: Cognitive Development above average cut-off	74	(46/28)	0.35	(0.48)	0.32	(0.48)	ns	-	0.06
*	Child receiving special services	74	(46/28)	-0.20	(0.40)	-0.18	(0.39)	ns	-	0.04

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

Table 5.8 - Results for High Treatment Group Triple P Participants/Non-Participants compared to the Low Treatment group: Child Development

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p[1]	Effect Size d
	ASQ Fine Motor Score								
	Participants	119	(46/73)	50.54	(9.62)	43.84	(13.48)	p<.01	0.56
	Non-Participants	101	(28/73)	41.96	(14.80)	43.84	(13.48)	ns	0.14
*	ASQ Fine Motor cut-off								
	Participants	119	(46/73)	-0.07	(0.25)	-0.26	(0.44)	p<.01	0.52
	Non-Participants	101	(28/73)	-0.25	(0.44)	-0.26	(0.44)	ns	0.02

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. p<.01, p<.05 and p<.10 indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

PARENTING

The findings for parenting outcomes according to Triple P status within the high treatment group are presented in Table 5.9. The results suggest that parental participation in Triple P had a limited positive impact on parenting outcomes. The high treatment group who took part in Triple P outperformed the non-participants on 19% (6/32) of the individual outcomes. Triple P participants were less likely than non-participants to display an authoritative parenting style as measured by the PSDQ (p<.01, d=0.61). They scored significantly higher on the PSDQ subdomain measures of regulation (p<.01, d=0.54), connection (p<.05, d=0.45), and autonomy (p<.10, d=0.41). Step-down tests of these measures indicated a significant joint effect driven by all three measures. Triple P participants were also less likely to have a high score for coercion within the authoritative parenting subdomain (p<.10, d=0.33). They were also more likely to view social competence as an important factor of school readiness (p<.05, d=0.61).

To explore these parenting results further, separate subgroup analyses examining the outcomes for Triple P participants and non-participants in the high treatment group to the low treatment group was conducted and is reported in Table 5.10. High treatment participants and non-participants did not differ significantly from low treatment participants on any of the six measures that were significantly different within the high treatment group. Therefore, the results suggest that the benefit of programme participation in Triple P on parenting outcomes was slight and inconsistent across the subgroups.

Table 5.9 - Results for Triple P Participants and Non-Participants in the High Treatment Group: Parenting

Variable	N	($n_{TRIPLEP}$/ $n_{NONTRIPLEP}$)	$M_{TRIPLE\ P}$ (*SD*)		$M_{NON\ TRIPLE\ P}$ (*SD*)		Individual Test p^1	Step -down Test p^2	Effect Size *d*
Parenting Daily Hassles (PDH)									
* PDH Parenting Tasks Score	74	(46/28)	-11.11	(3.61)	-12.18	(4.22)	ns	ns	0.28
* PDH Intensity Scale Score	74	(46/28)	-30.74	(10.68)	-32.96	(12.11)	ns	ns	0.20
* PDH Challenging Behaviour Score	74	(46/28)	-12.17	(4.90)	-12.68	(5.61)	ns	ns	0.10
* PDH Frequency Scale Score	74	(46/28)	-33.17	(7.02)	-33.79	(6.30)	ns	ns	0.09
Parenting Stress Index (PSI)									
* Parent-Child Dysfunctional Interactions	74	(46/28)	-17.65	(5.41)	-19.43	(7.03)	ns	ns	0.30
* Parental Distress	74	(46/28)	-23.37	(9.08)	-24.82	(6.69)	ns	ns	0.18
* Difficult Child	74	(46/28)	-22.5	(7.62)	-22.18	(7.44)	ns	ns	0.04
Parenting Styles and Dimensions Questionnaire (PSDQ)									
Authoritative Parenting	74	(46/28)	4.18	(0.49)	3.89	(0.51)	$p<.01$	$p<.05$	0.61
* Authoritarian Parenting	74	(46/28)	-1.52	(0.46)	-1.64	(0.49)	ns	ns	0.26
* Permissive Parenting	74	(46/28)	-2.26	(0.82)	-2.25	(0.70)	ns	ns	0.01
PSDQ Authoritative Parenting Subdomains									
PSDQ Regulation	74	(46/28)	4.07	(0.77)	3.67	(0.69)	$p<.01$	$p<.05$	0.54
PSDQ Connection	74	(46/28)	4.75	(0.34)	4.58	(0.46)	$p<.05$	$p<.10$	0.45
PSDQ Autonomy	74	(46/28)	3.73	(0.75)	3.41	(0.85)	$p<.05$	$p<.05$	0.41
PSDQ Authoritarian Parenting Subdomains									
* PSDQ Coercion	74	(46/28)	-1.32	(0.46)	-1.47	(0.51)	$p<.10$	ns	0.33
* PSDQ Hostility	74	(46/28)	-1.59	(0.53)	-1.68	(0.56)	ns	ns	0.17
* PSDQ Punitive	74	(46/28)	-1.65	(0.72)	-1.76	(0.64)	ns	ns	0.16
Parental Perception of Important School Readiness Traits									
Social Competence	71	(43/28)	0.74	(0.44)	0.46	(0.51)	$p<.05$	$p<.10$	0.61
Communication and General Knowledge	70	(43/27)	0.23	(0.43)	0.11	(0.32)	ns	ns	0.32
Physical Health and Wellbeing	70	(43/27)	0.40	(0.49)	0.30	(0.47)	ns	ns	0.21
Emotional Maturity	69	(42/27)	0.24	(0.43)	0.22	(0.42)	ns	ns	0.04
Language and Cognitive Development	71	(44/27)	0.32	(0.47)	0.37	(0.49)	ns	ns	0.11
Other Skills	69	(42/27)	0.07	(0.26)	0.15	(0.36)	ns	ns	0.26
TV Habits									
Mother talks to child about TV	67	(41/26)	1.00	(0.00)	0.96	(0.20)	ns	ns	0.32
* Maximum TV time allowed per day	46	(29/17)	-2.84	(1.6)	-3.18	(1.38)	ns	ns	0.22
TV/videos/DVDs per day (hours)	74	(46/28)	-2.36	(1.58)	-2.58	(1.48)	ns	ns	0.15
* Time TV is on in the home (hours)	74	(46/28)	-8.18	(4.08)	-8.57	(3.85)	ns	ns	0.10
Time spent watching TV with child (hours)	67	(41/26)	1.45	(1.08)	1.33	(1.39)	ns	ns	0.10
Child's TV time limited	74	(46/28)	0.63	(0.49)	0.61	(0.50)	ns	ns	0.05
* Time spent by child watching TV alone (hours)	74	(46/28)	-0.97	(1.2)	-0.90	(0.89)	ns	ns	0.07
Non Step-down Measures									
Age started school (months)	5	(3/2)	51.67	(3.21)	49.50	(0.71)	ns	-	1.05
* PSI Total Stress Score	74	(46/28)	-63.52	(19.52)	-66.43	(18.97)	ns	-	0.15
Child is on primary school waiting list	69	(43/26)	0.77	(0.43)	0.73	(0.45)	ns	-	0.08
Child has started primary school	74	(46/28)	0.07	(0.25)	0.07	(0.26)	ns	-	0.02
Worried about child's behaviour	74	(46/28)	0.13	(0.34)	0.14	(0.36)	ns	-	0.04
* PSI Stress cut-off	74	(46/28)	-0.09	(0.28)	-0.07	(0.26)	ns	-	0.06
Worried about child's language development	74	(46/28)	0.13	(0.34)	0.18	(0.39)	ns	-	0.14

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. 'ns' indicates the variable is not statistically significant. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. 's~' indicates that the variable was significant in a left-sided test. The variables are reported in order of the largest to the smallest T statistic within each Step-down category. [3] Indicates that the step-family was jointly significant in a left-sided test.

Table 5.10 - Results for High Treatment Group Triple P Participants/Non-Participants compared to the Low Treatment Group: Parenting

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p[1]	Effect Size d
	Authoritative Parenting								
	Participants	119	(46/73)	4.18	(0.49)	4.06	(0.59)	ns	0.23
	Non-Participants	101	(28/73)	3.89	(0.51)	4.06	(0.59)	ns	0.31
	PSDQ Connection								
	Participants	119	(46/73)	4.75	(0.34)	4.65	(0.47)	ns	0.23
	Non-Participants	101	(28/73)	4.58	(0.46)	4.65	(0.47)	ns	0.16
	PSDQ Autonomy								
	Participants	119	(46/73)	3.73	(0.75)	3.64	(0.79)	ns	0.12
	Non-Participants	101	(28/73)	3.41	(0.85)	3.64	(0.79)	ns	0.28
	PSDQ Regulation								
	Participants	119	(46/73)	4.07	(0.77)	3.89	(0.80)	ns	0.23
	Non-Participants	101	(28/73)	3.67	(0.69)	3.89	(0.80)	ns	0.28
*	PSDQ Coercion								
	Participants	119	(46/73)	-1.32	(0.46)	-1.37	(0.43)	ns	0.13
	Non-Participants	101	(28/73)	-1.47	(0.51)	-1.37	(0.43)	ns	0.22
	Social Competence								
	Participants	111	(43/68)	0.74	(0.44)	0.71	(0.46)	ns	0.09
	Non-Participants	96	(28/68)	0.46	(0.51)	0.71	(0.46)	ns	0.52

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] one-tailed (right-sided) *p* value from an individual permutation test with 100,000 replications. [2] one-tailed (right-sided) *p* value from a Step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. $p<.01$, $p<.05$ and $p<.10$ indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively.

Overall, Triple P participants slightly outperformed non-participants, however the difference between the two groups was relatively small. There were some differences between participants and non-participants in high and low treatment groups regarding child development. However, differences in parenting that were evident among participants and non-participants in the high treatment group were not evidenced in comparisons to the low treatment group.

5.7 Summary

The results of the interaction and subgroup analysis are summarised in Table 5.11. Overall, the analyses indicate that the *PFL* programme had some differential impacts on the participants depending on their characteristics. Regarding child development, the programme was most beneficial for the children of first time mothers and the children of mothers with lower cognitive resources. The programme did not affect participants differently based on child gender. Triple P participants in the high treatment group performed slightly better than non-Triple P participants in the high treatment group, but there was limited evidence of differences when compared to the low treatment group.

These results for the most part reflect expectations based on the literature. In keeping with the evidence, first time mothers obtained the greatest benefit from early intervention (Howard & Brooks Gunn, 2009). In other areas, the evidence is somewhat limited and thus expectations were unclear. In relation to gender, the results were not consistent with evidence from some other early intervention programmes (e.g. Anderson, 2008) as the programme was only marginally more effective at improving the development of boys than girls. The programme also impacted mothers differently according to their cognitive resources. Consistent with findings from other studies (Olds et al., 2002), the children of mothers with low cognitive resources benefitted more from the programme in terms of their development than the children of mothers with higher cognitive resources. In relation to Triple P, previous meta-analytic examinations of Triple P have found favourable effects in short (two month) to long-term (thirty-six month) studies (Sanders et al., 2014). The results from our evaluation found few differential effects by Triple P attendance, and were more in line with results from a single study by Eisner et al. (2012) which found no consistent effects of Triple P participation on parenting and child behaviour. The implications of these differences will be discussed in Chapter 6.

Table 5.11 - Summary of Subgroup Analyses

	Child Development								Parenting	
	Gender		Parity Status		Cognitive Resources		Triple P			
Significant differences indicated by interaction analysis	1/32 (3%)		6/32 (19%)		7/32 (22%)					
High treatment group: permutation analysis							2/32 (6%)		6/36 (17%)	
	Girls	Boys	Primiparous	Multiparous	High	Low	Participants	Non-Participants	Participants	Non-Participants
Follow-up subgroup analysis: Significant differences between high and low treatment groups	0/1 (0%)	1/1 (100%)	5/6 (83%)	0/6 (0%)	0/7 (0%)	6/7 (86%)	2/2 (100%)	0/2 (0%)	0/6 (0%)	0/6 (0%)

Chapter Six

Report Summary & Conclusions

6.1 Overview

This report presented the results of the effectiveness of the *Preparing For Life* programme between programme entry and when the *PFL* children were approximately forty-eight months of age. This report is the last to present findings based on interviews conducted with the *PFL* parents. It includes an analysis of the quantitative information derived from interviews with *PFL* participants and implementation data from *PFL*'s database. At this time point, the results differed from studies of other home visiting programmes. Based on the literature, we hypothesised that there would be moderate treatment effects in the areas of child development, parenting, maternal health and wellbeing, and household factors and SES, with limited effects on the other domains: child health, social support, the home environment, and childcare. Yet the main findings at forty-eight months were relatively limited across most domains. The section below summarises the results of the various analyses including the main results comparing the high and low treatment groups, the findings when inverse probability weighting was applied, and the results of the implementation, dynamic, and interaction and subgroup analyses.

OVERVIEW OF TREATMENT EFFECTS AT FORTY-EIGHT MONTHS

In total, 217 forty-eight month interviews were completed. The analyses focused on eight domains, incorporating 191 outcome measures. At forty-eight months we saw a drop in the number of positive findings compared to the previous time point at thirty-six months. Notably, findings in the areas of child development, parenting, maternal health and wellbeing, and household factors and SES were weaker than anticipated based on the literature. As hypothesised, findings in the other four domains were limited, although child health somewhat exceeded expectations. Positive significant differences between the high and low treatment groups were observed on 12% of all measures, and four of the 32 step-down categories (13%) remained significant in the multiple hypothesis analysis. One potential explanation for the reduction in the number of significant results is that the high treatment participants spent less time with their mentors between thirty-six and forty-eight months compared to previous time points, with an average of 8.7 home visits over the twelve month period. Thus, the weakening effects could be attributed to reduced treatment.

An alternative potential explanation for the reduced number of significant differences is spillover effects, for example, if the *PFL* material and advice intended for the high treatment group was provided to the low treatment group. However, our analysis of contamination using the blue-dye question suggests that such spillover did not occur. In addition, the comparison of the low treatment group with the no-treatment comparison group indicated few differences between the two, with the low treatment group outperforming the no-treatment group on 7% of measures, while the no-treatment group outperformed the low treatment group on 8% of measures. As the low treatment group did not systematically outperform the no-treatment group, who could not have been contaminated, it suggests that the low treatment group did not receive the supports and services designed for the high treatment group. It also suggests that the low level supports provided to the low treatment group were not effective in improving child outcomes.

When the inverse probability weighting procedure (IPW) was applied to account for differential attrition, the number of individual significant findings changed considerably. The weighting particularly impacted the domain of child development by almost tripling the number of significant individual findings, from 19% to 56% of all measures. The weighting also impacted the domains of child health, parenting, and maternal health and wellbeing to a lesser extent. Conversely, the application of IPW caused a reduction in the number of significant findings in the areas of the home environment, social support, and household factors and SES, while there was no change to the findings in childcare. This analysis suggests that the type of mothers in the high or low treatment groups who did not participate in the forty-eight month interview had different characteristics which may have led to biases in the main results. For example, the higher proportion of treatment effects in the weighted analysis could suggest that high treatment mothers who did not participate in the forty-eight month interview may have had children with better outcomes, or alternatively, low treatment mothers who did not participate in the forty-eight month interview had

children with poorer outcomes. Thus, once the results were weighted to ensure that these excluded participants were represented in the analysis, the number of treatment effects rose. This explanation is a possibility as considerable effort was invested in re-engaging participants (from both the high and low treatment groups) in the forty-eight month interview who did not participate in other recent waves of data collection. In general, this process resulted in more disengaged participants from the high treatment group completing the forty-eight month interview.

An analysis of the predictors of attrition/wave non-response found that those who participated in the forty-eight month interview in both groups had characteristics which would traditionally be associated with more positive child and family outcomes. For example, those who remained in the high treatment group were more likely to be employed, to have health insurance, to be older, to have a higher IQ and higher consideration of future consequences, and to be satisfied with their neighbourhood compared to those who dropped out. Meanwhile, those who remained in the low treatment group exercised more, were Irish, were not first time parents, had a more open personality, were older, and had higher education and a greater knowledge of child development. Thus, parents in the high and low treatment groups who did not participate in the interview had more risk factors. The main difference regarding the predictors of attrition was maternal IQ. Mothers in the high treatment group with a lower IQ were less likely to participate in the interview, while there were no differences in the IQ scores of participants and non-participants in the low treatment group. The IPW analysis reconfigured the high treatment group to account for these low IQ parents. If the children of low IQ parents benefitted most from the programme in terms of improving their development, this may account for the emergence of more treatment effects for child development once the weighting was applied. Indeed, the interaction and subgroup analysis also suggested that children of low IQ mothers benefitted most from the programme.

AIM OF THE CHAPTER

The purpose of this concluding chapter is to discuss and interpret the main results comparing the high and low treatment groups in the context of the full report. As such, the chapter integrates all of the findings including the results from the IPW, implementation, dynamic, comparison group, and interaction and subgroup analyses. In addition, the results are contextualised within the relevant literature. The remainder of this chapter is structured by each of the eight outcome domains.

6.2 Child Development

Based on the literature, we hypothesised that there would be moderate programme effects on child development at forty-eight months. The main results in this domain were limited, particularly in light of findings reported at previous time points. Children in the high and low treatment groups differed significantly on 19% of the child development measures (6/32). This represents a downturn in the overall trajectory of child development effects from baseline to forty-eight months. The positive programme effect on child development, which had been increasing steadily from baseline to twenty-four months and held constant through thirty-six months, was significantly reduced at forty-eight months. It is important to note however, that when inverse probability weighting was applied, the number of significant findings rose substantially to 56%, which was the largest finding on any domain to date. Consistent with the previous published reports, this section examines the findings before IPW was applied.

At forty-eight months, children in the *PFL* high treatment group demonstrated positive programme effects in the areas of cognitive development, behaviour, and fine motor skills. There was some evidence of consistency over time, particularly in the areas of cognitive development, behaviour and age-appropriate skills. In addition, some of the positive effects on child behaviour observed at both twenty-four and thirty-six months were no longer evident at forty-eight months.

The impact on cognitive development, as found on the DP-3 score, was present for both the overall score and the binary indicator of above average development. This is an important result as much of the findings in the literature relate to individual aspects of cognitive functioning, such as executive functioning (Olds et

al., 2004) and mental processing (Drazen & Haust, 1993), rather than the general cognitive development identified here. Furthermore, these effects on cognitive development are consistent, having also been reported at twenty-four and thirty-six months. As cognitive development is strongly associated with improved future outcomes, this finding merits particular attention. A review of the literature implies that home visiting programmes typically report minimal impact on early cognitive development (for further information see *Preparing For Life* Early Childhood Intervention: Assessing the Early Impact of *Preparing For Life* at 6 Months).

Two significant findings relating to child behaviour also emerged at forty-eight months. On the CBCL measure, high treatment children were less likely to score above the cut-off for clinically significant levels of internalising and externalising problems, however, there were no effects for any of the three continuous CBCL scores. At twenty-four and thirty-six months, there were some significant findings on both the continuous and cut-off scores. It is possible that these effects were linked to the Triple P programme which was delivered between twenty-four and thirty-six months, and that these effects were not sustained beyond the delivery of the Triple P programme.

The findings in relation to the internalising and externalising problem cut-offs are similar to those reported by other home visiting programmes (Connell et al., 2008; Olds et al., 2004). Research suggests that children may develop behavioural problems through early maladaptive parent-child interaction processes (Buke et al., 2002; Hinshaw, 2002), therefore it is possible that any potential behavioural problems were offset by improved early parenting skills in the high treatment group, as were found at previous time points. For example, at thirty-six months, treatment effects were found for a reduction in permissive parenting and authoritarian parenting. Additionally, the Tip Sheets on social-emotional development provided parents with the tools for encouraging children to express emotions. This may help to offset children's communication problems which, if left unchecked, could lead to clinical levels of internalising or externalising behaviours by forty-eight months.

An important aspect of children's school readiness skills is their communication and language development. Contrary to some published findings in the field (Drazen & Haust, 1993; Landry et al., 2008), the *PFL* programme has yet to have an impact on these outcomes. This result is somewhat surprising as the programme specifically targets this outcome, and furthermore, there is a reported link between children's language ability and their cognitive development (Gopnik & Meltzoff, 1987 as cited in Carr, 2006). Thus, while the high treatment children consistently perform well in terms of their cognitive ability, they show no significant difference regarding language. The vocabulary Tip Sheets used by mentors targeted children up to age three and were not designed to be distributed between thirty-six and forty-eight months. Therefore, that there was no specific work carried out on language skills after age three may explain the lack of results in this area.

A positive programme effect was found in the area of fine motor skills. While this domain has been investigated by other programmes, significant findings at forty-eight months have not been reported (Drazen & Haust, 1993). The programme also had an impact on the children's fine motor skills when they were twelve months old. The Tip Sheets delivered between thirty-six and forty-eight months specifically focused on developing fine motor skills in preparation for school, such as using scissors and drawing shapes, therefore it is likely that this effect was driven by the focus on such skills. However, it is important to note that the Cronbach alpha for the fine motor domain was lower than the acceptable level of 0.7, thus, this result should be interpreted with caution.

A number of additional analyses were carried out to investigate who was likely to benefit most from receiving the *PFL* programme. For child development outcomes, the programme was most beneficial for the children of first time mothers and the children of mothers who have lower cognitive resources. The limited literature available on the effectiveness of early intervention programmes by gender is quite mixed (e.g. Barnett, 1995; Niles et al., 2008). In particular, most of the existing literature has identified differential gender effects for older children (Anderson, 2008; Barnett, 1995; Eckenrode et al., 2010). In general, few differential programme effects by gender were found, suggesting that the *PFL* programme

impacted on the development of boys and girls in similar ways. High treatment boys demonstrated some improvements regarding pro-social behaviour, which is consistent with Niles et al. (2008) who found that boys who participated in the Chicago Child-Parent Center Preschool Program had better peer-related social skills as adolescents.

As many home visiting programmes target first time mothers, few studies have examine differential treatment effects by parity. Our findings support this approach as we found that the children of first time mothers appear to benefit more from the programme in terms of their development than the children of mothers with more than one child. Previous studies have recommended that early intervention could offset some of the difficulties associated with first time motherhood, including support from other mothers and health professionals (Darvill et al., 2010) and training and information, particularly in lower socioeconomic groups (Sayil et al., 2006). Through *PFL*, first time mothers received a number of the supports recommended in the literature, such as adapting to motherhood, infant care, and anxiety reduction (Sayil et al., 2006). Thus, it is likely that the support of the mentor helped the first time mothers to adapt to their new role, and thus had a positive, if indirect, influence on child development. Our results are also somewhat consistent with the few home visiting studies which have targeted all parents, including the Healthy Families America programme (DuMont et al., 2008), which found more treatment effects for first time mothers in terms of parenting behaviour.

The finding that the majority of the effects in the child development domain were found for children of lower cognitive resource mothers also adds strength to the small body of literature on which our hypothesis was based (Olds, 2002; Olds et al., 2007; Olds, Kitzman et al., 2004; Olds, Robinson et al., 2004). This result is a considerable outcome for the *PFL* programme in terms of its catchment area and target families, as disadvantage is often associated with lower cognitive resources which can be a difficult intergenerational cycle to break. Thus, specifically improving the developmental outcomes of children with low IQ mothers is a positive development for the programme with important policy consequences.

A comparison of the Triple P participants with non-participants yielded limited findings in terms of its impact on child development. Children whose parents took part in Triple P outperformed the children of non-participants on only 6% of measures. This could be perceived as a limited outcome as Triple P is specifically designed to promote positive parenting and, through, that to impact child development. This result may be explained by the timing of the intervention. As most families received Triple P before thirty-six months, by forty-eight months it is possible that any potential impact may have faded over time. However, a subgroup analysis by Triple P participation at thirty-six months found more effects for non-participants than participants in terms of the children's development. Thus, overall, at least in terms of developmental outcomes, Triple P appears to have had limited impact.

It is notable that when IPW was applied to the overall findings, i.e. when a larger weight was given to the participants who were under-represented in the sample due to attrition, there was a dramatic increase in the effects on child development from 19% of individual tests in the unweighted results to 56% in the weighted results. There was a similar shift in the multiple hypothesis test findings, from 33% to 67% in the weighted results. It is important to note that the main results for cognitive development (DP-3) and behavioural problems (CBCL cut-offs) were present in both the weighted and unweighted analysis; and that the additional results found in the weighted analysis were largely driven by differences on the individual CBCL subdomains.

6.3 Child Health

Of the 35 child health measures assessed as part of the main analysis, six (20%) were statistically significant in the hypothesised direction, such that the children in the high treatment group outperformed the low treatment group. This is consistent with and, arguably, slightly exceeds our hypothesis based on the literature that there would be limited findings in the area of child health at forty-eight months. The programme's impact on child health peaked at twenty-four months with significant results on 47% of measures, while this reduced to 24% at thirty-six months. However, there was consistency over time

regarding the impacted health outcomes. For example, the majority of the significant findings identified at forty-eight months, such as asthma and diet were found at previous time points, while several new results also emerged in terms of the child's weight and sleep habits.

In relation to previous literature, the results were not fully consistent with other home visiting programmes at forty-eight months. For example, an evaluation of the Nurse Family Partnership programme reported fewer ingestions and injuries in treated children, and a reduction in visits to the emergency department at forty-eight months (Olds et al., 1994). While the numbers of children in the *PFL* programme being injured and visiting casualty was indeed lower among the high treatment group, this did not reach statistical significance. In addition, our most notable physical health finding related to asthma, which has not been found in other studies (Klinnert et al., 2007). It is also noteworthy that asthma reduction was not a specific goal of the *PFL* programme, although Ireland has the fourth-highest prevalence of asthma worldwide (HSE, 2015). As reported at twenty-four months, children in the high treatment group were significantly less likely to have asthma at forty-eight months, with the low treatment group reporting double the percentage of cases (25% in low, 12% in high).

The reduced incidence of asthma among high treatment children may be linked to a number of factors. High treatment children were less likely to live in homes where people smoked around them. This was a consistent effect found at multiple time points, including forty-eight months. Thus, as there is a connection between passive smoking and asthma in childhood (Hofhuis et al., 2003), this is one potential mechanism for the effect. Additionally, the physical presence of the mentor in the home may serve as a mechanism by influencing the parents' decision-making around medical treatments. As the high treatment group were more likely to have a medical card (see Table 2.9 in Chapter 2), it could also be argued that without the financial implications associated with GP visits, they have attended the GP more readily for pre-asthmatic complaints such as chest infections, coughs and viruses which may have had a preventive effect. Furthermore, through early medical intervention for these complaints, they may have evaded the need for antibiotics – which are a risk factor for the development of asthma (Murk, Risnes, & Bracken, 2011). However, high treatment children had no more reported GP visits or chest infections at thirty-six or forty-eight months than low treatment children, and at twenty-four months they were in fact less likely to attend the GP or report chest infections. Thus, it is possible that the reduced incidence of asthma was linked to a healthier lifestyle in the household, including a healthier diet and a smoke-free environment.

Children in the high treatment group were more likely to consume the recommended daily amount of vegetables than those in the low treatment group. These findings are likely to be linked to the Tip Sheets which provided direct nutrition advice, in particular teaching parents to include "hidden" vegetables in their daily cooking for fussy children. Furthermore, it is possible that the nutrition classes provided to the high treatment families earlier in the programme are now yielding a longer-term effect. High treatment children were also less likely to be overweight according to their BMI. The latter point represents a change from the thirty-six month findings, where there were no differences regarding children's weight. It is important to note that the measures of weight and height were recorded by the interviewer during the interview, and therefore not subject to parental misreporting. However, the number of children who agreed to measurements taking place was lower than the total sample who participated in the forty-eight month interview, this should be borne in mind when interpreting the result. In addition to dietary advice, the Tip Sheets give clear recommendations for fostering an active lifestyle, encouraging healthy sleep habits, family activities, and cautioning against more sedentary pursuits, for example, television watching. Thus, the reduced BMI among the high treatment group was most likely caused by a combination of these factors. This is an important finding as increasing levels of childhood obesity is now a key policy concern in Ireland.

Sleep habits were previously measured at six months, with no significant differences between the two groups. At forty-eight months, high treatment children reportedly slept for longer each day and had fewer reported sleep problems than their low treatment counterparts. This is an important finding as poor sleep habits can negatively impact child behaviour and development (Carter et al., 2014). Moreover, as poor sleep habits are associated with obesity (Van Cauter & Knutson, 2008), it is possible that the findings for

positive sleep habits and BMI are connected. The Tip Sheets provided by the mentors included advice on a bedtime routine and a sleep diary. While these Tip Sheets were generally distributed prior to thirty-six months, sleep habits were not measured between six and forty-eight months, therefore it is not possible to ascertain exactly when this positive effect on sleep emerged. However, it is likely that by forty-eight months parents have developed a clear routine which is impacting their children's sleep favourably, and potentially having a further, indirect effect on the children's health and development.

Toilet training was measured for the first time at forty-eight months. High treatment children were more likely to be fully toilet trained than low treatment children. However, counter to our hypothesis, of those who had been toilet trained, the parents of low treatment children reported that their children finished toilet training at a younger age than their peers. While it is developmentally and socially preferable for a child to finish toilet training at an earlier age, there is a specific Tip Sheet related to toilet training which is generally distributed prior to thirty-six months and recommends not rushing the child, citing that toilet training generally occurs between two and a half and four years of age. Additionally, it recommends taking a break from training if the child is unwell or becomes upset. Thus, the later finishing age among high treatment children may be linked to these suggestions, with high treatment parents deliberately not rushing the child. Furthermore, the question about finishing training specifically asks parents to compare their child to their peers. Thus, it is possible that the peer norm is simply different in the low and high treatment groups. Regardless, the fact that more high treatment children are fully toilet trained represents a positive finding for the programme, as daytime wetting and soiling are associated with a number of developmental and behavioural difficulties (Joinson et al., 2008).

When IPW was applied to the main child health findings, there was a small increase in the proportion of individual significant results from 20% to 23%, and a larger increase in the multiple hypothesis test findings from 0% to 40%. The main gains were in relation to asthma and toilet training as discussed above.

6.4 Parenting

Mothers in the high and low treatment groups differed significantly on 6% or two of the 36 parenting outcomes measured at forty-eight months, and none of the multiple hypothesis tests reached significance. While moderate effects on parenting were hypothesised, these limited results represent a sharp decrease on almost all previous time points, in particular the thirty-six month assessment when high treatment mothers outperformed low treatment mothers on 26% of individual tests, with significant differences found on 43% of the step-down tests.

At forty-eight months, the high and low treatment groups differed significantly on two aspects of parenting. Firstly, as found at thirty-six months, mothers in the high treatment group reported fewer permissive parenting behaviours than those in the low treatment group. This is an encouraging finding as permissive parenting behaviours have been associated with negative developmental outcomes (Aunola & Nurmi, 2005; Petito & Cummins, 2000). However, a treatment effect on authoritarian parenting styles, which had been present at thirty-six months, was no longer significant at forty-eight months, albeit it was in the hypothesised direction. Furthermore, the authoritarian subdomains, which were significantly lower in the high treatment group at thirty-six months, now suggest a narrower gap between the groups, with no significant differences. This reduced impact on parenting behaviours is a key concern for the *PFL* programme whose logic model is based on improved parenting behaviour as the key mechanism of change with respect to improving the children's school readiness skills. As discussed above, this effect may be attributed to the reduced contact time between mentors and participants in the final year of the programme or the emphasis on child skills rather than parenting.

There were also fewer effects on the children's TV habits compared to thirty-six months. At forty-eight months only one significant difference was observed such that mothers in the high treatment group reported that their children spent less time watching TV alone than those in the low treatment group. Television viewing has been linked to a number of negative health and developmental outcomes for children (e.g. Christakis et al., 2004; DuRant et al., 1994; Hancox et al., 2005), and in keeping with the guidelines set

by the American Academy of Pediatrics, *PFL* recommend that children over two are not exposed to more than 2 hours of television or other media per day. Watching television with a parent enables TV to become a more interactive experience, as the parent can take opportunities to discuss the programme with the child and change the programme if it is deemed unsuitable, and accordingly, the *PFL* Tip Sheets suggest scheduling 'family' TV time. While this is a positive finding, a number of other effects in this domain were no longer significant, including placing a limit on the amount of time a child can spend watching TV, the length of time the TV is on in the home, and the amount of time spent by the child watching TV or DVDs daily. As these also correlate with direct suggestions in the Tip Sheets, it could have been anticipated that they would have remained significant at forty-eight months.

These results are somewhat in contrast with the literature as several home visiting programmes have reported positive effects on parenting at forty-eight months (Landry et al, 2008; Landry et al, 2012; Madden et al., 1984; Olds et al, 1994; Olds et al, 2004;). One potential explanation is that many of the results in the literature were based on measures of parenting from observations of parent-child interactions, while the results reported here are based on parent report. Nevertheless, the small number of parenting findings at this time point is somewhat counterintuitive. In particular, the dynamic analysis revealed a significant change over time between the high and low treatment groups on the Tasks subscale of the Parenting Daily Hassles Scale. While at thirty-six months the low treatment group exhibited a higher score than the high treatment group, indicating that the low treatment parents had more difficulty with tasks at this time, this trend had reversed by forty-eight months. Thus, by forty-eight months, high treatment parents were struggling more with parenting tasks. It is also surprising that the two groups did not differ regarding the traits considered important for success at school entry. In both groups, social competence was cited most frequently as being important for school success, followed by children's language and cognitive development, while emotional maturity and communication and general knowledge were cited less frequently by both groups of parents.

A number of the parenting measures used at thirty-six months were also used at forty-eight months, including the Parenting Daily Hassles scale, the Parenting Styles and Dimensions Questionnaire, and TV habits, yet the number of significant findings has diminished sharply. It is possible that the high treatment findings at thirty-six months had been impacted by recent participation in the Triple P programme, and that this influence had faded by forty-eight months. Taking a theoretical perspective, the theory of change underpinning *PFL* hypothesised that through altering parental behaviours child outcomes would be positively affected. However, these findings suggest that at forty-eight months, while there are several differences between the two groups regarding child outcomes, the parenting behaviours of the high and low treatment groups do not differ substantially. This implies that while child outcomes are changing, the proposed mechanism for these changes – improvements in parenting attitudes and behaviours - were not impacted by the programme.

As the Triple P programme was specifically designed to improve outcomes through the development of better parenting practices, its impact on parenting outcomes was specifically analysed. The comparison of Triple P participants to non-participants within the high treatment group indicated that the programme had some benefits on parenting outcomes. However, when further analyses were conducted by comparing Triple P participants to the low treatment group, no significant differences were present.

When IPW was applied, there was just one extra significant finding in the weighted parenting results that was not statistically different in the unweighted results. The IPW results indicated that parents in the high treatment group were more likely to watch television with their child. Thus, unlike the child development results, adjusting for differential attrition did not change the parenting results.

6.5 Home Environment

Of the four home environment measures examined at forty-eight months, two (50%) were significant, such that the high treatment group reported more positive outcomes than the low treatment group. As a percentage, this is in line with the literature, which suggests a moderate effect on the home environment at this time point. High treatment children were less likely to be exposed to cigarette smoke at home than low treatment children, and their families were less likely to be working with a social worker. These replicate findings identified at previous waves - the result in relation to smoking was found at thirty-six months, while the social worker finding was also present at twenty-four months. The dangers of passive smoking to children have been widely highlighted, not least due to the link between passive smoking and asthma. There was a *PFL* Tip Sheet which explained the risks of passive smoking and provided clear step-by-step advice on how to limit children's exposure to smoke. It is possible that this finding is linked to the asthma result in the child health section, whereby high treatment children were significantly less likely to suffer from asthma than low treatment children. This result represents a strong outcome for *PFL* as it has not hitherto been found in the international home visiting literature.

High treatment families were also less likely to be working with a social worker. Social workers are typically assigned to a family who are deemed to be socially at-risk, with the majority providing frontline childcare services such as identifying at-risk children (http://www.citizensinformation.ie/en/health/care_in_your_community/social_work_services.html. It is possible that, through assisting families and visiting regularly, the mentors could identify precursors of child risk and provide families with solutions before difficulties reached a level where social services were required. While the mentors do not formally discuss financial issues, alcohol or substance abuse, which could arguably lead to higher levels of familial risk, they do work on aspects of parenting, child behaviour, health and safety, therefore they may have sought to improve the family environment and reduce overall difficulties through providing strategies for dealing with typical family stressors. The "mentor as listener" model, which was identified in qualitative research with parents at twenty-four months, may help to explain this result. Aside from the Tip Sheets which guided the home visits, there was room for mothers to talk to the mentors about other family issues which may have offset problems before they reached a chronic level. It is notable then that when the IPW procedure was applied, this difference was no longer significant.

There were no significant differences between the two groups on a number of salient elements of the home environment which are critical for child development and health. For example, there were no differences on the Home Learning Environment (HLE) index – a measure of how often children engage with their parents on specific activities, including reading, library use, physical activities such as dance and sports, letters and numbers, singing, painting and drawing. While this was the first time that this particular measure was used in the *PFL* evaluation, similar instruments were used at previous time points, the most comparable being the interaction with baby/child scale which was used at six and eighteen months and yielded significantly different outcomes for the high treatment group at both junctures. That a similar effect was not found at forty-eight months is particularly counterintuitive given the programme's focus on promoting such activities. For example, the mentors encourage parents to interact with their children, and the benefits of reading and music are specifically outlined in Tip Sheets, although these are generally distributed prior to thirty-six months. Therefore, it is possible that by forty-eight months, the strength of these recommendations has dissipated and families are simply focusing on other areas of their children's development. In light of these concerns, each individual item on the HLE was subjected to further investigation for treatment effects, yet none of the individual items reached significance. Thus it is evident, and perhaps surprising, that the high treatment group are not engaging in these activities any more than the low treatment group.

There were also no significant differences between the two groups regarding the safety of the child's physical environment. By forty-eight months children are gaining independence, largely through mastering motor skills (Sheridan, 2004). Safety in the home through environmental modification, education and regulation is recommended up to forty-eight months by the World Health Organisation (Hyder et al., 2009). However, the practicality of making a home childproof for this age group is difficult: with increasing motor

skills, physical home safety barriers such as stairgates, fireguards etc., can potentially be manipulated by children, rendering them less effective. The Tip Sheets on child safety and supervision were generally distributed prior to thirty-six months, and by forty-eight months the focus was on issues of physical health safety, such as limiting sun and smoke exposure.

6.6 Maternal Health & Wellbeing

Of the 21 maternal health and wellbeing measures considered at forty-eight months, only three (5%) were statistically significant. This finding is counter to the hypothesis, based on the home visiting literature, that moderate effects would be identified. With the exception of thirty-six months, when treatment effects were found on 24% of maternal health and wellbeing measures, this particular domain has yielded consistently low findings, suggesting that the programme has had limited effects on the mother's wellbeing.

The first significant finding at forty-eight months related to the mother's general health. High treatment mothers were more likely than low treatment mothers to report that they were in good health compared to other women. The remaining significant findings were both related to alcohol use. High treatment mothers were less likely to report that they consumed more than 14 units of alcohol per week, and were less likely to report binge drinking, than low treatment mothers. This was a consistent finding over time, with reduced alcohol consumption among the high treatment group also noted at thirty-six, eighteen, and twelve months. These findings are important as poor maternal physical health is associated with a number of negative child outcomes (e.g. Lester et al., 2006; Osborn, 2007; Vannatta, Grollman, Noll, & Gerhardt, 2008). While the Tip Sheets do not address alcohol use directly at this timepoint, it is possible that by consuming less alcohol, the high treatment group were more conscious of the negative impact of alcohol use on family time.

The lack of findings in other areas – namely maternal mental health and maternal self-efficacy – represents a challenge to the programme. Maternal depression has been linked to behavioural problems and to lower vocabulary scores in children (Brennan, Hammen, Andersen, Bor, Najman, & Williams, 2000), in addition to elevated rates of childhood depression (Downey & Coyne, 1990), and low income families are at particular risk of mental health difficulties (Shonkoff & Phillips, 2000). At thirty-six months, high treatment mothers were at significantly lower risk of mental health difficulties as measured by the WHO-5 and Edinburgh Postnatal Depression scales. It is concerning that these effects had dissipated by forty-eight months. Furthermore, there were no reported significant differences for maternal self-efficacy or self-esteem. It is possible that these particular constructs were more challenging to change over time as they are often rooted in long-held personal beliefs and internal working models which are arguably more resistant to change. On the other hand, practical changes were more easily implemented, e.g. choosing to drink less alcohol or choosing not to smoke in front of the child. That the findings in relation to depression and wellbeing changed at the thirty-six month time point is difficult to account for in light of these most recent findings. Between the ages of zero and two, a number of Tip Sheets on self-care were provided, however the focus in later Tip Sheets was mainly on preparing the child for school. It is possible that the effects of the self-care advice was experienced at thirty-six months, but had dissipated by forty-eight months as the focus moved away from maternal wellbeing.

There was a small increase in findings when weighting was applied. The IPW results indicated that high treatment group mothers were more likely to use a valid form of birth control compared to the low treatment group. The weighting also yielded one additional significant finding for the multiple hypothesis test for Current Substance Use.

6.7 Maternal Social Support

Of the 14 measures included in the maternal social support domain, half were in the hypothesised direction and two (14%) were statistically significant. This is in line with the hypothesis which, in keeping with the literature, anticipated limited findings in this area. Both of the significant findings related to voting: high

treatment mothers were more likely to report having voted in the last local, European, and general elections than low treatment mothers. There were no significant findings regarding support from key people in the mother's life, including family, friends, neighbours, partners and/or the child's father, and indeed in many cases, the low treatment group reported higher levels of support than the high treatment group, although these differences failed to reach significance.

Social support has been highlighted as important for maternal physical and mental health (Beck, 2001; Berkman et al., 2000; Kawachi & Berman, 2001; Webster et al., 2011), and greater levels of maternal social support have been associated with positive outcomes for children (e.g. Melson et al., 1993, Slykerman et al., 2005). It is notable, then, that the programme does not appear to have impacted maternal social support, although the international literature indicates that this is a challenging area to change. It is possible that the mentor-parent relationship is directly affecting social support. The mother may be relying on the mentor as a proxy for social support from other individuals. It is also plausible that as the children grow older and participate in more formal activities, such as preschool, the mothers in both the high and low treatment groups derive the same level of support from their social network.

The results regarding social support have been very mixed over time. Early in the programme's evaluation, social support was relatively high among the high treatment group, with significant findings on 38% and 43% of measures at six and twelve months respectively. However, this reduced quite sharply at eighteen months and has remained low since. It is possible that the programme encouraged the high treatment group to avail of support from friends and family when their babies were young, yet as the children aged and the parents became more confident in their parenting skills, the level of support required was reduced. One potential concern was the lack of perceived support from the partner and/or the child's father. Up to twenty-four months, the Tip Sheets discussed the parents' relationships and encouraged self-care and time spent together. However, between thirty-six and forty-eight months, this area was not specifically addressed. The reported mean scores suggest that high treatment mothers were less satisfied with their partner's overall support and less likely to receive regular child maintenance payments than low treatment mothers, although these findings were not significant. It is possible that, through their involvement in the programme, high treatment mothers were more aware of the importance of social support, and were thus more likely to notice or comment on problems, difficulties or changes in this area. As more high treatment mothers reported living with a grandparent (see section 6.9: Household Factors and SES), it seems counterintuitive that they would report a lower level of support from their own parents unless they were more attuned to observing the nature and quality of support.

When the IPW procedure was applied, the two significant findings in relation to voting were rendered insignificant. Additionally, a further result was identified which was significant in the non-hypothesised direction. Specifically, the IPW results indicated that mothers in the high treatment group received less support from their partner than those in the low treatment group.

6.8 Childcare

There were no significant differences between the high and low treatment groups on the eight measures analysed in relation to childcare, which is consistent with our prior hypothesis. Moreover, the results for the IPW procedure did not differ from the unweighted results. This replicates the findings at previous time points. Much of the literature does not report on the impact of home visiting interventions on childcare, however our results are not consistent with one study of the Nurse Family Partnership programme which reported a negative effect on centre-based childcare attendance at forty-eight months (Olds et al, 2004).

The introduction of the free preschool year means that most children in Ireland now experience at least one year of preschool or centre-based care before starting formal schooling (Burke et al., 2012). High quality preschool childcare is associated with a number of positive child outcomes (Sylva et al., 2011), particularly for children from disadvantaged backgrounds (Sylva et al., 2004), with some negative behavioural factors associated with very early formal childcare (Loeb et al., 2007; National Institute of Child Health and Human Development, 2002). In the present study, by forty-eight months, 98% of the children in childcare

were in formal childcare. A higher proportion of the high treatment children were in Siolta-accredited centres, i.e. childcare centres that are adhering to the quality framework set out by the Irish government. While not significantly different, this suggests that they were receiving higher quality childcare than the low treatment children, and in keeping with the literature, it could be anticipated that this is a strong advantage given their relatively low socioeconomic status. In addition, a higher proportion of the high treatment mothers said that they were satisfied with the level of childcare their child was receiving. Through their involvement with *PFL* and regular home visits, it is possible that the high treatment mothers have a clear understanding of the markers of quality childcare, such as activities with the children and developmental resources, and are thus able to be discerning about their choice of facility. It is interesting, given the different levels of satisfaction, that the low treatment group are availing of a more expensive form of childcare. Indeed it is possible that the higher price is causing them to more strongly question the quality of the childcare provided.

6.9 Household Factors & SES

Of the 41 household and SES measures assessed, only two (5%) were statistically significant in the hypothesised direction, while four (10%) were significant in the non-hypothesised direction. This is counter to the hypothesis, as moderate findings were anticipated in this domain. Household factors and SES are considered central to child developmental outcomes, with multiple studies highlighting the negative effects of economic and social disadvantage on children's trajectories (e.g. Siddiqui et al., 2007).

High treatment mothers reported significantly fewer family mental health issues and risks in the "other" category than low treatment mothers. In the non-hypothesised direction, high treatment mothers were more likely to report living with a grandparent, having a medical card, and being in receipt of household unemployment benefit than low treatment mothers. High treatment mothers were also more likely to report having difficulties with suicidal thoughts in their family. Overall, these findings suggest that the high treatment group were at a higher level of financial disadvantage than the low treatment group, which potentially places the high treatment children at higher risk of developmental difficulties. Furthermore, differences across the high and low treatment families have changed over time. At thirty-six months, the low treatment group experienced more economic difficulties such as unemployment and financial distress, although this pattern was reversed by forty-eight months. However, the lack of significant differences between the two groups on all other measures in this domain must be taken into account. The finding in relation to suicidal thoughts among family members (although we cannot determine among which family members, including the mother herself), while affecting a very small number of participants, is concerning in light of its gravity. It is possible that there is a link between the increased suicidal thoughts and the few indicators of financial disadvantage reported among the high treatment group.

When the IPW method was applied, a number of changes emerged. The significant findings in relation to mental health issues or other risks were not replicated in the IPW results, while the negative difference in relation to residing with a grandparent also lost its significance. The IPW results also showed four significant differences which were not statistically different in the unweighted results, only one of which was in the hypothesised direction.

6.10 Future Reports

The final report will provide an overview of the *PFL* findings from baseline to forty-eight months, and will examine the children's school readiness skills as they enter primary school. This report will provide the final overview of the *PFL*'s success in improving the development, health, and wellbeing of children and their families as children commence formal schooling.

References

Please see the following website for the reference list:

http://geary.ucd.ie/preparingforlife/

preparing for life

planning together for our children

The ATLANTIC Philanthropies

An Roinn Leanaí agus Gnóthaí Óige

Department of Children and Youth Affairs

Northside Partnership