preparing for life

Early Childhood Intervention

Assessing the Impact of *Preparing For Life* at 36 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 Thirty-six Months

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

By
PFL EVALUATION TEAM AT THE UCD GEARY INSTITUTE FOR PUBLIC POLICY
December, 2014

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 seven years (Ailbhe Booth, Carly Cheevers, Maria Cosgrove, 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. Results from Chapter 7 of the present report were submitted by Lauren Smith to the UCD School of Psychology as part of the requirements for the M. Psych. Sc. 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. In addition, we gratefully acknowledge funding from the Irish Research Council Collaborative Project Scheme which supported the Day in the Life of a *Preparing For Life* Parent sub-study. 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 seven 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 and books, 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 Program. 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 thirty-six months of age.

Recruitment and Baseline Characteristics

At baseline, 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, and 99 comparison group participants. Tests of baseline differences between the high and low *PFL* treatment groups found that the two 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 (SES).

Findings from the Six Month Report

In total, 257 six month interviews (n*Low* = 90; n*High* = 83; n*LFP* = 84) were completed. As found in other studies of home visiting programmes, there were limited significant differences between the high and low treatment groups (14%) at six months. 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, quality of the home environment, and social support, which corresponds directly to information provided by the *PFL* mentors during qualitative interviews. However, the programme had no significant impact on pregnancy behaviour, infant birth weight, breastfeeding, and child development. While attrition from the programme was low and participant satisfaction was high, the level of engagement was less than anticipated, with parents in the high treatment group receiving 14 home visits between programme entry and six months. Mothers with relatively higher cognitive resources received more home visits and may have benefited more from the programme at six months than those with lower cognitive resources.

Findings from the Twelve Month Report

Altogether, 247 twelve month interviews (n*Low* = 83; n*High* = 82; n*LFP* = 82) were completed. Limited significant differences between the high and low treatment groups (8%) were found. Note that the measures used at the six and twelve month evaluations differed, therefore the two time points may not be directly compared. The high treatment group reported somewhat better outcomes than the low treatment group. There were some significant findings in all domains, apart from parenting and the home environment, and the largest number of significant results was found in the domains of child health and social support. Findings from the dynamic analyses were limited. The level of attrition between six and twelve months was low; however, engagement was lower than prescribed, with the high treatment groups receiving seven programme visits on average between six and twelve months. Overall, participant satisfaction was high,

and although the risk of contamination was high, there was little evidence of contamination between the high and low treatment groups at twelve months.

Findings from the Eighteen Month Report

In total, 225 eighteen month interviews (n*Low* = 74; n*High* = 80; n*LFP* = 71) were completed. Overall, the findings from the eighteen month evaluation were consistent with similarly timed evaluations of other home visiting programmes. Of the outcomes analysed, 14% showed significant differences between the high and low treatment groups. Based on the literature, we hypothesised that treatment effects would be found in the domains of the home environment, parenting, child health, and child development. The results suggest support for our hypotheses, as significant effects were found in these domains. We also found limited effects in the domains of social support and maternal health. Attrition and disengagement were low between twelve and eighteen months - no attrition was experienced in the high treatment group or the comparison group and only 2% attrition was experienced in the low treatment group. As experienced in other periods, engagement was lower than prescribed, with the high treatment groups receiving six programme visits on average between twelve and eighteen months.

Findings from the Twenty-four Month Report

In all, 239 twenty-four month interviews (n*Low* = 84; n*High* = 82; n*LFP* = 73) were completed. There were positive, significant differences between the high and low treatment groups on 21% of measures. These findings represented the largest proportion of significant individual tests reported to date. The significant findings in the domains of child development and child health supported, and exceeded, our hypotheses, such that the number of significant positive outcomes in these domains doubled compared to the eighteen month evaluation. This is a key finding as the programme impacted multiple areas of the child's physical and social development. There were also positive findings in the areas of parenting and home environment. The dynamic analysis found that only the measure of satisfaction with the programme showed a significant interaction between treatment and time in the hypothesised direction. There was no attrition in the high or low treatment groups between eighteen and twenty-four months, and on average participants received six home visits. Although the risk of contamination was high, a direct measure suggests that the level of contamination was low and did not bias the twenty-four month results.

Thirty-six Month Report

The aim of this report is to test whether the *PFL* programme had an impact on parent and child outcomes at and before thirty-six months, and to provide a detailed review of implementation practices regarding attrition, dosage, participant engagement, and programme effectiveness. It is important to note that in addition to the home visiting programme, high treatment mothers and their partners were also offered a second treatment, the Triple P Positive Parenting Program, between twenty-four and thirty-six months. Triple P is a gold standard parenting programme which seeks to promote healthy parenting practices and positive attachment relationships by focusing on the home environment, parent self-awareness, and parenting techniques. Thus, the results of the thirty-six month evaluation should be interpreted in light of this additional treatment. This report also outlines results from two sub-studies, designed and carried out to complement the main evaluation. The first provides a targeted investigation of the impact of the *PFL* programme on maternal wellbeing, while the second investigates the experience of fatherhood within the *PFL* community and father figure involvement in the *PFL* programme.

Table ES.1 - Summary of Main Findings at Six, Twelve, Eighteen, Twenty-four, & Thirty-six Months

PFL Low – *PFL* High	Proportion of Measures Significantly Different									
	Six Months		Twelve Months		Eighteen Months		Twenty-four Months		Thirty-six 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
Child Development	0% (13)	0% (2)	7% (28)	20% (5)	16% (25)	0% (6)	34% (41)	22% (9)	33% (39)	38% (8)
Child Health	10% (30)	0% (3)	17% (23)	0% (4)	24% (17)	67% (3)	47% (17)	50% (2)	24% (21)	33% (3)
Parenting	23% (22)	20% (5)	0% (16)	0% (2)	20% (10)	50% (2)	18% (17)	0% (3)	26% (34)	43% (7)
Home Environment	36% (22)	50% (2)	0% (6)	0% (1)	33% (21)	67% (3)	50% (2)	~	40% (15)	50% (2)
Maternal Health & Wellbeing	5% (20)	25% (4)	4% (28)	25% (4)	5% (19)	0% (3)	6% (16)	0% (3)	24% (17)	33% (3)
Social Support	38% (13)	0% (2)	43% (7)	0% (2)	8% (12)	0% (3)	10% (19)	0% (4)	5% (19)	0% (4)
Childcare	7% (14)	0% (2)	~	~	0% (16)	0% (2)	0% (7)	0% (1)	0% (17)	0% (2)
Household Factors & SES	0% (26)	0% (5)	3% (32)	0% (5)	8% (23)	0% (5)	13% (47)	29% (7)	14% (42)	0% (6)
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)

Impact of *PFL* at Thirty-six Months: Main Results

In total, 217 thirty-six month interviews (n*High* = 75; n*Low* = 76; n*LFP* = 66) 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 204 outcome measures.

Table ES.1 summarises the *PFL* results at six, twelve, eighteen, twenty-four, and thirty-six months. Based on the literature, we anticipated moderate positive effects in the areas of child development, child health, and parenting. However, given the implementation of Triple P we anticipated a greater number of favourable effects in the domains of parenting and child development, as compared with previous waves of the *PFL* evaluation and other studies evaluating home visiting programmes in isolation. Consistent with the literature, we also hypothesised few positive effects in the areas of the home environment, maternal health, and household factors and SES. Regarding maternal social support and childcare, which were only investigated by one other programme, we hypothesised that effects in both domains would be limited. Positive significant differences between the high and low treatment groups were observed on 22% of measures, and nine of the 35 step-down categories (26%) remained significant in the multiple hypotheses analysis, including aspects of child development, child health, parenting, the home environment, and maternal health and wellbeing. The findings at thirty-six months are consistent with, and slightly exceed, those at twenty-four months and represent the largest proportion of significant individual tests and step-down tests found to date.

Consistent with the literature and our hypotheses, we found moderate positive effects in the domain of child development. The results are consistent with the twenty-four month findings, and suggest that the programme is having a positive and sustained impact on multiple areas of the child's development such as cognition and problem behaviours. There were also significant positive effects in the area of child health, which supported our hypothesis, although fewer effects were found at thirty-six months than twenty-four months. The parenting and home environment domains also showed significant positive findings, with the results in the home environment domain exceeding our hypotheses. The results in the area of maternal health and wellbeing also supported our hypotheses, such that the number of significant positive outcomes in this domain tripled compared to the twenty-four months results. This is noteworthy as only one other study (Shaw et al., 2009) investigated and reported positive effects on maternal health and wellbeing at thirty-six months.

INDIVIDUAL FACTORS FOUND TO BE SIGNIFICANT AT THIRTY-SIX MONTHS INCLUDE THE FOLLOWING:

Children in the high treatment group exhibited stronger problem solving skills and cognitive development than the low treatment group. They also exhibited fewer problem behaviours such as somatic complaints, aggressive behaviour, sleep problems, other problems (e.g. toileting problems), and clinically significant levels of externalising problems. High treatment children also tended to have developmental outcomes which were more similar to each other as a group than low treatment children, who exhibited more extreme scores.

Children in the high treatment group were reported to have had fewer accidents and were less likely to have stayed in hospital overnight in the last year, or to have been diagnosed with a chronic illness. They were also more likely to consume proteins and to meet dietary guidelines compared to children in the low treatment group. Children in the high treatment group were also less likely to be exposed to cigarette smoke.

Children in the high treatment group spent less time watching TV & DVDs and were less likely to spend time watching TV alone than those in the low treatment group. Of the mothers who reported imposing a limit on TV viewing time, those in the high treatment group reported allowing their children to watch less TV, and having the TV on for a shorter amount of time during the day than those in the low treatment group.

Mothers in the high treatment group were less likely to engage in behaviours associated with either authoritarian or permissive parenting than mothers in the low treatment group. They were also less likely to engage in punitive or hostile parenting. However, they had higher scores on an indicator of parental coldness and an indicator of indifference/neglect towards their child.

Mothers in the high treatment group were less likely to experience symptoms of depression, less likely to report drinking alcohol in the previous year, and more likely to have reduced their smoking between twenty-four and thirty-six months. They were also more likely to report greater levels of wellbeing when compared to the low treatment group.

In terms of household factors, mothers in the high treatment group were more likely to classify themselves as homemakers, be optimistic about their financial situation, and less likely to be unemployed. Families in the high treatment group were less likely to be living in social housing and more likely to report higher satisfaction with their neighbourhood. Their home environment was more likely to be safe and organised with a regular routine in place. Their homes were more likely to promote learning and development and be accepting of less than optimal behaviour.

PFL Implementation Analysis at Thirty-Six Months

ATTRITION

The level of official attrition from *PFL* between baseline and thirty-six months was 16% across the whole sample. Official attrition was slightly higher among the high treatment group (19%) than among the low treatment group (16%), and it was lowest among the comparison group (12%). Importantly, the level of official attrition was minimal between the twenty-four and thirty-six month interview rounds, with no attrition experienced in the high or low treatment groups and only 2% attrition in the comparison group. In addition to those who dropped out, 19% of the sample did not complete a thirty-six 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 19% and 16% respectively, while the corresponding rate for the comparison group was 12%. Total non-completion (attrition and disengaged) at thirty-six months was very similar among the high (35%) and the low treatment groups (36%), and slightly lower among the comparison group (33%). However, non-completion is 25% for the low treatment group and 28% for the high treatment group if we restrict our attention to the sample that completed a baseline survey.

In order to test for non-random attrition, we compared the baseline characteristics of those who participated in the thirty-six months survey to those who did not. Overall, there was some evidence of systematic differences between these groups, such that more disadvantaged participants were difficult to contact or had dropped out of the programme by thirty-six months. For example, mothers in the high treatment group who did not participate in the survey were less likely to have been employed at baseline and had lower levels of cognitive resources. However, as shown in previous reports, the majority of individual characteristics were not associated with attrition. To account for any potential bias due to differential attrition, the main outcome analyses were re-estimated using an inverse probability weighting (IPW) technique. Slightly fewer of the individual tests showed significant differences between the high and low treatment groups when the weighting was applied (20%), as compared with the unweighted results (22%). However, a slightly higher number of multiple hypothesis tests were significant in the weighted results (29%) as compared to the unweighted results (26%).

ENGAGEMENT

Families in the high treatment group received an average of 46 home visits from the *PFL* mentors between programme intake and thirty-six months, with each visit lasting slightly over 1 hour on average. The number and duration of visits were roughly similar across each time period. The proportion of prescribed home visits delivered was gradually declining until twenty-four months, however, this pattern was reversed between twenty-four and thirty-six months when the proportion of prescribed visits delivered rose above

50%. On average, participants met their mentor just over once a month between twenty-four and thirty-six months.

Few individual participant characteristics were associated with the frequency or duration of home visits. Consistent with previous reports, mothers with higher cognitive resources participated in more home visits and spent more total time in visits. This suggests that engagement may be related to the mother's ability to understand the programme materials and recognise the potential need for the programme in their lives. In addition, mothers who smoked during pregnancy had fewer visits and spent less total time in visits. Factors such as age, marital status, employment status, and socio-emotional functioning were not associated with engagement in *PFL*.

SATISFACTION

Overall, participant satisfaction with the programme at thirty-six months was high. As expected, the high treatment group reported greater satisfaction with the programme compared to the low treatment group. However, the low treatment group still reported relatively high levels of satisfaction despite the minimal supports received. Both groups were most satisfied with the programme as a whole, their child's progress and the amount of help they receive from the programme. However, both groups reported being least satisfied with how the programme has improved their relationships with their partner. This is consistent with previous findings.

MISREPORTING

Differential misreporting between the high and low treatment group was measured using a bogus question which asked participants their knowledge of a fake child development term. The high treatment group were more likely than the low treatment group to claim to have heard the term. This suggests that members of the high treatment group may be more likely to provide answers which they feel portray a better image of themselves as parents. However, this result is in contrast to the findings of no differences on social desirability across the groups reported at previous time points.

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, including neighbours with similar aged children neighbours taking part in *PFL*. In addition, the majority of participants in both groups claimed to share their *PFL* materials with others. The direct measure of contamination suggested that all three groups (high, low, and comparison) had similar levels of parenting knowledge. This suggests that the chosen contamination question was ill-suited to determine whether contamination took place.

Additional Thirty-Six Month Analyses Results

Additional analyses were conducted to explore different aspects of the data not captured in the main analysis. These included a comparison of the thirty-six month outcomes of the low treatment group to the comparison group, and the thirty-six month dynamic analysis which examined changes in child and parent outcomes over time.

Overall, the mixed results of the low treatment group and the comparison group analysis support the study design as it suggests that the low treatment group did not perform systematically better than the comparison group across most domains. Of the 193 items analysed, there were statistically significant findings in the hypothesised direction for 19 measures (10%) and there were 28 statistically significant differences in the non-hypothesised direction (15%). These results indicated that the low treatment group was not outperforming the comparison group and suggest that either contamination between the high and low treatment group was minimal, or that the low treatment supports had little effect on the participants.

A number of standardised instruments were collected at multiple time points which allowed us to compare the outcomes of the high and low treatment groups over time. The goal of dynamic analyses is to determine whether the programme's impact increases or diminishes from wave to wave. Overall, nine (14%) of the 64 measures upon which dynamic analyses were conducted, yielded significant interactions between interview wave and treatment. Seven (11%) of these findings were in the hypothesised direction implying the strengthening of positive treatment effects over time. However, two (3%) were in the non-hypothesised direction, suggested that the programme's impact diminished over time in these areas. These hypothesised findings were concentrated in four areas: child development, maternal health and wellbeing, the home environment, and participant satisfaction; while both the non-hypothesised findings were found in the home environment.

Triple P Positive Parenting Program Results

Triple P was offered to all high treatment group participants when their *PFL* child was two years old. Triple P aims to promote healthy parenting practices and positive parent-child attachment and can be delivered at different levels. Participants were offered Selected Triple P (Level 2), Triple P Discussion Groups (Level 3), Primary Care (Level 3), and Group Triple P (Level 4). Of the families who took part in the twenty-four month interview, 59% (n=48) participated in some form of Triple P. Parents who participated differed in a number of ways at baseline to those who did not participate, such that they were more likely to be living with their own parents and were more likely to be in paid employment at baseline. They also had higher cognitive resources and were less likely to have smoked during pregnancy, yet were more likely to have a mental health condition.

Group Triple P, which consists of five 2-hour group discussion sessions facilitated by the *PFL* mentors and three individual phone calls, was the most popular type of Triple P offered. Of the 43 mothers and 16 partners who attended Group Triple P, 84% and 50% respectively completed the prescribed dosage. The Triple P Discussion Groups, which required less time commitment and consisted of four 90-minute sessions on specific topics, were attended by 21 mothers and two partners. Of the four topics covered, bedtime routine was the most popular, followed by the fighting and aggression. Of the six *PFL* families who took part in Primary Care Triple P, which consisted of four weekly 60-minute sessions delivered in the home by *PFL* mentors, four families completed the allocated series. Finally, Selected Triple P, which consisted of three 90-minute seminars introducing one of the Triple P strategies such as The Power of Positive Parenting, was attended by two families.

A sub-group analysis of Triple P participants and non-participants revealed that both groups outperformed the low treatment group on 19% of the individual tests. Triple P participants outperformed the low treatment group primarily in the areas of parenting and the home environment, while non-participants outperformed the low treatment group on a higher proportion of measures on the child development and child health domains.

A Day in the Life of a *Preparing For Life* Parent Study

During 2013 an additional sub-study, A Day in the Life of a *Preparing For Life* Parent, was conducted. The study sought to complement the main evaluation by providing a targeted investigation of the impact of the *PFL* programme on maternal wellbeing.

A multi-method approach was employed utilising a unique combination of measurement techniques including a day reconstruction method which records participants' activities and emotional states over the course of the day (Day Reconstruction Method; Kahneman et al., 2004), global questions about mood and life satisfaction, a standardised measure of parenting stress (Parenting Stress Index Short Form; Abidin, 1995) and biomarkers of emotional arousal and movement. Results pertaining to biomarkers will be considered in the forty-eight month report. 102 *PFL* mothers ($n_{Low} = 56$; $n_{High} = 46$), who were at various stages in the *PFL* programme, participated in the sub-study. There were a small number of differences

between participants and non-participants, indicating that mothers who chose to take part in the study were somewhat more disadvantaged than non-participants.

The results from the day reconstruction method showed that high treatment mothers reported higher levels of experienced positive emotion compared to the low treatment group, yet only for times when they were not with their *PFL* child. Consistent with these results, high treatment mothers reported higher global judgements of positive mood across the study day than low treatment mothers, yet not for times spent with their child(ren). There were no treatment effects for negative aspects of wellbeing including experienced negative affect and parenting stress. Finally, there were no treatment effects for participants' judgements of their life satisfaction. Overall, these results indicate that *PFL* is generating an impact on some aspects of everyday positive wellbeing of mothers.

Father Figures Focus Group Results

As part of the on-going *PFL* process evaluation, qualitative research was conducted with fathers of and father figures to *PFL* children (hereafter referred to as father figures). The aim of the study was to investigate the experience of fatherhood within the *PFL* community, and to determine father figure involvement in the *PFL* programme. Specifically, focus groups and interviews were conducted with 10 father figures ($n_{Low} = 4$; $n_{High} = 6$). The findings were analysed using thematic analysis as outlined by Braun and Clarke (2006). Six main themes were identified: a familial learning curve, role embodiment, the ecology of the father figure role, barriers to being a father figure, looking towards the future, and *PFL*.

A key finding was the similarity between the high and low treatment participants on their thoughts about fatherhood. Participants saw themselves as being on a learning curve with their families (theme 1), such that they were continuously learning, their children were developing, and the rest of the family was also adapting to the child and to the father figure's role. Participants described fatherhood as a multi-faceted role which was challenging, yet enjoyable and rewarding (theme 2). Participants saw their father figure role as being influenced by a number of factors including their own father, their relationship with the child's mother, the wider family environment, and society in general (theme 3). A number of barriers to engaging with their children were identified such as time, work, and the mother working as a gatekeeper in cases where relationship breakdown had occurred (theme 4). Future plans for children included instilling confidence to follow their dreams, and having more opportunities in life than the father figures themselves (theme 5).

The focus groups produced very rich data concerning the experiences of fatherhood among the *PFL* community, but relatively less data on father involvement in the *PFL* programme. Thus a key finding was the lack of knowledge and involvement of the father figures in *PFL*. Father figures were generally supportive of the programme, but felt disconnected from it (theme 6). Participants from the high treatment group reported a desire to be more involved with mentors, and participants from both treatment groups would like the opportunity to meet other father figures through discussion groups.

Conclusion

The thirty-six month evaluation of *PFL* suggests that the programme is well embedded in the community and is progressing very well, in line with, and in some cases exceeding results from other home visiting programmes at this time period. Over one-fifth of the individual tests and one-quarter of the step-down tests were statistically different across the high and low treatment group, representing the largest proportion of significant results to date. As hypothesised, there were a number of positive, significant findings in the areas of child development and parenting. There were also significant findings in the areas of child health, while results in the domains of the home environment and maternal health and wellbeing exceeded our expectations. Between twenty-four and thirty-six months the high treatment group were offered an additional treatment in the form of the Triple P Positive Parenting Program. Of the families who took part in the twenty-four month interview, over half participated in some form of Triple P. The Triple P

participants and non-participants outperformed the low treatment group on roughly the same proportion of individual tests, suggesting no supplementary benefit of the additional treatment. However, effects were concentrated on parenting and home environment measures for the Triple P group, and child health and development measures for the non-participating group. The highest level of attrition/disengagement was experienced at the thirty-six months interview; however, IPW analyses, which adjusted for differential attrition, revealed similar results overall. A sub-study conducted to investigate the impact of *PFL* on maternal wellbeing, found that high treatment mothers reported higher levels of experienced positive affect during time spent without their *PFL* child than low treatment mothers. Similarly high treatment mothers also reported higher global assessments of positive mood for the study day than low treatment mothers. However, no group differences were identified for experienced negative emotions or life satisfaction. The qualitative findings from the father figure focus groups identified few differences between the high and low treatment groups regarding their perceptions of fatherhood or their knowledge or involvement with *PFL*. However, participants displayed rich narratives about what it means to be a father figure in the *PFL* community and were generally supportive of the programme.

The reports of the six, twelve, eighteen, twenty-four, and thirty-six 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 Thirty-six Months

Kirsty is now three; she continues to live at home with her older brother and her mam and dad who are unmarried but in a long-term committed relationship. Her mam is happy with the neighbourhood she lives in and the support they get from family who live nearby. Kirsty is a lively, active child and her development is on track. She can now use four to five word sentences and loves to show people the numbers and colours that she knows. She likes to watch TV, but watches less TV than the other kids in her neighbourhood, and mostly watches shows that are age appropriate. There is a well-established routine in her house and her mam takes an active role in her learning and development. She now attends a childcare centre and her mother is very satisfied with the level of care that Kirsty receives there. She is in good health and is eating the recommended amount of protein for her age.

This year Kirsty's mam took part in the Triple P Program that taught her how to cope with difficult behaviour. She now knows how to avoid using punishments and harsh discipline and loves to engage with Kirsty on a regular basis. Kirsty and her mam still see their mentor about once a month, in their home. Kirsty's dad would like to be more involved in these visits but finds it difficult to find the time due to work and other commitments. Kirsty's mam has good mental wellbeing and she spends much of her day in a happy mood. Although she is relatively healthy, she has been to the GP four times in the last year. She doesn't use drugs and she only drinks occasionally and is optimistic about her future finances.

Chapter One

Background of the *PFL* Programme Thirty-Six Month Evaluation

1.1 Introduction

This report is the fifth 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 . The six month 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.[1] The six, twelve, eighteen, and twenty-four 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 thirty-six month evaluation, including new measures utilised as part of the thirty-six month interview, the results of the impact evaluation at thirty-six months, and new implementation data collected between twenty-four and thirty-six months. This report also includes the results of new sub-studies that were conducted as part of the *PFL* evaluation including the results of a qualitative study investigating the perspectives of father figures of the *PFL* children, and a study investigating maternal wellbeing among the *PFL* participants. In addition, between twenty-four and thirty-six months the *PFL* high treatment group were offered the Triple P Positive Parenting Program (Sanders et al., 2003). This report considers the impact of this additional treatment on child and parent outcomes.

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, and twenty-four months. It then presents a table of relevant findings from the literature on the impact of home visiting programmes at thirty-six months of age. Updated hypotheses are presented, as well as information regarding the collection of thirty-six 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 two 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 Six Month Evaluation

The six month evaluation suggested that the programme was developing well. In total, 257 six month interviews (n*Low* = 90; n*High* = 83; n*LFP* = 84) were completed. Analysis across eight domains revealed there were limited significant differences between the high and low treatment groups (14%). In addition, five of the 29 step-down categories (17%) remained significant in the multiple hypotheses analysis. These results are consistent with the programme evaluation literature which finds few treatment effects at this 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

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

directly to information on the *PFL* Tip Sheets delivered to participants during this period. Participants, on average received 14 home visits between baseline and six months, thus the intervention may not have been sufficiently intensive to generate significant treatment effects at this early stage. There were low levels of attrition (10% dropped-out and 8% disengaged) and high participant satisfaction, indicating that programme engagement was high which may result in positive future outcomes.

1.4 Summary of 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. 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 less significant differences reported 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 seven home visits between six and twelve months, which is lower than anticipated yet equivalent to the number of visits delivered during the first six months. There was minimal attrition between six and twelve months (two participants dropped-out) and participant satisfaction was high. This suggests that while engagement among participants is lower than anticipated (10% of the sample did not complete the twelve month interview), they are satisfied with the level of support they are receiving and they are choosing to remain in the programme.

1.5 Summary of 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. A similar number of significant differences were found at six and eighteen months, compared to those reported at twelve months, which is most likely due to the use of the same measures at both time points (e.g. the HOME). In total, 14% of the outcomes analysed were significant in the hypothesised direction at eighteen months, while 6% were statistically significant in the non-hypothesised direction, such that the low treatment group outperformed the high treatment group on these measures. 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 other evaluations of home visiting programmes failing to identify significant findings in these realms at eighteen months. Of the 27 step-down categories, five (18.5%) were statistically significant. The participants in the high treatment group received an average of 27 visits between recruitment during pregnancy and when the infant was eighteen months. The number and timing of mentor's home visits indicated that the average number of home visits was broadly similar over time. In terms of disengagement a total of 17% of the sample did not complete an eighteen month interview. At 15%, the level of attrition from *PFL* between baseline and eighteen months was quite low across the whole sample. Importantly, attrition was minimal between the twelve and eighteen month interview rounds.

1.6 Summary of Twenty-Four Month Evaluation

The twenty-four month findings represented the largest proportion of significant individual tests reported to date. A total of 239 (n*Low* = 84; n*High* = 82; n*LFP* = 73) interviews were completed. 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 hypotheses 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. The level of attrition between baseline and twenty-four months was quite low at 15% across the whole sample. None of the high or low treatment group dropped out between eighteen months and twenty-four months, and only one of the comparison group participants dropped out during this period. A total of 13% of the sample did not complete a twenty-four month interview.

1.7 Evidence on Short-term Effectiveness of Home Visiting Programmes

Previous reports reviewed the evidence on the effectiveness of home visiting programmes on outcomes observed up to twenty-four months of age. This section reviews the evidence on outcomes reported between twenty-four and thirty-six months. Several evaluations of home visiting interventions measure outcomes assessed when the child is thirty-six months old. Favourable results from these evaluations are observed predominantly in the domain of child development and school readiness; however these results are not consistent across programmes.

Table 1.1 reflects the outcomes from home visiting programmes from twenty-four to thirty-six months postpartum. 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 that were rated as either:

(1) High: 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: 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 below presents the findings observed between twenty-four and thirty-six months postpartum 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 outlined above 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 Early Outcomes for Home Visiting Programmes at 18-24 Months.

Outcome	Author	Sample Size	Programme
Child Development & School Readiness	Love et al. (2002)	744-746	Early Head Start – Home Visiting (EHS-HV)
	Roggman et al.(2009)	161	Early Head Start – Home Visiting (EHS-HV)
	Roggman & Cook (2010)	143	Early Head Start – Home Visiting (EHS-HV)
	Jones Harden et al. (2012)	950	Early Head Start – Home Visiting (EHS-HV)
	Fergusson, Horwood et al. (2005)	367-391	Early Start (New Zealand)
	Shaw et al. (2006)	92	Family Check-Up
	Connell et al. (2008)	662	Family Check-Up
	Shaw et al. (2009)	642	Family Check-Up
	Landsverk et al. (2002)	412	Healthy Families America
	King et al. (2005)	513	Hawaii Healthy Start
	Schwarz et al. (2012)	271	MOM Program
	Olds et al. (1994)	226	Nurse Family Partnership (NFP)
	Wagner & Clayton (1999)	363	Parents as Teachers
Child Health	Love et al. (2002)	950	Early Head Start – Home Visiting (EHS-HV)
	Fergusson, Horwood et al. (2005)	388-391	Early Start (New Zealand)
	Wagner & Clayton (1999)	352	Parents as Teachers
Positive Parenting Practices	Love et al. (2002)	744-950	Early Head Start – Home Visiting (EHS-HV)
	Jones Harden et al. (2012)	950	Early Head Start – Home Visiting (EHS-HV)
	Fergusson, Horwood et al. (2005)	391	Early Start (New Zealand)
	Dishion et al. (2008)	662	Family Check-Up
	Landsverk et al. (2002)	412	Healthy Families America
	Duggan, McFarlane et al. (2004)	541	Healthy Families America
	Wagner & Clayton (1999)	335	Parents as Teachers

	Measures used	Sig. Finding [Effect Size *d*]	Effect	Timing
	BSID MDI (standard score, score <85), Peabody Picture Vocabulary Test [PPVT]-III (standard score, score <85), parent-child semi-structured play (engagement, sustained attention, negativity toward parent), parent-child puzzle challenge task (engagement, persistence, frustration), Bayley Behaviour Rating Scale [BRS] (emotional regulation, orientation/engagement), CBCL	Parent-child semi-structured play (engagement)	Favourable	36 months
	BSID MDI	BSID MDI	Favourable	36 months
	Child aggression	None	None	36 months
	Parent reported that child has an Individualised Education Plan, Bayley Scales of Infant Development - Mental Devlopment Index [BSID MDI], Peabody Picture Vocabulary Test III (receptive vocabulary), Child Behavior Checklist [CBCL] (aggressive behaviour), engagement during play	Engagement during play [.19]	Favourable	36 months
	Early childhood education attendance, externalising score, internalising score, total behaviour score, Wechsler Preschool and Primary Scale of Intelligence [WPPSI] (verbal IQ, performance IQ, total IQ)	Duration of early childhood education attendance [.22], total internalising score [.26], total behaviour score [.24]	Favourable	36 months
	CBCL (aggression scale, physical aggression, destructive scale)	None	None	36 months
	CBCL (internalising, externalising, emotional reactivity, anxiety, depression, somatic problems, social withdrawal, transitions from internalising/externalising to "normal" from age 2 to 3)	None	None	36 months
	Eyberg Problem behaviour, CBCL (externalising, internalising)	Eyberg Problem behaviour [.23], CBCL (externalising [.23], internalising [.21])	Favourable	36 months
	Stanford Binet (short term memory, abstract/visual reasoning, verbal reasoning, verbal reasoning memory, sum of area, partial composite, communication survey), CBCL (anxious/depressed, withdrawn, sleep problems, somatic problems, aggressive, destructive, behaviour problems, externalising problems, internalising problems)	CBCL (somatic problems T score) [-.24]	Favourable	36 months
	Preschool Language scale [PLS-3] (total score, prevalence of severe delays, prevalence of any delays)	None	None	36 months
	WPPSI-III (verbal score, performance score, language score, full scale IQ)	None	None	33 months
	Stanford Binet	None	None	36 months
	Developmental Profile II [DPII] (physical development scale, cognitive development scale, communication development scale, self-help development scale, social development scale), Peabody Picture Vocabulary Test [PPVT]	DPII (self-help development scale – mean months differential [.25])	Favourable	36 months
	Child's health status, percentage of children in fair or poor health	None	None	36 months
	Number of visits made to family doctor in past 36 months, percent of children up to date with immunisations, percent of children up to date with well-child checks, percent attended hospital for any other reason, percent breastfed for six months or more, percent smoke-free home/smoke-free area, number of home safety features, percent with dental service	Number of visits made to family doctor in past 36 months [.24], percent of children up to date with well-child checks [.25], percent with dental service [.20]	Favourable	36 months
	Well-child visit, treated for illness, treated for injury, emergency room visit	None	None	36 months
	Parent-child play, parent-child semi-structured play (supportiveness, detachment, intrusiveness, negative regard), parent-child puzzle challenge task (supportive presence, quality of assistance, detachment, intrusiveness), daily routines items, percentage that usually use a car seat correctly, severity of discipline practices, reading practices, PSI (parental distress, parent-child dysfunctional interaction)	Parent-child semi-structured play (supportiveness), PSI (parental distress)	Favourable	36 months
	Parental discipline strategies, routines, parent-child observations, Parental Stress Index [PSI]	Parent supportiveness [.16], PSI [.14]	Favourable	36 months
	Child Rearing Practices Report [CRPR] (positive parenting attitude, non-punitive attitudes, total parenting score)	CRPR (positive parenting attitude [.26], non-punitive attitudes [.22], total parenting score [.27])	Favourable	36 months
	Parents positive behaviour support (indicated by: observations of parent involvement, positive reinforcement, prompting, and structuring, parent engagement, and coder impressions of proactive parenting)	Parents positive behaviour support [.33]	Favourable	36 months
	Nursing Child Assessment Satellite Training [NCAST] (child and caregiver total score), CTS-PC (nonviolent discipline)	None	None	36 months
	Conflict Tactics Scale Parent Child [CTS-PC] (occurrence and number of nonviolent discipline and time outs used)	None	None	36 months
	Knowledge of Infant Development Inventory [KIDI], Parenting Sense of Competence [PSOC] (total score, parenting satisfaction, parenting self-efficacy)	None	None	36 months

Outcome	Author	Sample Size	Programme
Reductions in Child Maltreatment	Lowell et al. (2011)	157	Child FIRST
	Roggman & Cook (2010)	143	Early Head Start – Home Visiting (EHS-HV)
	Jones Harden et al. (2012)	950	Early Head Start – Home Visiting (EHS-HV)
	Fergusson, Horwood et al. (2005)	382-388	Early Start (New Zealand)
	Landsverk et al. (2002)	382	Healthy Families America
	Duggan, McFarlane et al. (2004)	541	Healthy Families America
Home Environment	Love et al. (2002)	744-950	Early Head Start – Home Visiting (EHS-HV)
	Jones Harden, et al. (2012)	950	Early Head Start – Home Visiting (EHS-HV)
	Shaw et al. (2006)	92	Family Check-Up
	Landsverk et al. (2002)	412	Healthy Families America
	Duggan, McFarlane et al. (2004)	541	Healthy Families America
	Olds et al. (1994)	202-238	Nurse Family Partnership
	Wagner & Clayton (1999)	335	Parents as Teachers
Maternal Health	Fergusson, Horwood et al. (2005)	388	Early Start (New Zealand)
	Shaw et al. (2009)	651	Family Check-Up
	Landsverk et al. (2002)	412	Healthy Families America
	Duggan, Fuddy et al. (2004)	548	Healthy Families America
	Wagner & Clayton (1999)	343	Healthy Steps
	Schwarz et al. (2012)	271	MOM Program
Social support	Schwarz et al. (2012)	287	MOM Program
Household Factors and SES	Love et al. (2002)	950	Early Head Start – Home Visiting (EHS-HV)
	Fergusson, Horwood et al. (2005)	388	Early Start (New Zealand)
	Landsverk et al. (2002)	412	Healthy Families America
	Schwarz et al. (2012)	271	MOM Program
Reductions in Juvenile Delinquency, Family Violence, and Crime	Landsverk et al. (2002)	382	Healthy Families America
	Duggan, McFarlane et al. (2004)	548	Healthy Families America

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
	Family involvement with Child Protective Services [CPS]	Family involvement with CPS	Favourable	36 months
	Physical punishment	Physical punishment	Favourable	36 months
	Emergency Room visits due to accident or injury	None	None	36 months
	Hospital attendance, CTS-PC (physical assault by parent), parent report of agency contact for abuse or neglect	Percent attended hospital for accident, injury or poisoning [.22], CTS (percent severe/very severe physical assault by any parent) [.26]	Favourable	36 months
	CTS-PC (prevalence and frequency of: neglect, psychological aggression, mild physical assault, physical abuse	CTS-PC (psychological aggression frequency [-.27], mild physical assault frequency [-.29]	Favourable	36 months
	CTS-PC (psychological aggression, minor physical assault, severe physical assault, severe physical abuse, common corporal/verbal punishment, assault on child's self-esteem, hitting with an object, extreme physical abuse, shook child, neglect), substantiated CPS reports (abuse or neglect, threatened abuse neglect or harm)	None	None	36 months
	HOME (total score, internal physical environment, warmth, support of language and learning, harshness), Family Environment Scale [FES] family conflict	None	None	36 months
	Home Observation for the Measurement of the Environment [HOME]	None	None	36 months
	HOME (involvement)	HOME (involvement) [.27]	Favourable	36 months
	HOME (quality)	None	None	36 months
	HOME (acceptance, responsivity,)	None	None	36 months
	Hazardous exposures observed in home, HOME (total score)	Hazardous exposures observed in home	Favourable	34 months
	HOME (total score, parental responsivity, acceptance, organisation, appropriate play materials, involvement, opportunities for stimulation, language and literacy promoting behaviours, discipline)	HOME (acceptance of child's behaviour [-.28]	Unfavourable	36 months
	Pregnancy history, incidence of depression, smoking, substance use, family/social relationship problems	None	None	36 months
	Center for Epidemiologic Studies Depression scale [CES-D] (maternal depressive symptoms)	CES-D (maternal depressive symptoms) [.18]	Favourable	36 months
	Confidence in relationships, problem substance use, drinks per day, CES-D (total score, threshold for depression), MHI, repeat pregnancy, PSI	None	None	36 months
	CES-D, PSI, Mental Health Inventory [MHI] (score <67), drug use, problem alcohol use	None	None	36 months
	Mother had additional births	None	None	36 months
	Number of subsequent pregnancies, number of other children, Beck Depression Inventory [BDI-II] scale	None	None	33 months
	Referral to early intervention, receipt of early intervention, time to referral for early intervention, time to first receipt of early intervention	Referral to early intervention [0.23], receipt of early intervention [0.87], time to referral for early intervention, time to first receipt of early intervention	Favourable	33 months
	Continuous biological father presence child age 14-36 months, continuous male presence child age 14-36 months	Continuous male presence child age 14-36 months	Unfavourable	36 months
	Percent single-parent family, percent separated from partner, mean number of family changes, percent welfare dependent, mean family income per week, percent income inadequate/very inadequate, mean amount of debt	None	None	36 months
	Mother attended school, earned HS degree of currently enrolled, employed, someone in household employed welfare receipt, WIC receipt, food stamps, emergency food, section 8 housing, homeless shelter, legal aid, child support enforcement, use of respite care, adult education	Mother attended school [.25]	Favourable	36 months
	Maternal level of education, number of months employment in last year, number of months school in last year, number of months in last 33 months, receipt of public assistance, use of homeless service, monthly income	None	None	33 months
	CTS (partner violence resulting in injury, incident of physical assault frequency and prevalence), restraining order, family court	None	None	36 months
	CTS (partner psychological abuse, physical abuse, incident resulting in injury)	None	None	36 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	5 out of 8 programmes	• Improved engagement during play, including parent-child semi-structured play • Fewer behavioural problems • Improved cognitive development
Child Health (including reductions in child maltreatment)	4 out of 5 programmes	• Fewer GP visits • Up to date with well-child checks • Attending dentist • Reduced involvement with Child Protective Services (USA) • Reduced hospital attendance for accident, injury, or poisoning • Reduced occurrence of physical punishment, assault & psychological aggression by a parent
Parenting	3 out of 5 programmes	• Reduced parental stress • Improved parental supportiveness of child during semi-structured play • Increased positive, non-punitive parenting attitudes • Increased parent positive behaviour support
Home Environment	2 out of 5 programmes	• Greater parental involvement with child in the home • Reduced child exposure to hazards in home
Maternal Health	1 out of 5 programmes	• Reduced maternal depressive symptoms
Maternal Social Support	1 out of 1 programme	• Increased referral to and receipt of early intervention • Earlier referrals and receipt of intervention
Childcare	1 out of 1 programme	• Longer early childhood education attendance
Household Factors & SES	1 out of 4 programmes	• Increased likelihood of mother having attended school

1.8 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 hypothesise that the programme will have an effect on other child and family outcomes (e.g. social support, service use, maternal health and wellbeing). 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 thirty-six months of age are informed by the evidence described. Results from previous studies indicate that at thirty-six months home visiting programmes have moderate effects on child development, child health, and parenting. Consistent with these findings, we expect that the impact of *PFL* on these three domains will also be moderate. However, given that high treatment families were offered Triple P training at twenty-four months we expect a greater number of favourable effects in the domains of child development and parenting at thirty-six relative to twenty-four months. As effects on maternal social support and childcare were each only investigated by one programme, it is difficult to ascertain the impact of *PFL* on these areas. We hypothesise that the effects in both domains will be limited. As reported in other studies, we expect to find limited positive effects in the areas of the home environment, maternal health, and household factors and SES. Consistent with the evaluation design, we anticipate few significant differences between the low treatment group and the comparison group.

1.9 Description of Thirty-Six Month Survey & Data Collection Process

Between July 2011 and July 2014, a sixth research interview was conducted by the *PFL* evaluation team. The interviews took place when the *PFL* child was between two weeks before their third birthday and up to six months after their birthday. In total, 217 thirty-six month interviews (n*High* = 75; n*Low* = 76; n*LFP* = 66) were completed. The average age of the target child at the time of the interview was 36.96 months old (SD = 5.83 weeks). The dropout rate between twenty-four and thirty-six months was minimal. None of the high or low treatment group dropped out and only two of the comparison group participants dropped between twenty-four and thirty-six months. However, disengagement rates across the high and low treatment groups were 16% and 19% respectively, and the corresponding rate for the comparison group was 21%. This represents the highest level of disengagement since the programme commenced. A comprehensive analysis of attrition rates may be found in Chapter 3 of this report.

The thirty-six 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 either in a local community centre. The majority of participants completed the interview in their homes (88% high treatment group, 82% in the low treatment group, and 100% in the comparison community). Each participant was given a €20 shopping voucher after the thirty-six month interview was completed as a thank you for taking part in the interview.

A number of questions/measures asked in the interview had also been used at previous time points, while a number of questions/measures were new. The repeated questions related to family demographics and socioeconomic profile, maternal physical and psychological health, substance use by the mother, family risk factors, the home environment, parenting stress, use of childcare, child motor skills, cognitive development, behavioural, and emotional functioning, social-emotional development, and child health. New questions added to the thirty-six month interview included items related to neighbourhood quality and crime, TV habits, diagnoses of chronic illnesses and physical disabilities, parental quality of life and romantic relationship quality, parental acceptance, and parenting styles and dimensions.

The thirty-six month survey was divided into 10 modules, each containing questions with a common theme.

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

Similar to previous reports, this report focuses on eight domains incorporating 43 categories and 204 outcome measures. The domains and categories within each domain are – child development (Ages Stages Questionnaire, Brief Infant Toddler Social and Emotional Assessment, Infant and Toddler Social and Emotional Assessment, Developmental Profile-3, Achenbach Child Behaviour Checklist), child health (child physical health, mother's health decisions for her child), parenting (Parenting Daily Hassles Scale, Parenting Styles and Dimensions Questionnaire, Parental Acceptance – Rejection Questionnaire, child protective services involvement, activities with child, parental attitudes toward education, parental monitoring of TV), home environment (Home Observation for Measurement of the Environment, Supplement to the HOME for Impoverished Families, Family Environment Scale), maternal health and wellbeing (maternal physical health, maternal mental health, drug and alcohol use), social support (father involvement, support from relatives, friends and neighbours, participation in community services), childcare and service use (childcare use, type & satisfaction, and service use), and household factors and SES (household factor measures, parental education, parental employment, household finances and expectations of

future finances, Difficult Life Circumstances, mother's satisfaction with neighbourhood, Neighbourhood Quality Evaluation Scale, Neighbourhood Criminal Events Scale). Note that while the same domains are investigated at each time point the measures included in the thirty-six month report may differ to those included at previous time points.

1.10 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 thirty-six 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 which examines changes in child and parent outcomes over time. Chapter 5 presents a summary of the results from the Triple P Positive Parenting Program implemented with the *PFL* high treatment group between twenty-four and thirty-six months. Chapter 6 presents a summary of the findings from an auxiliary study on maternal wellbeing. Chapter 7 presents a summary of the findings from a qualitative investigation of *PFL* with male role models. Chapter 8 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 thirty-six 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 thirty-six 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 four previous reports. This chapter contains relevant literature for the new measures which are not included in previous reports and considers the relevance and impact of previous home visiting programmes on all measures at thirty-six 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 reports at twelve, eighteen, and twenty-four months 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 presented in the report. 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 thirty-six months.

2.2 Methods & Description of Outcome Tables

A full description of the methodology used to analyse each wave of outcomes 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 hypotheses testing, and the procedure for dealing with missing data[2].

[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 twenty-four month data is very low; less than 7% of data were missing for any item in each psychometric scale, with the majority of scales missing less than 1% of data overall. 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, observations with missing data were excluded from the analysis. On average, over 99% of data were present for single item measures.

The following information is included in the outcomes tables presented in this report and provides a reference for interpreting the results.

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	The mean responses for the low treatment group and high treatment group are compared in multiple ways. In this chapter the data are first grouped by *PFL* treatment status (low treatment and high treatment) to examine thirty-six month differences within the *PFL* cohort. In Chapter 3 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-value*s 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, all the analyses comparing the thirty-six month outcomes of the high and low treatment groups use an alternative approach called Permutation-based hypothesis testing. This approach has been found to be appropriate for small samples and was used to analyse data for a similar evaluation of Perry Preschool Program by Heckman and colleagues (2010). The individual *p-value* represents the probability of observing differences between 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 .10 is considered to be statistically significant. A *p-value* of less than 0.10 (10%), 0.05 (5%), and 0.01 (1%) conveys that the probability that the difference between the two groups is due to chance is less than 10%, 5%, or 1% respectively. Low *p-value*s (i.e., significant results) would be a positive outcome indicating that the high treatment group is outperforming the low treatment group. *p-value*s 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 166 outcome measures are considered in this report, it is possible that we may reject some of these null hypotheses 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. To illustrate the Step-down procedure, consider the null hypothesis of no treatment effect for a set of, say, K outcomes jointly. The complement of the joint null hypothesis is the hypothesis that there exists at least one hypothesis out of K that we reject. We apply the analysis 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 family-wise error rate (FWER), the probability of yielding one or more false positives out of a set of hypotheses 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 0.2 to 0.8 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

Early child development is one of the most important phases of life which strongly influences wellbeing, mental health, physical health, competence in literacy and numeracy, criminality, and economic participation throughout life (Illig, 1998; WHO, 2007). It is a period of great opportunity, however it is also a period of great vulnerability to negative influences, and constitutes a unique phase for capitalising on developmental forces (WHO, 2007). Research has identified strong associations between a family's economic, educational, social, emotional, environmental and parenting resources and practices and the increased likelihood of cognitive and language delays, behavioural problems, and undesirable life course outcomes for children (see Illig, 1998 for a review). For example, a study investigating the prevalence of language delays in a US sample of socioeconomic at-risk children found that at age three years, one in 10 children, four times the expected proportion, had severe delays in language development, while nearly half of all children met criteria for at least a mild degree of language delay (King et al., 2005). Evidence suggests that early intervention programmes are an effective means of reducing these types of socioeconomic inequalities and promoting healthy child development (Gomby, 2005; Kahn & Moore, 2010). Although physical, cognitive, and social development are interrelated they will be considered here separately as a framework to discuss the research in the area.

PHYSICAL DEVELOPMENT: GROSS AND FINE MOTOR SKILLS

Physical development is usually thought of in terms of locomotion and postural development on the one hand and manipulative skills on the other (Carr, 2006). The development of these skills is important in facilitating the growing child's need and desire to explore their environment and to seek out new interactions with people and things (Williams et al., 2013). At three years of age children can typically stand on one (preferred) foot for five seconds, stand or walk on tiptoe and walk alone up stairs using alternating feet, coming down stairs two feet to a step (Sheridan, 2004; American Academy of Pediatrics, 2013). They can turn around obstacles and corners while running and also while pushing and pulling large toys (Sheridan, 2004). In the recently published Growing Up in Ireland (GUI) study, most three year old Irish children were able to throw a ball overhand (95%) and the majority were also able to stand on one leg (87%) (Williams et al., 2013). The authors also noted that children who played physically active games at home tended to have better gross motor skills at age three years (Williams et al., 2013).

In terms of manipulative skills or fine motor skills, three year old children tend to hold pencils near the point in their preferred hand, between the first two fingers and thumb (Sheridan, 2004). They can copy a circle, imitate a cross, and tend to draw a person with a head and one or two other features or parts (American Academy of Pediatrics, 2013, Sheridan, 2004). They can build a tower of nine or 10 cubes and by three years six months they can build a bridge of three cubes using two hands cooperatively (Sheridan, 2004). The GUI study found that most Irish three year olds were able to use a pencil and play with small objects such as a jigsaw puzzle (Williams et al., 2013). They also found that about half of the children (51%) were using a pincer-type grip and that girls had better early fine motor skills than boys (Williams et al., 2013).

COGNITIVE AND LANGUAGE DEVELOPMENT

Although cognitive and language development are to some degree independent of each other many of the skills interact, for example the statement 'all gone' depends on the child developing the concept of object permanence (Gopnik & Meltzoff, 1987 as cited in Carr, 2006). Three year olds typically have a large vocabulary intelligible even to strangers, but speech still contains many immature phonetic substitutions and unconventional grammatical forms (Sheridan, 2004). They can give their full name, sex, and sometimes age (Sheridan, 2004). Children of this age tend to use personal pronouns and plurals correctly and also most prepositions (Sheridan, 2004). They can count by rote up to 10 or more, but have little appreciation of quantity beyond two or three (American Academy of Pediatrics, 2013; Sheridan, 2004). The GUI study found that Irish children were somewhat ahead of a standardised sample in terms of cognitive development as assessed by the British Ability Scale (Williams et al., 2013). They also found that on average, boys fared worse on tests of cognitive ability and parent-report indicators of speech and

language problems than girls (Williams et al., 2013). Interestingly, children in the study who did not meet expected scores on developmental measures at nine months were at greater risk of lagging behind their peers at three years (Williams et al., 2013).

PERSONAL, SOCIAL, AND EMOTIONAL DEVELOPMENT

When socio-emotional milestones are not negotiated successfully, children are at risk of multiple behaviour problems and poor academic performance (Denham et al., 2009). Healthy social-emotional development relates to children's capacity to experience, manage, and express emotion, form close and secure relationships, and to explore and learn from their environments (Zero to Three, 2001).

By three years of age children's general behaviour is amenable and they can be affectionate and confiding, often showing affection towards younger siblings (Sheridan, 2004). They tend to like helping adults with domestic activities and they make an effort to keep surroundings tidy (Sheridan, 2004). They display an increasingly inventive fantasy play and engage in make-believe play with other children, showing an understanding of sharing their playthings (American Academy of Pediatrics, 2013; Sheridan, 2004). Pre-schoolers also increasingly use language for regulating emotions and show increased insight into others emotions (Saarni, 1999 as cited in Carr, 2006). Children are becoming more independent and can often dress and undress themselves (American Academy of Pediatrics, 2013; Sheridan, 2004). The GUI study found that children in Ireland had significantly fewer difficulties on the total difficulties scale of the Strengths and Difficulties Questionnaire (SDQ) at three years compared to children participating in the (UK) Millennium Cohort Study (William et al., 2013). They also found that the prevalence of childhood behavioural problems was strongly correlated with socio-demographic and family characteristics, the latter including parenting dimensions of warmth and hostility (Williams et al., 2013).

IMPACT OF HOME VISITING INTERVENTIONS ON CHILD DEVELOPMENT AT THIRTY-SIX MONTHS

Overall, the majority of home visiting programmes find positive effects on child development at the thirty-six month time period. The Early Head Start programme found a number of positive impacts on child development from age two to three including cognitive and language development (Love et al., 2002; Roggman et al., 2009) and social-emotional development, with significantly fewer children in the treatment group scoring in the at-risk range of developmental functioning (Jones Harden et al., 2012; Love et al., 2002). Fergusson, Horwood et al. (2005) assessed the Early Start Programme in New Zealand and found statistically significant reductions in problem behaviour scores for externalising, internalising, and total behaviour problems in children receiving the intervention. They also assessed cognitive development, and although children in the programme scored slightly higher on all tests of IQ, none reached statistical significance. The Family Check-Up is a brief family intervention which found that children involved in the programme showed less externalising and internalising problem behaviour than those in the control group (Shaw et al., 2009). Reporting on the Healthy Families America programme in California, Lansverk et al., (2002) found that mothers of children in the intervention group reported fewer somatic problems than those in the control group. The Parents as Teachers programme also found positive effects such that children involved in the study showed greater persistent problem solving with novel tasks (Drotar et al., 2009) and greater self-help skills as measured by the Developmental Profile II scale (Wagner & Clayton, 1999) than children in the control group.

However, a number of studies found no programme effects on child development at thirty-six months. In contrast to above, one study reporting on the Early Head Start programme in Bear River found no programme effects in the area of child development (Roggman & Cook, 2010). The MOM home visiting programme also found no significant differences in cognitive outcomes at age thirty-three months between the children in the intervention and control groups (Schwarz et al., 2012). Two studies reporting on the Family Check Up programme located in different sites than above, found no programme effects on social and emotional outcomes at thirty-six months (Shaw et al., 2006; Connell et al., 2008). In addition, a study reporting on the Hawaii Healthy Start programme found no differences in the prevalence of language delays between children who did and did not receive home visitation (King et al., 2005). Finally, the Nurse Family Partnership also found no differences between groups in terms of intellectual functioning as measured by the Standford Binet IQ test (Olds et al., 1994).

2.3.1 Child Development Instruments

AGES AND STAGES QUESTIONNAIRE

Child development in the *PFL* evaluation was assessed using the thirty-six month version of the Ages and Stages Questionnaire, except in one case where it was necessary to use the forty-two month version of the 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=0.49) measures the child's understanding of language, naming of items and word combinations. The Gross Motor domain (α=0.49) measures the child's walking, running, and jumping movements. The Fine Motor domain (α=0.69) assesses the child's finger and hand movements, including stacking and threading. Problem Solving (α=0.55) measures the child's ability to follow instruction, pretence, and problem solving. Finally, the Personal-Social domain (α=0.47) provides a rating of eating skills, solitary play, and self-awareness. During the interview, the interviewer asked the mother questions related to different activities her child is capable of. The mother responded by indicating if her child exhibits 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 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 = 38.7; Gross Motor = 35.7; Fine Motor = 30.0; Problem Solving = 38.6; and Personal-Social = 38.7). 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 are 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=0.82) 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 if 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 was 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 59 and suggests that children with scores above this cut-off may be at risk. In line with this cut-off score, a binary variable was calculated to illustrate if the child was at risk of poor socio-emotional development.

BRIEF INFANT-TODDLER SOCIAL AND EMOTIONAL ASSESSMENT

The Brief Infant-Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2006) is a shortened version of the Infant-Toddler Social and Emotional Assessment (ITSEA). The 42-item screening tool measures social-emotional/behavioural problems and delays in competence in children aged twelve to thirty-six months. However it should be noted the *PFL* children sampled at this time point were between thirty-six and forty-two months old. The BITSEA yields a Problem score (α=0.87) and a Competence score

(α=0.71). Competencies (11 items) include areas of attention, compliancy, mastery, motivation, pro-social peer relations, empathy, play skills and social relatedness. The Problem behaviour score includes three subscales: Externalising (6 items: α=0.79), Internalising (8 items: α=0.65), and Dysregulation Problems (8 items: α=0.69). In addition, there are two scales which identify possible clinical problems: the Autism Spectrum Disorder (ASD) scale (17 items: α=0.40) which identifies whether the child may benefit from an assessment for ASD and the 'red flag' scale (14 items: α=0.72) which indicates a potential clinically significant problem. The interviewer asked mothers to verbally rate each item on a 3-point scale (0=not true/rarely, 1=somewhat true/sometimes, 2=very true/often). Items were summed to obtain a Problem score whereby higher scores indicate greater levels of social-emotional or behavioural problems and a Competence score whereby lower scores indicate possible delays/deficits.

INFANT-TODDLER SOCIAL AND EMOTIONAL ASSESSMENT: PEER INTERACTION: EXPERIENCES WITH OTHER YOUNG CHILDREN

The Infant-Toddler Social and Emotional Assessment (ITSEA; Carter & Briggs-Gowan, 2006) is a 168-item questionnaire which is completed by parents or caregivers of children aged twelve to thirty-five months old. As with the BITSEA, it should be noted that the *PFL* children sampled were above the manual stated age range for this measure. The items cover four domains (externalising, internalising, dysregulation, and competence), which are then divided into 17 subscales. The *PFL* evaluation includes two subscales - Pro-Social Peer Relations (α=0.63) and Peer Aggression (α=0.74), incorporating 11 items. These subscales focus on the child's behaviour around other children, for example: 'Takes turns when playing with others'. For each item the participant was asked to describe her child's behaviour over the last month with the following options not true/rarely, somewhat true/sometimes, or very true/often. These were scored as 0, 1, and 2 respectively. Items were summed to obtain a Pro-Social Peer Relations score whereby higher scores indicate greater levels of positive peer competencies and engagement in pro-social activities, and a Peer Aggression score whereby higher scores indicate greater levels of negative peer relations.

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 an 'other problems' (α=0.83) category. These 8 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 are scored as 0, 1, and 2 respectively. From the 100 questions, eight raw scores are produced (seven syndromes and other category, as above). Further, the raw scores of the emotionally reactive, anxious/depressed, somatic complaints, and withdrawn subscales are totalled as an Internal Problems score. Correspondingly, the attention problems and aggressive behaviour syndromes are totalled to produce the External Problems score. Finally, the totals of all seven syndromes plus the other problems subscale are 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. 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=0.84) using a 38-item scale. Each of the items refers to tasks which require cognitive skill and are arranged in order of difficulty, for example: 'Does your child point to at least 20 things or pictures when they are named' 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 above the average score, 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

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

ASQ SCORES

Within the ASQ scores category, all six child development measures were in the hypothesised direction and one of these, ASQ Problem Solving score, was statistically significant. The high treatment group scored an average of 49.79 on this subscale while the low treatment group scored an average of 45.61 ($p<.05$, $d=.37$) indicating that children in the high treatment group were more likely to display developmentally appropriate problem-solving 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 statistically significant ($p<.10$) and that the effect was driven by the significant results found for the ASQ Problem Solving score.

ASQ CUT-OFF SCORES

Within the ASQ cut-off scores category, which measures the proportion of children at risk of development delay, five of the six measures were in the hypothesised direction. One of these differences was statistically significant. 11% of children in the high treatment group were at risk of developmental delay regarding problem solving skills, compared with 22% of children in the low treatment group ($p<.05$, $d=.29$). The step-down test showed that the joint effect of all six measures in the ASQ Cut-off Scores category was not statistically significant

BITSEA

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

BITSEA SUBDOMAINS

Two of the five BITSEA subdomains were in the hypothesised direction, but none reached statistical significance. The step-down test showed that the joint effect of the BITSEA subdomains was not statistically significant.

ITSEA

Within the ITSEA category, one measure was in the hypothesised direction and one was in the non-hypothesised direction, however, neither was statistically significant.

CBCL DOMAINS

Within the CBCL category, all three measures were in the hypothesised direction and two were statistically significant. The high treatment group scored an average of 22.45 on the total score, while the low treatment group scored an average of 26.82 ($p<.10$, $d=.23$), indicating that children in the high treatment group were reported to display less behavioural problems. In addition, the high treatment group scored an average of 7.61 on the external problems score, compared with the low treatment score of 9.24 ($p<.10$, $d=.23$). 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, two of the three measures were in the hypothesised direction and statistically significant. Only 1% of the high treatment group were rated at the clinical level for the CBCL total score, compared with 8% of the low treatment group ($p<.05$, $d=.31$), indicating that the high treatment group were less likely to have a behavioural problem. In terms of external problems, only 1% of the high treatment group were rated as having problems at the clinical level, compared with 7% of the low treatment group ($p<.10$, $d=.27$). 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 total problems cut-off ($p<.10$) and the external problems cut-off ($p<.10$).

CBCL SUBDOMAINS

Within the CBCL subdomains category, all eight measures were in the hypothesised direction and four were statistically significant. The high treatment group scored an average of 1.26 on the somatic complaints subdomain, compared with the low treatment score of 1.91 ($p<.05$, $d=.34$). The high treatment group scored on average 2.22 on the sleep problems domain compared to an average score of 2.79 for the low treatment group ($p<.10$, $d=.23$). In addition, the high treatment group scored significantly higher on the other problems subdomain ($p<.10$, $d=.22$) and the aggressive behaviour subdomain ($p<.10$, $d=.22$). The step-down test showed that the joint effect of the CBCL subdomains was statistically significant ($p<.10$) and was driven by the somatic complaints result.

NON STEP-DOWN MEASURES

All four of the non step-down measures were in the hypothesised direction. There were three statistically significant differences between the high and low treatment groups. The high treatment group scored an average of 101.90 on the ASQ standardised total score, compared to the low treatment score of 97.91 ($p<.10$, $d=.27$). Furthermore, 53% of the high treatment group were scored as above average on the DP-3 compared to 36% of the low treatment group ($p<.05$, $d=.35$). Finally, the high treatment group scored on average 114.66 on the DP-3 standardised score as compared to the low treatment group who scored on average 109.82 ($p<.05$, $d=.34$). These indicate that the high treatment group were displaying more advanced cognitive abilities than the low treatment group.

In addition to examining the difference in means between the high and low treatment groups, tests were conducted to examine the equality of standard deviations between the groups. On 11 of the 28 continuous child development measures the high treatment group had a significantly lower variance than the low treatment group. This implies that the high treatment group's outcomes exhibited a lower degree of dispersion around the group mean, relative to the low treatment group.

	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 Problem Solving Score	147	(73/74)	49.79	(10.46)	45.61	(12.52)	*p<.05*	p<.10	0.37
	ASQ Communication	150	(75/75)	53.33	(7.94)	52.00	(9.23)	ns	ns	0.16
	ASQ Personal Social Score	150	(75/75)	55.00	(6.15)	53.87	(8.57)	ns	ns	0.15
	ASQ Gross Motor Score	150	(75/75)	54.67	(8.63)	53.33	(9.67)	ns	ns	0.15
*	ASQ Social-Emotional Score	150	(75/75)	31.53	(25.42)	34.33	(28.59)	ns	ns	0.10
	ASQ Fine Motor Score	147	(73/74)	45.00	(15.57)	43.51	(15.14)	ns	ns	0.10
ASQ Cut-off scores										
*	ASQ Problem Solving Cut-off	147	(73/74)	0.11	(0.31)	0.22	(0.41)	*p<.05*	ns	0.29
*	ASQ Communication Cut-off	150	(75/75)	0.04	(0.20)	0.08	(0.27)	ns	ns	0.17
*	ASQ Social-Emotional Cut-off	150	(75/75)	0.09	(0.29)	0.12	(0.33)	ns	ns	0.09
*	ASQ Personal Social Cut-off	150	(75/75)	0.03	(0.16)	0.04	(0.2)	ns	ns	0.07
*	ASQ Fine Motor Cut-off	147	(73/74)	0.16	(0.37)	0.18	(0.38)	ns	ns	0.03
*	ASQ Gross Motor Cut-off	150	(73/74)	0.07	(0.25)	0.07	(0.25)	ns	ns	0.00
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)										
	BITSEA Competence Score	151	(75/76)	19.03	(2.68)	18.61	(2.86)	ns	ns	0.15
*	BITSEA Problem Score	151	(75/76)	7.37	(5.51)	8.03	(7.74)	ns	ns	0.10
BITSEA subdomains										
*	BITSEA External Problems	151	(75/76)	1.15	(1.56)	1.45	(2.29)	ns	ns	0.15
*	BITSEA Dysregulation	151	(75/76)	2.84	(2.36)	3.22	(3.09)	ns	ns	0.14
*	BITSEA Internal Problems	151	(75/76)	1.59	(1.71)	1.59	(1.67)	ns	ns	0.00
*	BITSEA Red Flag	151	(75/76)	2.37	(2.29)	2.34	(2.74)	ns	ns	0.01
*	BITSEA Autism	151	(75/76)	3.52	(3.21)	3.47	(3.37)	ns	ns	0.01
Infant-Toddler Social and Emotional Assessment (ITSEA)										
	ITSEA Pro-social Behaviour	149	(74/75)	1.67	(0.33)	1.62	(0.38)	ns	ns	0.14
*	ITSEA Aggression	149	(74/75)	0.19	(0.27)	0.21	(0.32)	ns	ns	0.07
Child Behaviour Checklist (CBCL) domains										
*	CBCL Total Score	150	(74/76)	22.45	(15.79)	26.82	(21.49)	*p<.10*	ns	0.23
*	CBCL External Problems	150	(74/76)	7.61	(5.72)	9.24	(8.42)	p<.10	ns	0.23
*	CBCL Internal Problems	150	(74/76)	6.16	(5.83)	7.20	(6.74)	ns	ns	0.17
Child Behaviour Checklist (CBCL) domains cut-off scores										
*	CBCL Total Score cut-off	150	(74/76)	0.01	(0.12)	0.08	(0.27)	*p<.05*	*p<.10*	0.31
*	CBCL External Problems cut-off	150	(74/76)	0.01	(0.12)	0.07	(0.25)	*p<.10*	*p<.10*	0.27
*	CBCL Internal Problems cut-off	150	(74/76)	0.07	(0.25)	0.08	(0.27)	*ns*	*ns*	0.04
Child Behaviour Checklist (CBCL) subdomains										
*	CBCL Somatic Complaints	150	(74/76)	1.26	(1.59)	1.91	(2.22)	*p<.05*	p<.10	0.34
*	CBCL Sleep Problems	150	(74/76)	2.22	(2.26)	2.79	(2.75)	*p<.10*	ns	0.23
*	CBCL Other Problems	150	(74/76)	6.46	(4.62)	7.59	(5.77)	p<.10	ns	0.22
*	CBCL Aggressive Behaviour	150	(74/76)	5.89	(4.79)	7.14	(6.73)	p<.10	ns	0.22
*	CBCL Attention Problems	150	(74/76)	1.72	(1.52)	2.09	(2.14)	ns	ns	0.20
*	CBCL Anxious/Depressed	150	(74/76)	1.89	(1.89)	2.08	(1.91)	ns	ns	0.10
*	CBCL Emotionally Reactive	150	(74/76)	1.88	(2.09)	2.05	(2.17)	ns	ns	0.08
*	CBCL Withdrawn	150	(74/76)	1.14	(1.56)	1.16	(1.64)	ns	ns	0.01
Non Step-down Measures										
	ASQ Standardised Total Score	147	(73/74)	101.90	(13.84)	97.91	(15.62)	*p<.10*	-	0.27
	DP3: Cognitive Development above average cut-off	150	(74/76)	0.53	(0.50)	0.36	(0.48)	*p<.05*	-	0.35
	DP3: Cognitive Development standardised score	150	(74/76)	114.66	(14.39)	109.82	(14.02)	p<.05	-	0.34
*	Child receiving special services	151	(75/76)	0.11	(0.31)	0.17	(0.38)	ns	-	0.19

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

Child health is one of the primary outcomes of the *PFL* evaluation and an area which many home visiting programmes seek to improve by supporting families and facilitating access to community-based services for parents and their children. People living in disadvantaged areas are at increased risk of developing a number of preventable diseases such as heart disease, diabetes, respiratory infections, and obesity (Galobardes et al., 2004; Komro et al., 2011; German & Latkin, 2012) and are also more likely to engage in behaviours which may have an adverse effect on their health and the health of their children.

GENERAL HEALTH

Weight is a good indicator of general health and nutrition. It reflects the combined effects of energy intake from food and energy output. At thirty-six months it is suggested that an average boy weighs 14.3kg and measures 95.4cm and an average girl weighs 13.9kg and measures 94.4cm (Centers for Disease Control and Prevention clinical growth charts: http://www.cdc.gov/growthcharts/clinical_charts.htm). GUI reported that the average weight of a thirty-six month old was 15.6kg and the average height was 96.2cm, with a quarter of all three year olds measuring as overweight or obese (Williams et al., 2013). Research has demonstrated that children living in socially disadvantaged areas may be at increased risk of weight problems such as obesity (Greves Grow et al., 2010). A report published by the HSE noted that, in developed countries such as Ireland, levels of obesity are higher among lower socioeconomic groups. Although Type 2 diabetes and other weight-related diseases were previously associated with middle age, they are now being increasingly observed in children (HSE, 2005).

LONG TERM HEALTH

The GUI study reported that almost 16% of three year olds had a longstanding illness, disability, or other ongoing health condition, and the most commonly reported longstanding illness among three year olds in Ireland was respiratory illness (Williams et al., 2013). A doctor diagnosis of asthma was reported by 6% of the sample followed by eczema and skin allergies which were reported by 4%. The study also reported that the extent to which the experience of chronic illness hampered children's daily activities appeared to increase with social disadvantage.

ILLNESS, ACCIDENTS, AND INJURIES

Data from GUI indicated that asthma is currently the most commonly reported illness among three year olds (Williams et al., 2013). The contraction of illnesses and use of medication such as antibiotics can have negative implications for future child health (Johnson et al., 2005; Marra et al., 2006). Almost two-thirds (65%) of the GUI sample had received at least one course of antibiotics in the preceding twelve months, and 16 percent of three year olds in the study had experienced an accident or injury that required hospital treatment or admission over their lifetime (Williams et al., 2013). In Ireland, no immunisations are scheduled for thirty-six months. At four to five years the 4 in 1 vaccine (diphtheria, tetanus, whooping cough [pertussis] and polio), and MMR vaccine (measles, mumps, and rubella) are administered by the GP or school (HSE, 2013a).

APPROPRIATE FOOD

One way of tackling weight issues such as obesity is through diet. Balanced nutrition early in life is essential for health later in life (Schwartz et al., 2011). A varied, healthy diet throughout childhood is important for physical development, cognitive development, and health. Feeding difficulties can affect child nutrition and toddlers need to consume a healthy diet consisting of the right amounts of dairy, fruit/vegetables, grains, fats, and protein (Gottesman, 2002). The GUI study provides evidence of a social gradient in relation to children's diet (Williams et al., 2013). This study found that children of less-educated caregivers were more likely to have consumed energy-dense food such as hamburgers and crisps, but less likely to have eaten fresh fruit or vegetables, in the 24 hours preceding the interview.

IMPACT OF HOME VISITING INTERVENTIONS ON CHILD HEALTH AT THIRTY-SIX MONTHS

A number of home visiting programmes have sought to improve child health at thirty-six months, and evaluations of these programmes have found mixed results. Evaluations of a number of large scale home visiting programmes, including the Early Head Start and Parents as Teachers programmes, found no favourable effects in relation to child health. These studies measured the child's health status, percentage of children reported as having fair or poor health, well-child visits, treatment for illness and/or injury, and emergency room visits (Wagner & Clayton, 1999; Love et al., 2002). Fergusson et al. (Fergusson, Horwood et al., 2005; Fergusson, Grant et al., 2005) found that the Early Start programme had favourable effects on the number of visits made by a family to the doctor in the past thirty-six months, the percentage of children up to date with well-child checks, and the percentage of children enrolled with dental services. The programme was not found to have any effect on the percentage of children who had attended hospital for any reason, the percentage of children breastfed for six months or more, or the percentage of children living in a smoke free home. An evaluation of Early Start by Fergusson, Horwood, and colleagues (2005) found that the programme had a favourable effect on the percentage of children who attended hospital for accidents, injuries, or poisoning. Few home visiting programmes have examined programme effects on chronic illnesses. The Childhood Asthma Prevention Study, a home visiting programme for wheezing infants from low-income families, has been shown to reduce environmental exposure and improve illness management; however, the programme has not been found to decrease asthma among children with early wheezing as a whole (Klinnert et al., 2005, 2007).

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 had an accident, chest infection, asthma, skin problems, or an ear infection in the past 12 months.

LONG TERM CHILD HEALTH

Two binary variables were created based on whether the child had any ongoing diagnosed chronic illness and 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. A binary variable was also created to signify whether or not the child was meeting the overall dietary guidelines. For each food groups 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 recommendation as indicated by the binary diet variable. 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 food such as sugars and fats).

WEIGHT AND HEIGHT

The child's current weight and height were measured by the interviewer during the thirty-six month interview. To ensure consistency the *PFL* evaluation team 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 approximately half of the *PFL* sample. It was not always possible for the *PFL* evaluation team 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

Five of the eight measures in the child health in last 12 months category were in the hypothesised direction, and two of the measures were statistically significant in the hypothesised direction. 14% of children in the high treatment group were reported to have had an accident compared with 22% of the low treatment group ($p<.10$, $d=.23$). 5% of the high treatment group were reported to have stayed in hospital for at least one day compared with 12% of the low treatment group ($p<.10$, $d=.23$). The step-down test showed that the joint effect of the eight measures in this category was not statistically significant.

LONG TERM CHILD HEALTH

One of the two measures in the long term child health category was in the hypothesised direction, and this measure was statistically significant. 15% of children in the high treatment group were reported to have a chronic illness compared with 24% of the low treatment group ($p<.10$, $d=.22$). Overall, the step-down test showed that the joint effect of both measures in this category was statistically significant. The joint effect finding was driven by the significant result found for the 'child has chronic illness' measure ($p<.10$).

MEETING DIETARY GUIDELINES

All five measures in the meeting dietary guidelines category were in the hypothesised direction, one of which was statistically significant. 32% of the high treatment group were found to be meeting the dietary guidelines for protein compared with 19% of the low treatment group ($p<.05$, $d=.32$). The step-down test showed that the joint effect of the five measures in this category was not statistically significant.

NON STEP-DOWN MEASURES

Of the six non step-down measures, five were in the hypothesised direction and one measure was significant. 15% of children in the high treatment group were reported to be meeting the recommended dietary guidelines, compared with 8% of children in the low treatment group ($p<.10$, $d=.22$).

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 an accident	150	(74/76)	0.14	(0.34)	0.22	(0.42)	*p<.10*	*ns*	0.23
* Stayed in hospital for at least one day	150	(74/76)	0.05	(0.23)	0.12	(0.33)	*p<.10*	*ns*	0.23
* Had chest infection	150	(74/76)	0.28	(0.45)	0.36	(0.48)	*ns*	*ns*	0.15
* No. of health problems taken to GP/health centre/casualty	150	(74/76)	1.36	(1.17)	1.49	(1.18)	*ns*	ns	0.10
* Had asthma	150	(74/76)	0.16	(0.37)	0.18	(0.39)	ns	ns	0.06
Child has good health	150	(74/76)	0.88	(0.33)	0.87	(0.34)	ns	ns	0.03
* Had skin problems	150	(74/76)	0.11	(0.31)	0.11	(0.31)	ns	ns	0.01
* Had an ear infection	150	(74/76)	0.30	(0.46)	0.24	(0.43)	ns	ns	0.14
Long Term Child Health									
* Has a chronic illness	150	(74/76)	0.15	(0.36)	0.24	(0.43)	*p<.10*	p<.10	0.22
* Has a physical disability	150	(74/76)	0.01	(0.12)	0.00	(0.00)	*ns*	ns	0.17
Meeting Dietary Guidelines									
Protein	149	(74/75)	0.32	(0.47)	0.19	(0.39)	*p<.05*	ns	0.32
Vegetables	149	(74/75)	0.32	(0.47)	0.24	(0.43)	ns	ns	0.19
Fruits	149	(74/75)	0.61	(0.49)	0.56	(0.50)	ns	ns	0.10
Dairy	149	(74/75)	0.66	(0.48)	0.63	(0.49)	ns	ns	0.07
Grains	149	(74/75)	0.61	(0.49)	0.57	(0.50)	ns	ns	0.07
Non Step-down Measures									
Meeting Dietary Guidelines	149	(74/75)	0.15	(0.36)	0.08	(0.27)	p<.10	-	0.22
Diet Quality Score	149	(74/75)	39.77	(14.28)	35.88	(11.37)	ns	-	0.14
Child's current height (cm's)	67	(34/33)	96.31	(3.78)	96.01	(3.76)	ns	-	0.08
* Child's current weight (kg's)	68	(35/33)	15.35	(2.06)	15.33	(1.76)	ns	-	0.02
* BMI Score	63	(33/30)	16.64	(1.47)	16.72	(1.47)	ns	-	0.06
* BMI Overweight	63	(33/30)	0.24	(0.44)	0.27	(0.45)	ns	-	0.06

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.

2.5 Parenting

Parenting plays a vital role in child development. Evidence suggests that a number of aspects of parenting, such as the quality of parent-child attachment and the combination of control and warmth, have strong, significant effects on children's later psychological adjustment (Carr, 2006). Many home visiting programmes target parenting as a primary outcome and interventions directly aim to improve parenting through strengthening parenting skills (Kendrick et al., 2000). A variety of aspects of parenting have been discussed in depth in previous reports. This report includes an overview of the aspects of parenting assessed at thirty-six months that were not included in previous reports.

PARENTING BEHAVIOURS/ATTITUDES

The beliefs and practices of parents play a significant role in child development (Miller et al., 1996). Traditionally, research has focused on the conceptualisation of parenting patterns and has identified parenting styles based on parents' relative use of responsiveness and demandingness to parent their children. Parenting styles characterised by a combination of high responsiveness and high control are most often associated with positive child outcomes (e.g., Baumrind, 1991; Hetherington et al., 1999; Taylor et al., 2004), while those associated with low responsiveness and/or high control are commonly associated with negative developmental outcomes (Aunolo & Nurmi, 2005; Petito & Cummins, 2000). Meta-analytic results indicate that specific parenting attitudes and behaviours such as parental rejection, warmth, aversion, and granting of autonomy, all significantly influence child anxiety (McLeod et al., 2007). Furthermore, the promotion of a secure attachment relationship, sensitive, responsive, and stimulating parenting is most likely to promote optimal psychological and behavioural functioning (O'Brien et al., 1989). Thus, enhancing sensitive and responsive parenting among high risk families may promote positive development for children who are at risk of poor developmental outcomes, as well as prevent parental abuse and neglect. To this effect, systematic and meta-analytic reviews of home visiting programme have indicated that while the results are mixed across programmes there is some evidence that home visiting positively impacts parenting behaviours and attitudes (Howard & Brooks-Gunn, 2009; Nievar et al., 2010; Sweet & Applebaum, 2004).

PARENTAL ATTITUDES TOWARDS EDUCATION

While there are many factors that influence academic performance, one of the strongest predictors of student motivation is parental involvement. Specifically, children whose parents take an interest in their school progress are more likely to succeed academically and are less likely to be early school leavers (Cotton & Wikelund, 1989; U.S. Department of Education 1994). Parental investment in education begins in the early years and includes the provision of support, encouragement of learning outcomes, active engagement with the learning process, and behaviours such as enrolment in quality preschools (Watson, Brown et al., 1983). Furthermore, parents who have a more positive attitude towards education and are actively involved in their child's education are more likely to engage in behaviours which promote positive beliefs about school and education. For example, parents who are involved in their child's academic success are more likely to praise school related achievements and are more likely to show interest and spend time asking about school or reviewing homework, thus demonstrating that school and education are worthy of adult interest (Hoover-Dempsey & Sandler, 1995). Parental attitudes to education are not commonly used as an outcome in home visiting interventions, however some schools in the US have demonstrated that the practice of parent-teacher home visits from school staff is an effective way promoting parental involvement in their child's academic life and improve outcomes for children (National Education Association, 2011).

PARENTAL MONITORING OF TELEVISION

Children learn social skills through interactions with others (Vygotsky, 1978) and interactions with parents are associated with long-term developmental and behavioural outcomes (Tamis-LaMonda et al., 2001). Indeed, early childhood is a particularly important time for the development of attention and behavioural self-regulation, however excessive television viewing may impede this process (Pagani et al., 2009). There is substantial evidence highlighting the negative impact of excessive television viewing on a variety of outcomes (Christakis et al., 2004; DuRant et al., 1994; Hancox et al., 2005; Pagani et al., 2009; Zimmerman et al., 2005). Specifically, higher levels of television viewing have been found to be associated with reduced physical activity and greater proportion of body fat (DuRant et al., 1994). In addition, television viewing is linked to deficiencies in skills that are essential for academic success such as attention and psychosocial wellbeing in later childhood (Christakis et al., 2004; Pagani et al., 2009). Excessive television viewing has also been linked to poorer educational achievement and participation in bullying (Hancox et al., 2005; Zimmerman et al., 2005).

The American Academy of Pediatrics recommends that children aged two years and older should not be exposed to more than two hours of television or media entertainment per day. However, the reality is that many children engage in much more than the recommended level of media exposure. An analysis conducted

in the US found that 41% of children aged twenty-four to thirty-five months old interacted with media devices, such as televisions or videogames, for longer than the recommended duration. This study also found that lower maternal education was associated with higher rates of children's exposure to television (Certain & Kahn, 2002). An investigation by the Independent Regulator and Competition Authority for the UK communications industries found that 51% of three to four year olds have an established cut-off time after which they are no longer allowed to watch television (Ofcom, 2013).

Despite the evidence that television has a strong influence on child development, parental monitoring of television or exposure to television is not often included as an outcome measure in home visiting evaluations. An evaluation of the Healthy Steps programme included an outcome relating to imposing a limit on television watching. They found that, at 30 months, parents who received the Healthy Steps intervention, which included post natal home visits, were less likely to allow their children to watch more than one hour of television per day than parents who received usual care. They were also more likely to impose a limit than the parents who received the Healthy Steps + Prepare intervention, which entailed the Healthy Steps intervention plus an additional three prenatal home visits (Johnston et al., 2006).

IMPACT OF HOME VISITING INTERVENTIONS ON PARENTING AT THIRTY-SIX MONTHS

Several randomised control trial evaluations of home visiting programmes have reported positive impacts on parenting at thirty-six months (Dishion et al., 2008; Fergusson, Horwood et al., 2005; Jones Harden et al., 2012; Love et al., 2002). Many of these findings are based on observational measures of parent-child interactions. Two evaluations of Early Head Start found that parents who were engaged in the programme were more likely to demonstrate supportiveness in play based interactions with their children (Jones Harden et al., 2012; Love et al., 2002). Similarly, Dishion et al. (2008) reported that, when under observation, parents in the Family Check-up intervention were more likely than the control group to display positive behaviour support in the form of parental involvement, positive reinforcement, prompting, structuring, engagement, and proactive parenting. However, an evaluation of the Healthy Families America programme reported no significant differences between intervention and control groups using an observational measure of caregiver-child interactions, at thirty-six months (Landsverk et al., 2002).

Non-observational measures of parenting have been used by some studies, however the results are somewhat mixed. An evaluation of the Early Start New Zealand programme using the Child Rearing Practices Report found that the intervention group were more likely to indicate the use of positive parenting practices towards their child (Fergusson, Horwood et al., 2005). In addition, Jones Harden and colleagues (2012) found that participants in the Early Head Start programme were less likely to report parental distress. However, both Duggan, McFarlane, and colleagues (2004) and Landsverk and colleagues (2002) assessed the impact of the Healthy Families America programme on violent discipline practices and found no significant differences at thirty-six months. Additionally, an evaluation of the Oxfordshire home visiting study, a UK based intervention, reported that significant improvements in maternal sensitivity and infant cooperativeness in the intervention group, which were identified at twelve months, were not maintained at thirty-six months (Barlow et al., 2008). In summary the results regarding the impact of home visiting programmes on parenting at thirty-six months are mixed. However the evidence suggests that some home visiting programmes can have a favourable impact on parenting beliefs and practices.

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 is proposed to give 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 main global measures, a Frequency scale (α=0.87) score which indicates the frequency of typical hassle events and an Intensity scale (α=0.92) score which reflects the parent's subjective appraisal of how much of a hassle she finds the event to be. Two further subscales are also calculated; parenting hassles related to parenting tasks (α=0.82), that is hassles related to typical tasks or duties a parent may be exposed to, and parenting hassles (α=0.83) related to challenging behaviour by a child.

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. These parenting typologies are based on a 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=0.74), Authoritative (12 items; α=0.85), and Permissive (5 items; α=0.67). The parenting style with the highest mean determines the respondents' parenting style. In addition, the PSDQ yields three authoritative parenting subdomains: connection (α=0.77), regulation (α=0.73), and autonomy (α=0.71); as well as three authoritarian parenting subdomains: punitive (α=0.51), hostility (α=0.46), coercion (α=0.66).

PARENTAL ACCEPTANCE & REJECTION QUESTIONNAIRE – SHORT FORM

The Parental Acceptance and Rejection Questionnaire – Short Form (PARQ; Rohner, 1991) is a 24-item measure assessing the way in which parents act toward their child. Answers were given on a four point scale ranging from almost always true of me to almost never true of me. Item 13 was reverse scored and item scores were summed to yield subscale scores for warmth/affection (8 items; α=0.69), hostility/aggression (6 items; α=0.23), indifference/neglect (6 items; α=0.32), and undifferentiated rejection (4 items; α=0.26). Subscale scores were then summed to an aggregate total score (α=0.59).

ACTIVITIES/INTERACTIONS WITH CHILD

Mothers were asked 16 questions (α=.62) relating to how often they did certain activities (e.g., singing songs, dancing, telling stories) with their child. These items were taken from the My Baby and Me program and Parenting for the First Time program (Centers for the Prevention of Child Neglect, 2000). Answers were given on a 6-point scale ranging from 0 representing not at all to 5 signifying more than once a day. A scale representing the frequency of the mother's interaction with her child was created by taking an average of all responses, with higher scores indicating more interaction.

PARENTAL ATTITUDES TOWARDS EDUCATION/PRIMARY SCHOOL PLANNING

Participants' attitudes to education were assessed by asking them how beneficial they felt it was to have a good education. Specifically, participants were asked to rate their level of agreement, on a 5-point scale from strongly agree to strongly disagree, as to whether a good education would be beneficial for their child, and whether a good education had been beneficial for themselves. Two binary variables were created. Participants were also asked to estimate the age at which they thought their child would leave full-time education and to select from a list of 11 options, ranging from no formal qualification to postgraduate

qualification, the highest level of education that they would like their child to achieve. A binary variable was created from this information to represent whether or not the mother had reported that she would like her child to achieve a third level degree. Additionally, participants reported whether they had their child on a waiting list for primary school, and if so how long they had been on the list.

PARENTAL MONITORING OF TV

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 watched. From the mother's report of the shows watched by the child a binary variable was created which indicated whether or not the child was exposed to content which was age-inappropriate. Additionally, mothers were asked to report their child's favourite show, whether they limited 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

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

PARENTING STYLES AND DIMENSIONS QUESTIONNAIRE (PSDQ)

Two of the three PSDQ subscales were in the hypothesised direction, and both indicated statistically significant differences. Mothers in the high treatment group scored an average of 1.48 on the Authoritarian Parenting subscale, while the low treatment group scored an average of 1.62 ($p<.05$, $d=.32$). In addition mothers in the high treatment group scored an average of 2.15 on the Permissive Parenting subscale, while mothers in the low treatment group scored an average of 2.40 ($p<.05$, $d=.32$). This indicates that mothers in the high treatment group were less likely to engage in behaviours associated with authoritarian or permissive parenting. Furthermore, the step-down test showed that the joint effect of the three PSDQ subscales was statistically significant ($p<.10$), driven by the significant results for Authoritarian and Permissive Parenting.

PSDQ AUTHORITATIVE PARENTING SUBDOMAINS

One of the three PSDQ Authoritative Parenting subdomains was in the hypothesised direction, however there were no significant differences between the high and low treatment group. In addition, the step-down test showed that the joint effect of the three PSDQ Authoritative subdomains was not statistically significant.

PSDQ AUTHORITARIAN PARENTING SUBDOMAINS

All three of the PSDQ Authoritarian Parenting subdomains were in the hypothesised direction, and two indicated statistically significant differences. Mothers in the high treatment group scored an average of 1.59 on the PSDQ Punitive subdomain, whereas mothers in the low treatment group scored an average of 1.82 ($p<.05$, $d=.36$). Mothers in the high treatment group scored an average of 1.53 on the PSDQ hostility subdomain, while mothers in the low treatment group scored an average of 1.71 ($p<.05$, $d=.31$). This indicates that mothers in the high treatment group were less likely to engage in punitive or hostile parenting tactics. Furthermore, the step-down test showed that the joint effect of the three PSDQ Parenting subdomains was statistically significant ($p<.05$), driven by the significant results for the PSDQ Punitive and Hostility subdomains.

PARENTAL ACCEPTANCE AND REJECTION QUESTIONNAIRE (PARQ)

None of the three PARQ subscales were in the hypothesised direction and one indicated a significant difference between the high and low treatment group in the non-hypothesised direction. Mothers in the high treatment group scored an average of 15.52 on the PARQ Indifference/Neglect subscale, whereas mothers in the low treatment scored an average of 14.59 ($p<.05$, $d=.38$). This indicates that mothers in the high treatment group were more likely to show indifference or neglect towards their children. In addition, the step-down test showed that the joint effect of the three PARQ subscales was statistically significant in the non-hypothesised direction ($p<.10$), driven by the significant result for the PARQ Indifference/Neglect subscales.

MATERNAL ATTITUDES TOWARDS EDUCATION

Four of the five Maternal Attitudes Towards Education measures were in the hypothesised direction, and one indicated a statistically significant difference. 66% of mothers in the high treatment group had a positive attitude towards the education that they received, compared with 51% of the low treatment group ($p<.05$, $d=.31$). The step-down test showed that the joint effect of the five measures in the Maternal Attitudes Towards Education category was not statistically significant.

TV HABITS

Six of the eight TV habits measures were in the hypothesised direction; four of which indicated a statistically significant difference. Children in the high treatment group spent significantly less time watching TV, videos, or DVDs than children in the low treatment group, 2.28 hours compared with 2.73 hours ($p<.05$, $d=.29$). In addition, children in the high treatment group spent an average of 0.79 hours watching TV alone, whereas children in the low treatment group spent an average of 1.42 hours watching TV alone ($p<.01$, $d=.54$). The maximum TV time allowed per day for children in the high treatment group, 1.90 hours, was also significantly lower than that for the low treatment group, 2.55 hours ($p<.01$, $d=.60$). Furthermore, households in the high treatment group reported that the TV was on for significantly less time during the day even if no one is watching it. Specifically, those in the high treatment group reported that the TV was on for 7.97 hours per day, while the low treatment group reported that the TV was on for 9.40 hours per day ($p<.05$, $d=.31$). Finally, the step-down test showed that the joint effect of the eight TV Habits measures was statistically significant ($p<.01$), driven by the significant results for Time spent watching TV, videos, or DVDs, Time spent watching TV alone, Maximum TV time allowed per day, and Time TV is on in the home.

NON STEP-DOWN MEASURES

One of the four non step-down measures was in the hypothesised direction, with one significant difference in the non-hypothesised direction: mothers in the high treatment group scored an average PARQ Total score of 139.58, while mothers in the low treatment group scored an average of 138.76 ($p<.10$, $d=.25$).

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)										
	Parental Self-Efficacy	165	(81/84)	8.74	(0.96)	8.42	(1.23)	$p<.05$	ns	0.29
	Baby Comparison Score	165	(81/84)	7.04	(2.10)	6.41	(2.27)	$p<.05$	ns	0.29
	Parental Impact	165	(81/84)	7.97	(2.23)	7.50	(2.55)	ns	ns	0.20
	Parental Warmth	165	(81/84)	8.69	(1.38)	8.48	(1.55)	ns	ns	0.15
Parenting Styles and Dimensions Questionnaire (PSDQ)										
*	Authoritarian Parenting	150	(74/76)	1.48	(0.37)	1.62	(0.48)	$p<.05$	$p<.10$	0.32
*	Permissive Parenting	150	(74/76)	2.15	(0.78)	2.40	(0.78)	$p<.05$	$p<.10$	0.32
	Authoritative Parenting	150	(74/76)	4.12	(0.62)	4.12	(0.62)	ns	ns	0.00
PSDQ Authoritative Parenting Subdomains										
	PSDQ Connection	150	(74/76)	4.74	(0.51)	4.65	(0.56)	ns	ns	0.18
	PSDQ Regulation	150	(74/76)	3.89	(0.88)	3.87	(0.84)	ns	ns	0.02
	PSDQ Autonomy	150	(74/76)	3.73	(0.87)	3.85	(0.91)	ns	ns	0.14
PSDQ Authoritarian Parenting Subdomains										
*	PSDQ Punitive	150	(74/76)	1.59	(0.64)	1.82	(0.63)	*p<.05*	$p<.05$	0.36
*	PSDQ Hostility	150	(74/76)	1.53	(0.51)	1.71	(0.64)	$p<.05$	$p<.10$	0.31
*	PSDQ Coercion	150	(74/76)	1.31	(0.45)	1.33	(0.48)	ns	ns	0.03
Parental Acceptance & Rejection Questionnaire (PARQ)										
	PARQ Warmth/Affection	150	(74/76)	9.03	(1.61)	9.03	(2.32)	ns	ns3	0.00
*	PARQ Undifferentiated Rejection	150	(74/76)	12.91	(1.01)	12.86	(1.23)	ns	ns	0.04
*	PARQ Hostility/Aggression	150	(74/76)	20.58	(1.21)	20.34	(1.16)	ns	ns	0.20
*	PARQ Indifference/Neglect	150	(74/76)	15.52	(1.25)	14.59	(1.52)	s~	ns	0.38
Maternal Attitudes Towards Education										
	Believes education helped her in life	150	(74/76)	0.66	(0.48)	0.51	(0.50)	$p<.05$	ns	0.31
	Believes a good education will help her child get ahead	150	(74/76)	0.89	(0.31)	0.84	(0.37)	ns	ns	0.15
	Age thinks child will leave education	144	(72/72)	19.69	(2.16)	19.39	(2.17)	ns	ns	0.14
	Would like child to have a third level degree	150	(74/76)	0.68	(0.47)	0.62	(0.49)	ns	ns	0.12
	Child is on primary school waiting list	148	(73/75)	0.36	(0.48)	0.41	(0.50)	ns	ns	0.12
TV Habits										
*	Time spent by child watching TV alone	140	(71/69)	0.79	(0.91)	1.42	(1.42)	$p<.01$	$p<.01$	0.54
*	Maximum TV time allowed per day	83	(42/41)	1.90	(1.03)	2.55	(1.16)	$p<.01$	$p<.05$	0.60
*	Time TV is on in the home	143	(73/70)	7.97	(4.31)	9.40	(4.87)	$p<.05$	ns	0.31
*	Time spent by child watching TV/videos/ DVDs per day	150	(74/76)	2.28	(1.45)	2.73	(1.73)	$p<.05$	ns	0.29
	Child's TV time is limited	145	(73/72)	0.59	(0.50)	0.58	(0.50)	ns	ns	0.01
	Mother talks to child about TV	120	(59/61)	0.98	(0.13)	0.98	(0.13)	ns	ns	0.00
*	Child watches age-inappropriate content	141	(71/70)	0.61	(0.49)	0.60	(0.49)	ns	ns	0.01
	Time mother spends watching TV with child	117	(57/60)	1.25	(1.06)	1.28	(0.91)	ns	ns	0.03
Non Step-down Measures										
	Interaction with child	151	(75/76)	2.71	(0.45)	2.69	(0.45)	ns	-	0.04
	Worried about child's language development	151	(75/76)	0.16	(0.37)	0.17	(0.38)	ns	-	0.03
	Worried about child's behaviour	151	(75/76)	0.07	(0.25)	0.09	(0.29)	ns	-	0.09
*	PARQ Total Score	150	(74/76)	139.58	(3.22)	138.76	(3.38)	s~	-	0.25

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.

2.6 Home and Family Environment

During the early years the quality of the home and family environment, incorporating the experiences and exposures it allows, is recognised as being critical in fostering brain development (DiPietro, 2000; Halfon et al., 2001; Maggi et al., 2010). Between the ages of two and three children are developing rapidly, especially in terms of language, and thus the home and family environment is particularly important during this period (Tamis-LeMonda et al., 2006; WHO, 2007). One study investigating the impact of the home environment on child development revealed stronger associations between learning experiences and language skills for younger as compared to older children (Bradley et al., 2001a, 2001b).

The home and family environment encompasses a range of social, cognitive, and physical factors, which can impact on child development and family functioning (Illig, 1998). Characteristics of a child's social environment include parental responsiveness, warmth, nurturance, sensitivity, in addition to parent-facilitated promotion of autonomy and restriction (Bradley & Caldwell, 1984; Edwards et al., 2010). Many of these factors have been associated with cognitive and social outcomes for young children and their later academic achievements (Bradley et al., 1989; Clark & Ladd, 2000; Edwards et al., 2010; Farah et al., 2008; Steinbert et al., 1992). The family environment also includes the sense of cohesion within the family, conflict, and communication patterns (Moos & Moos, 2009). Studies have found that children living in families with low cohesion or high conflict are more likely to have a difficult temperament or conduct problems (Lopez & Thurman, 1993; Stadelmann et al., 2007).

Cognitive factors of the home environment which can facilitate child development include access to a variety of stimulating play materials and activities and the provision of opportunities for perceptual and cognitive activities (Bradley & Caldwell, 1976; Farah et al., 2008). A stimulating and rich environment can predict children's language, cognitive development, and academic achievement, especially combined with parental participation (Arnold et al., 1994; Fagan & Iglesias, 1999; Farah et al., 2008; Sheridan & Knoche, 2008). Conversely, exposure to low levels of environmental stimulation is associated with cognitive, social, and behavioural delays in children (Bradley & Caldwell, 1984; WHO, 2007). The literacy environment, including provision of age-appropriate materials, across the first three years of a child's life has been linked with children's cognitive and language skills at thirty-six months (Rodriguez et al., 2009). Environmental stimulation is also associated with intrinsic motivation, later language development, mathematics, and reading ability (Farah et al., 2008; Gottfried et al., 1998).

Physical factors including safety, exposure to toxins, deprivation, overcrowding, chaos, and disorganisation, have all been shown to impact on aspects of child development (Chen et al., 2002; Evans, 2006; Moos & Moos, 2009). Second hand smoke exposure during early childhood is related to a variety of health difficulties both during infancy and later in life (Been et al., 2013; Gergen et al., 1998; Mannino et al., 2003). A child-safe physical environment can be protective and allows opportunities for children to explore, developing independence and autonomy (Carr, 2006; Edwards et al., 2010). Overcrowding, chaos, and confusion within the home are associated with psychological distress and behavioural adjustment problems at school (Ackerman et al., 1999; Adam, 2004; Petrill et al., 2004; Evans et al., 2010; Vernon-Feagans et al., 2012). Household disorganisation over children's first three years of life is also a significant predictor of both receptive and expressive language at thirty-six months, independent of factors such as poverty, maternal literacy, and depression (Vernon-Feagans et al., 2012).

When issues arise within a child's home and family environment, such as social and emotional problems, or issues of domestic violence or child abuse, a family may be put in contact with a social worker (http://hse.ie/eng/services/list/4/ChildrenandFamilyServices/Roles_.html#Social). While many home visiting evaluations do not include a measure of social worker involvement, the Healthy Families America programme investigated substantiated child protective services reports of abuse or neglect or threatened abuse, neglect, or harm (Duggan, McFarlane et al., 2004). They found no differences between the intervention and control groups. Four other home visiting programmes reported favourable effects in relation to child maltreatment; a common reason for a social worker to become involved with a family. The Child FIRST home visiting programme was found to have a favourable effect on family involvement with child protective services, measured by mother's reported involvement with services and state records

(Lowell et al., 2011). Roggman and Cook (2010) found a favourable effect for the Early Head Start home visiting programme relating to the reduction in the use of physical punishment. An evaluation of Early Start by Fergusson, Horwood et al. (2005) found that the programme had a favourable effect on the percentage of children who attended hospital for accidents, injuries, or poisoning and physical assault by any parent as measured by the Conflict Tactics Scale (Straus et al., 1996). The Healthy Families America programme was found to have a favourable effect on the frequency of psychological aggression and mild physical assault as measured by the Parent-Child Conflict Tactics Scale (Straus et al., 1998).

At thirty-six months, evidence of the impact of home visiting programmes on other factors of the home and family environment is mixed. In the Family Check-up programme, Shaw et al. (2006) reported that mothers receiving treatment showed increases in involvement in child behaviour from ages two to four compared with reductions for those in the control condition. While this study used a subscale of the HOME to measure parent involvement, they used an adapted version of the scale using only observable items and this should be considered when comparing results. Olds et al. (1994) reported that at thirty-six months, children involved in the Nurse Family Partnership programme lived in homes with fewer hazards for children. While they also used the full HOME inventory scale they found no other significant treatment effects. The Early Head Start home visiting programme reported that the home environment of families receiving the programme did not differ in comparison to the control group (Love et al., 2002). However, they report that parents involved in the combined home visiting and centre-based programme provided significantly more support for language and learning than control group parents. They also reported a favourable effect in terms of the HOME total score for these families (Love et al., 2002). An overview of the Early Head Start programme across 17 sites found no significant differences in the home learning environment between participants involved in the programme and those in the control group (Jones Harden et al., 2012). Fergusson, Horwood et al. (2005) found no effects in the percentage of children living in a smoke free home at thirty-six months. The Healthy Families America programme evaluation found no programme impacts on the intervention children's home environment (Landsverk et al. 2002; Duggan, McFarlane et al., 2004). Finally, Wagner et al. (1999) found that the Parents as Teachers programme had an unfavourable effect on the home environments, reporting a negative effect on the HOME subscale that measures acceptance of children's behaviour.

2.6.1 Home Environment Instruments

QUALITY OF THE HOME ENVIRONMENT

The Infant-Toddler version of the Home Observation for Measurement of the Environment (HOME; Caldwell & Bradley, 2003) is a 45-item instrument completed by a trained interviewer. It measures the stimulation potential of the child's home environment and may be used as a substitute for reliance on social class as an indicator of quality of the child's home environment. The HOME Inventory comprises six domains. Responsivity (11 items, α=0.51) illustrates the degree to which a parent is responsive to the child's behaviour. Acceptance (8 items, α=0.49) represents parental acceptance of negative behaviour from the child and avoidance of unnecessary punishment. Organisation (6 items, α=0.12) pertains to the degree of routine in a family's schedule, safety of the environment, and community supports utilised. The Learning Materials domain (9 items, α=0.22) assesses the appropriateness of play materials for the child. Involvement (6 items, α=0.58) illustrates the degree to which the parent is involved in the child's learning and promotes child development. Finally, the Variety domain (5 items, α=0.36) assesses visitation of people and attendance of activities that introduce variety into the child's life. Each item was scored by a trained interviewer as true or not. Items were scored based on observations while in the home (18), direct questioning in a semi-structured interview manner (15), or either of the two (12). For items where this was not possible, the mother was directly asked the question in an interview format. If the item was true it was scored as a 1, if it was not true it was scored as 0. Scores for each domain of the HOME Inventory were obtained by averaging the responses to each question in that domain, with higher scores indicating a more nurturing home environment. In addition, an overall total HOME score (α=0.66) was obtained by calculating the average of all responses. Some of the HOME items cannot be completed if the interview does not take place in the participant's home or the participant's child is not present. For this reason some participants will not have a score for all HOME domains.

The Supplement to the HOME for Impoverished Families (SHIF; Ertem, Avni-Singer, & Forsyth, 1996) consists of 20 (α=0.53) binary choice items which are summed to form a total score. The SHIF was designed by clinicians with extensive experience working with impoverished families in urban areas, to address limitations in the HOME scale when used with this population. Specifically, the SHIF provides additional information relating to whether the basic needs of the child are adequately being met, such as whether the child has a safe and consistent place to sleep and the organisation of the child's environment and daily routine, such as whether a child eats one meal per day with mother or whether the child is not left to self-feed. The measure was developed to be used in conjunction with the HOME and to be administered by a trained interviewer.

FAMILY ENVIRONMENT SCALE

The Family Environment Scale (FES; Moos & Moos, 2009) is a 90-item inventory which measures perceived real, ideal, and expected family environments, across three dimensions; Relationship, Personal Growth and System Maintenance. In the *PFL* evaluation, only two of the dimensions were measured using 45 items - the Relationship dimension includes cohesion (9 items, α=0.64), expressiveness (9 items, α=0.46), and conflict (9 items, α=0.65), and the System Maintenance dimension includes organisation (9 items, α=0.20) and control (9 items, α=0.75). Mothers were asked to answer each of the 45 statements by choosing whether it was true or false for their family. The responses were matched to the score card (each item has a preferred direction, either true or false), and the scores were converted into a standard score using the FES conversion table.

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/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/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.

HOME OBSERVATION FOR MEASUREMENT OF THE ENVIRONMENT

Of the six subscales on the Home Observation for Measurement of the Environment (HOME) measure, three were in the hypothesised direction and were statistically significant. The high treatment group scored 5.35 on average on the organisation subscale while the low treatment group scored 5.05 on average (p<.01, d=.40). This suggests that children in the high treatment group have more regularity and predictability in their families schedule, their physical environment and their families' use of community services. The high treatment group scored significantly higher (4.20) than the low treatment group (3.73) on the Involvement subscale (p<.05, d=.31) which measures the extent to which the parent is actively involved in the child's learning and development. Thirdly, the high treatment group scored significantly higher (6.31) than the low treatment group (5.89) on the Acceptance subscale (p<.05, d=.38), which is designed to assess parental acceptance of negative behaviours and situations. The supplement measure to the HOME designed for impoverished families was not found to be statistically significant.

FAMILY ENVIRONMENT SCALE (FES) - STANDARDISED

Two of the five measures in the FES standardised category were in the hypothesised direction, with the Organisation subscale being statistically significant ($p<.10$, $d=.23$). The high treatment group scored 50.36, compared with the low treatment group average score of 48.43. This indicates that the high treatment group reported more emphasis on clear organisation and structure in planning family activities and responsibilities.

NON STEP-DOWN MEASURES

Two of the three non step-down measures were in the hypothesised direction. The total HOME score, based on all six subscales, was in the hypothesised direction and statistically significant with the high treatment group scoring on average 38.37 and the low treatment group scoring on average 36.98 ($p<.05$, $d=.41$). In addition, the proportion of children in the high treatment group (36%) exposed to cigarette smoke in the home was significantly less than the number of children in the low treatment group (53%; $p<.05$, $d=.34$).

Table 2.4 Results for High and Low Treatment Groups: Home 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
	Home Observation for Measurement of the Environment (HOME)									
	Organisation	149	(73/76)	5.35	(0.60)	5.05	(0.90)	*p<.01*	p<.05	0.40
	Involvement	148	(72/76)	4.20	(1.51)	3.73	(1.48)	*p<.05*	ns	0.31
	Acceptance	92	(43/49)	6.31	(1.04)	5.89	(1.19)	p<.05	ns	0.38
	Supplement to the HOME for Impoverished Families (SHIF)	148	(72/76)	17.73	(1.60)	17.56	(1.50)	ns	ns	0.26
	Variety	149	(73/76)	4.07	(1.02)	4.11	(0.95)	ns	ns	0.04
	Learning Materials	148	(72/76)	8.42	(0.71)	8.49	(0.56)	ns	ns	0.10
	Responsivity	144	(72/72)	9.61	(1.37)	9.76	(1.12)	ns	ns	0.12
	Family Environment Scale Standardised									
	Organisation Standard Score	150	(74/76)	50.36	(8.71)	48.43	(7.89)	p<.10	ns	0.23
	Expressiveness Standard Score	150	(74/76)	54.64	(11.66)	53.87	(9.48)	ns	ns	0.07
*	Conflict Standard Score	150	(74/76)	70.57	(9.74)	69.68	(9.21)	ns	ns	0.08
	Control Standard Score	150	(74/76)	52.09	(6.65)	52.99	(8.18)	ns	ns	0.12
	Cohesion Standard Score	150	(74/76)	55.82	(11.31)	57.71	(8.80)	ns	ns	0.19
	Non Step Down Measures									
	Total HOME	92	(43/49)	38.37	(3.64)	36.98	(3.21)	p<.05	-	0.41
*	Child exposed to cigarette smoke	149	(74/75)	0.36	(0.48)	0.53	(0.50)	p<.05	-	0.34
*	Social worker working with family	150	(75/75)	0.04	(0.20)	0.03	(0.16)	ns	-	0.07

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.

2.7 Maternal Health & Wellbeing

General maternal health as well as psychological wellbeing may influence a child's learning, development, and behaviour (Mensah & Kiernan, 2010). Parental physical illness may disrupt parenting practices such as support, reinforcement, and changes in household routines, which may in turn affect child functioning (Armistead et al., 1995). Maternal health and wellbeing may be positively affected by the level of social support available and the quality of the mother's relationship with the child's father (Bögels & Phares, 2008; Lamb, 2000). Physical and psychological health difficulties are often concurrent and self-reported general health by mothers has been shown to reflect both physical and psychological symptoms (Mensah & Kiernan, 2010; Zubaran et al., 2010). Maternal mental health has been consistently reported to have an effect on parenting and child development (Waylen & Stewart-Brown, 2010; Letourneau et al., 2013). Letourneau et al. (2013) suggest that exposure to maternal depression at different developmental periods may negatively affect child outcomes in different ways. The effects of exposure to maternal depression are also likely to be additive (Sohr-Preston & Scaramella, 2006). Maternal depression occurring when a child is two to three years of age may lead to increased risk for anxiety by the time the child is ten to eleven years of age (Letourneau et al., 2013). Mothers with severe or chronic depression are also less able to engage in symbolic play and likely to read less frequently to their toddlers (see Sohr-Preston & Scaramella, 2006 for a review). Poor maternal mental health may also work additively alongside other issues, such as substance abuse and domestic violence, to increase the risk of child behaviour problems (Whitaker et al., 2006).

SUBSTANCE USE

While the negative effects of substance use (i.e. smoking, drinking alcohol, and drug use) during pregnancy have been well documented, the effects of substance use in the home during a child's preschool years has received less attention. The effects of inhaling second hand smoke pose a greater risk to children than adults as they breathe more quickly and have smaller airways which are still developing along with their immune systems (Been et al., 2014; Robinson & Kirkcaldy, 2007). A growing body of literature suggests that exposure to second hand smoke during the preschool years may leave young children at a higher risk of developing a number of illnesses including lower respiratory tract infections, wheezing, and asthma (Burke et al., 2012; Li et al., 1999). The HSE notes higher rates of cigarette smoking among lower socioeconomic groups in Ireland (HSE, 2013b). Drug and alcohol abuse in the home may create an unstable environment and have detrimental effects on a child's development. A recent report from the HSE found that those from lower social classes were more vulnerable to family problems due to others drinking (HSE, 2011).

IMPACT OF HOME VISITING INTERVENTIONS ON MATERNAL HEALTH AND WELLBEING AT THIRTY-SIX MONTHS

Home visiting programmes have had limited effects on maternal health and wellbeing at thirty-six months. Evaluations of a number of large scale home visiting programmes, including the MOM Program, Early Start, Healthy Families America, and Healthy Steps found no favourable effects in relation to maternal health and wellbeing. These studies measured the number of subsequent pregnancies, number of other children in the family, pregnancy history, incidence of depression, smoking, substance use, family/social relationship problems, confidence in relationships, drug use, problem substance use, drinks per day, parental distress, and mental health including depression and anxiety (Schwarz et al., 2012; Fergusson, Horwood et al., 2005; Duggan, Fuddy et al., 2004; Landsverk et al., 2002; Wagner & Clayton, 1999). Two home visiting programmes found effects in the area of maternal mental health and wellbeing at thirty-six months. Love and colleagues (2002) found that Early Head Start had a favourable effect on parental distress as measured by the Parenting Stress Index. Additionally, when investigating the impact of the Family Check-up on maternal depressive symptoms, Shaw and colleagues (2009) found that mothers in the intervention group reported a significantly greater decrease in depressive symptoms using the Center for Epidemiological Studies on Depression Scale (CES-D; Radloff, 1977).

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 if they had used any health services in the past 12 months from a list provided by the researcher. In addition, participants were asked how many times they visited the GP in the last 12 months (not including visits for their child). Participants were asked 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.

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. The four responses to each question were rated 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 being at high risk for depression.

WHO-5 INDEX

The WHO-5 Index is a 5-item measure (α=0.91) 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 on this measure, with scores below 13 indicating poor wellbeing.

FUTURE OUTLOOK

Future outlook was assessed using the Future Outlook Inventory (FOI; Cauffman & Woolard, 1999), an 8-item measure (α=0.74) of an individual's consideration of future events and consequences. The FOI measures future orientation, in particular an individual's tendency to recognise and consider future events or consequences. Mothers rated how true each item was of themselves on a 4-point scale ranging from never true to always true. The scores on each item were summed and divided by eight to give a mean score, with higher scores indicating a higher degree of future consideration and planning.

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 twenty-four and thirty-six months. Positive numbers indicate a reduction in smoking, whereas negative numbers indicate that the participant is smoking more. The average change is reported.

2.7.2 Maternal Health & Wellbeing Results

Table 2.5 presents the results comparing the high and low treatment groups on the maternal health and wellbeing domain.

MATERNAL PHYSICAL HEALTH AND HEALTH BEHAVIOURS IN PAST 12 MONTHS

Of the two measures in the Maternal Physical Health and Health Behaviours in the past 12 months category, both were in the non-hypothesised direction and failed to reach statistical significance. The step-down test showed that the joint effect of the two measures was not statistically significant.

MATERNAL MENTAL HEALTH

All three of the measures in the Maternal Mental Health category were in the hypothesised direction and two were statistically significant. Lower scores on the EPDS were reported by mothers in the high treatment (6.35) group compared to the low treatment (8.01) group ($p<.05$, $d=.30$). In addition, mothers in the high treatment (61.08) group had higher scores than those in the low treatment (55.32) group on the WHO-5 ($p<.10$, $d=.24$). Overall, the step-down test showed that the joint effect of these three measures was statistically significant ($p<.10$). This effect was driven by the significant result found on the EPDS.

SUBSTANCE USE

Three of the four measures in the Substance Use category were in the hypothesised direction and one showed a statistically significant difference between the groups. 84% of the high treatment group reported having drank alcohol in the past 12 months compared with 92% of the low treatment group ($p<.10$, $d=.26$). The step-down test showed that the joint effect of the four measures in this category was not statistically significant.

NON STEP-DOWN MEASURES

All eight of the non step-down measures were in the hypothesised direction. One of these measures revealed a statistically significant difference such that the high treatment group reported a reduction in the number of cigarettes smoked between twenty-four and thirty-six months ($p<.10$, $d=.32$). This equates to the high treatment group smoking approximately one cigarette less per day in this time period.

Table 2.5 - Results for High and Low Treatment Groups: Maternal Health and 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										
*	No. of GP visits	148	(74/74)	3.89	(4.80)	3.73	(5.61)	ns	ns	0.03
	Good health compared with other women	150	(74/76)	0.82	(0.38)	0.87	(0.34)	ns	ns	0.12
Maternal Mental Health										
*	Edinburgh Postnatal Depression Score for past 7 days	151	(75/76)	6.35	(5.86)	8.01	(5.40)	p<.05	p<.10	0.30
	WHO-5 Percentage Score	150	(74/76)	61.08	(24.56)	55.32	(24.36)	p<.10	ns	0.24
	Future Outlook Inventory	150	(74/76)	2.67	(0.60)	2.62	(0.51)	ns	ns	0.10
Substance Use										
*	Drank alcohol in past 12 months	150	(74/76)	0.84	(0.37)	0.92	(0.27)	p<.10	ns	0.26
*	Drug use in past 12 months	150	(74/76)	0.03	(0.16)	0.05	(0.22)	ns	ns	0.13
*	More than 14 units of alcohol consumed per week	150	(74/76)	0.18	(0.38)	0.21	(0.41)	ns	ns	0.09
*	Currently a smoker	150	(74/76)	0.53	(0.50)	0.50	(0.50)	ns	ns	0.05
Non Step-down Measures										
*	Below WHO-5 Score of 13	150	(74/76)	0.28	(0.45)	0.38	(0.49)	ns	-	0.21
*	Edinburgh Postnatal Depression Cut-off	151	(75/76)	0.27	(0.45)	0.34	(0.48)	ns	-	0.16
	Been pregnant since birth of *PFL* child	150	(74/76)	0.27	(0.45)	0.18	(0.39)	ns	-	0.21
	Currently using a valid form of birth control	148	(72/76)	0.69	(0.46)	0.66	(0.48)	ns	-	0.08
	New pregnancy planned	12	(7/5)	0.43	(0.53)	0.40	(0.55)	ns	-	0.06
*	Binge drinking (> 6 units in any sitting at least once per week)	150	(74/76)	0.20	(0.40)	0.24	(0.43)	ns	-	0.08
	Reduction in smoking between 24 and 36 months (number of cigarettes)	71	(36/35)	0.92	(3.99)	-0.31	(3.81)	p<.10	-	0.32
*	Number of cigarettes per day	78	(40/38)	11.85	(6.19)	11.87	(5.57)	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. *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. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.8 Maternal Social Support

Social support may serve as a protective factor for both maternal and child wellbeing (Levitt, Webber, & Cherie, 1986) as it can reduce the impact of negative situations and events (Cobb, 1976). It can be emotional, informational, and instrumental (House & Kahn, 1985, cited in Lutz et al., 2012) and can come from a number of sources, most notably emotional support and childcare assistance from the mother's spouse or partner and the maternal grandmother (Levitt et al., 1986). Social support may also be provided by the other members of the family, close friends, and the wider community. Social support may be a particularly important protective factor for individuals residing in disadvantaged communities, many of whom are at greater risk of poor physical and mental health (e.g., Bradley & Corwyn, 2002).

Emotional, behavioural or financial support offered to the child's mother may also have an impact on the child. By functioning as a source of emotional support for the mother, fathers can enhance the quality of the mother-child relationship, and facilitate positive adjustment for the child (Lamb, 2000; Bögels & Phares, 2008). The presence of a non-resident father has been reported to reduce parenting stress and depression in mothers of three year olds from disadvantaged communities, which in turn reduced child behaviour problems at age five (Jackson et al., 2013). Fathers' presence was assessed using mothers' ratings of how satisfied they were with the child's father in relation to love, time spent with the child, and money provided for the child. In addition, the presence of a child's father in infancy has been associated with less internalising problems in three year olds (Flouri & Malmberg, 2012). It should be noted that the relationship between father's presence and child development is complicated as father's presence does not guarantee quality interactions with the child.

The family is the first social system that a child is a part of and involves networks of reciprocal relationships and alliances (Bronfenbrenner, 1979). The social system of the family consists of a network of connected relationships, each of which affects and is affected by one another. As such, a mother's relationship with her partner may affect a child's development. A good relationship between a child's mother and her partner may serve as a protective factor in this domain. Where a child is experiencing difficulties, these problems may be maintained due to unresolved or unacknowledged marital (relationship) discord (Carr, 2006).

VOTING BEHAVIOUR

As in previous *PFL* evaluation reports, voting behaviour was assessed at the thirty-six month time point as a measure of social participation. In Ireland, electoral turnout is reportedly associated with socioeconomic class, with greater turnout displayed by higher SES groups (Kavanagh, 2005). To our knowledge, there are no previous reports of the impact of home visiting programmes on voting behaviour.

IMPACT OF HOME VISITING INTERVENTIONS ON SOCIAL SUPPORT AT THIRTY-SIX MONTHS

Although promoting social support is not the primary outcome of many home visiting interventions, it is often cited as a secondary or mediating outcome (Hodnett & Roberts, 2007; Kearney et al., 2000). A number of home visiting evaluations report positive effects on maternal social support (Barlow et al., 2003; De la Rosa et al., 2005; McCurdy, 2001), however, few report outcomes at thirty-six months. Schwarz et al. (2012) report an increase in referral to, and receipt of, early intervention services in mothers who participated in the MOM programme at thirty-three months. The MOM program also had a favourable effect on the time to referral for early intervention and the time to receipt of early intervention services. This programme aims to increase participation in child primary health care services and to promote participation in early intervention programmes, with early intervention services representing social support.

2.8.1 Maternal Social Support Instruments

FATHER 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.

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 do not live with them. A binary variable was created, indicating whether mothers met with them most days or less frequently.

MATERNAL SOCIAL SUPPORT INDEX

Maternal social support was assessed using an adapted version of the Maternal Social Support Index (MSSI; Pascoe et al., 1988). The adapted version consists of nine items related to maternal perception of help with daily tasks (α=0.79). The participant responded to each item by indicating who does different household tasks in her household, with the response options: I generally do it, someone else and I generally do it, or someone else generally does it. Items were summed to obtain a total score, with higher scores representing more perceived support by the mother.

RELATIONSHIP QUALITY INDEX

Participants' romantic relationship quality was assessed using the Quality of Marriage Index (QMI; Norton, 1983) renamed as the Relationship Quality Index (RQI). The RQI is a 6-item questionnaire (α=0.96) which measures the mother's satisfaction with her relationship with her current partner. For items 1 to 5 mothers rated statements about their relationship with their partner on a 7-point scale from very strongly disagree to very strongly agree. For item 6 participants rate the degree of happiness that best described their relationship with their partner on a scale of 1 to 10 with 1 representing unhappy and 10 indicating very happy. Scores on the RQI were summed to obtain a total score. Scores on this measure range from 6 to 45, with lower scores on this measure indicating a more negative evaluation.

BEECH CENTER FAMILY QUALITY OF LIFE SCALE

The Beech Center Family Quality of Life Scale (FQOL; Hoffman et al., 2006) is a 25-item measure (α=0.94) of family quality of life. The *PFL* evaluation excluded the disability-related support subdomain (4 items), yielding a 21 item measure of family quality of life. The measure yields an aggregate total score and scores for four subscales: family interaction (6 items, α=0.89), parenting (6 items, α=0.87), emotional wellbeing (4 items, α=0.78), and physical/material wellbeing (5 items, α=0.74). Items were rated on a 5-point Likert scale of satisfaction ranging from very dissatisfied to very satisfied. Higher scores indicate higher quality of life.

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. The results of all three measures were not statistically significant. The step-down test showed that the joint effect of the three measures in this category was not statistically significant.

SOCIAL SUPPORT

Four of the six 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 six measures in this category was not statistically significant.

FAMILY QUALITY OF LIFE

Three of the four subdomains of the Beech Family Quality of Life Scale (FQOL) were in the hypothesised direction. None of the subdomains were statistically significant. The step-down test showed that the joint effect of the four measures in this category was not statistically significant.

VOTING BEHAVIOUR

Both of the measures in the Voting Behaviour were in the hypothesised direction; however neither measure revealed statistically significant differences between the groups. The step-down test showed that the joint effect of the measures was not statistically significant.

NON STEP-DOWN MEASURES

Three of the four non step-down measures were in the hypothesised direction and one was statistically significant. 90% of the high treatment group reported that child maintenance payments from the child's father were paid regularly compared to 72% of the low treatment group ($p<.10$, $d=.47$).

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 child's father	146	(71/75)	0.68	(0.47)	0.67	(0.47)	ns	ns	0.02
Father has daily contact with child	149	(73/76)	0.67	(0.47)	0.68	(0.47)	ns	ns	0.03
Support from partner	105	(50/55)	0.84	(0.37)	0.89	(0.31)	ns	ns	0.15
Social Support									
Support from parent	142	(71/71)	0.70	(0.46)	0.61	(0.49)	ns	ns	0.21
Support from relatives	151	(75/76)	0.49	(0.50)	0.42	(0.50)	ns	ns	0.15
Support from friends	151	(75/76)	0.20	(0.40)	0.18	(0.39)	ns	ns	0.04
Maternal Social Support Index	150	(74/76)	25.38	(4.94)	25.18	(5.00)	ns	ns	0.04
Meet friends (most days/less)	150	(74/76)	0.46	(0.50)	0.47	(0.50)	ns	ns	0.03
Support from neighbours	147	(73/74)	0.07	(0.25)	0.09	(0.29)	ns	ns	0.10
Beech Center Family Quality of Life Scale (FQOL)									
Physical/Mental Wellbeing	150	(74/76)	21.54	(2.71)	21.13	(2.56)	ns	ns	0.16
Emotional Wellbeing	150	(74/76)	16.49	(2.59)	16.20	(2.54)	ns	ns	0.11
Parenting	150	(74/76)	26.30	(3.16)	26.22	(2.73)	ns	ns	0.03
Family Interaction	150	(74/76)	26.36	(3.78)	26.46	(2.62)	ns	ns	0.03
Voting Behaviour									
Voted in last General Election	150	(74/76)	0.61	(0.49)	0.53	(0.50)	ns	ns	0.17
Voted in last Local Elections and European Elections	148	(73/75)	0.59	(0.50)	0.51	(0.50)	ns	ns	0.17
Non Step-down Measures									
Child's father pays maintenance	64	(32/32)	0.63	(0.49)	0.56	(0.50)	ns	-	0.13
Child maintenance is paid regularly	38	(20/18)	0.90	(0.31)	0.72	(0.46)	$p<.10$	-	0.47
FQOL Total Score	150	(74/76)	85.64	(11.16)	85.08	(9.08)	ns	-	0.06
Relationship Quality Index	105	(50/55)	39.24	(6.92)	39.36	(4.75)	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. *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. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.9 Childcare and Service Use

CHILDCARE

There is a complex interaction between the type, timing, and quality of childcare on child outcomes (NICHD, 2004; Sylva et al., 2011). While centre-based childcare can be a predictor of healthy cognitive development when of good quality, (Sylva et al., 2011), it has also been associated with higher incidences of externalising behaviours in children (Loeb et al., 2007; NICHD, 2004).

Evidence from the US suggests that children who begin childcare between the ages of two to three years display the greatest academic benefit, while those who commence at a younger age experience negative behavioural effects (Barnett, 1995; Loeb et al., 2007). In addition, the duration of time spent in childcare, specifically the number of hours attending centre-based childcare per week, is reported to be the most consistent predictor of socio-behavioural problems (National Institute of Child Health and Human Development, 2002). However, the impact of time spent in childcare on problem behaviours appears to be stronger at twenty-four months than at thirty-six months, when the effect is greatly reduced (NICHD, 2002). Additionally, child temperament was seen to be exacerbated by lengthy periods of centre-based childcare, such that children who were classified as 'easily frustrated' before starting childcare displayed more externalising behaviours and 'distressed' children more internalising behaviours (Crockenberg, 2005). A study conducted by Barnes and colleagues (2010) examining the experiences of childcare in a UK sample reported no evidence of adverse consequences of childcare in the first three years, and some limited evidence of benefits.

While there is substantial documentation of the advantages and disadvantages of formal childcare on cognitive development and school readiness up to thirty-six months, childcare use is not generally reported as an outcome in evaluations of home visiting programmes. Thus, there is an absence of research reporting the impact of home visiting programmes on childcare uptake.

SERVICE USE

In low socioeconomic communities there are often higher rates of emergency services usage and lower rates of engagement with preventative care services (Hubert, 2010). Engaging with health services, particularly preventative services, can reduce the instances of illness and emergency room visits and hospitalisations (Leventhal et al., 2000). Home visiting programmes can promote children's health by linking families to other services and encouraging the use of preventative health service, such as prenatal care, checkups, and immunisations (Gomby et al., 1999). However, assessing the impact of home visiting programmes on service use is difficult. Service use is often not explored as a distinct subdomain and often subtle indicators of service use, such as hospital admissions, are collapsed into larger domains such as child health. Thus the comparability of findings across programmes is limited. In spite of this at thirty-six months there is some evidence that home visiting increases preventative service use. Specifically, children participating in home visiting interventions are more likely to have a dentist and are more likely to be referred to, and receive, early intervention services than those in the control group (Fergusson, Horwood et al., 2005; Schwarz et al., 2012).

2.9.1 Childcare and Service Use Instruments

CHILDCARE

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 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.

SERVICE USE

Participants in the *PFL* cohort were asked if they had ever used any of the 63 services listed. Services were grouped into the following domains: emergency services, health services, family services, employment/adult education services, community information services, residents' association services, childcare services, and other services. Scores for each domain represent the number of services ever used by participants in that domain. In addition, a variable representing the total number of services mothers indicated using was created. Note that these questions were not asked of the comparison community.

2.9.2 Childcare and Service Use Results

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

CHILDCARE USE

Within the total sample, 75% of participants reported using either formal or informal childcare arrangements by thirty-six months with no reported difference across the high and low treatment groups. Six of the seven measures in the Childcare Use category were in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the six measures in the Childcare Use category was not statistically significant.

SERVICE USE

Seven of the nine measures in the service use category were in the hypothesised direction, however none of these effects indicated statistically significant differences between the high and low treatment groups. In addition, the step-down test showed that the joint effect of the nine measures in the service use category was not statistically significant.

NON STEP-DOWN MEASURES

The measure relating to use of any type of childcare was in the hypothesised direction, but was not statistically significant.

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										
	Childcare cost per hour (€)	101	(54/47)	2.26	(2.37)	1.86	(1.18)	ns	ns	0.21
	Age started childcare (months)	111	(58/53)	23.47	(9.54)	21.49	(10.60)	ns	ns	0.20
	Uses formal childcare	112	(58/54)	0.97	(0.18)	0.93	(0.26)	ns	ns	0.18
*	Uses grandmother care	112	(58/54)	0.03	(0.18)	0.06	(0.23)	ns	ns	0.10
	Satisfaction with childcare	111	(58/53)	0.86	(0.35)	0.85	(0.36)	ns	ns	0.04
	Child attends Síolta accredited centre	106	(56/50)	0.63	(0.49)	0.62	(0.49)	ns	ns	0.01
	Hours per week in childcare	111	(57/54)	20.19	(7.32)	20.56	(7.34)	ns	ns	0.05
Service Use										
	Childcare Services	151	(75/76)	0.99	(0.86)	0.82	(0.86)	ns	ns	0.20
	Community Information Services	151	(75/76)	1.19	(1.36)	0.96	(1.27)	ns	ns	0.17
	Employment/Adult Education Services	151	(75/76)	0.72	(1.17)	0.54	(0.94)	ns	ns	0.17
	Total no. of services	151	(75/76)	7.40	(4.07)	6.79	(4.47)	ns	ns	0.14
	Other Services	151	(75/76)	0.73	(0.45)	0.68	(0.47)	ns	ns	0.11
*	Emergency Services	151	(75/76)	0.32	(0.70)	0.38	(1.25)	ns	ns	0.06
	Health Services	151	(75/76)	2.47	(1.45)	2.41	(1.67)	ns	ns	0.04
	Residents Associations' Services	151	(75/76)	0.04	(0.20)	0.04	(0.20)	ns	ns	0.00
	Family Services	151	(75/76)	0.95	(0.57)	0.96	(0.72)	ns	ns	0.02
Non Step-down Measures										
	Uses any type of childcare	150	(74/76)	0.78	(0.41)	0.71	(0.46)	ns	-	0.17

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. [3] Indicates that the step-family was jointly significant in a left-sided test.

2.10 Household Factors & SES

There is substantial documentation of the impact of household and SES inequalities on children's skills and capabilities. These inequalities begin early in life and operate indirectly by influencing the availability of resources (housing, medical care), lifestyle and attitudes (neighbourhood quality, accidents), and the social and emotional context in which the child develops (parenting skills, marital disharmony, family composition) (Fonagy & Higgit, 2000; Foster et al., 2005). In the short term, it is evident that household, SES inequalities, and neighbourhood factors can have profound effects on child development and wellbeing (Bradley & Corwyn, 2002; Letourneau et al., 2011; Leventhal & Brooks-Gunn, 2000). For example, lower SES has been associated with increased internalising and externalising behaviours in children, and can negatively impact cognitive and language development (Kagan, 1992; Nagin & Tremblay, 2001). Furthermore, it can have detrimental long-term effects on future success in school, academic achievement, and even the likelihood of employment later in life (Duncan et al., 2007; Hirsch, 2007; Najman et al., 2004; Rouse et al., 2005). A comprehensive review of the literature surrounding the area can be found in the six month *PFL* report.

There are a limited number of home visiting evaluations which assess household factors and SES between twenty-four and thirty-six months. Furthermore, those that do report only a small number of significant findings and results tend to be inconsistent across studies. For example, Landsverk et al. (2002) in their evaluation of the Healthy Families America programme reported that at thirty-six months mothers in the

intervention group were more likely to be in education than those in the control group. However, they did not find significant effects on a variety of other household and SES indicators. Additionally, an evaluation of the MOM programme and the New Zealand-based Early Start programme found no significant programme effects following the analysis of a large number of household and SES factors (Fergusson, Horwood et al., 2005; Schwarz et al., 2012). In contrast, Love et al. (2002) in an evaluation of the Early Head Start programme focused on father involvement and found that between fourteen and thirty-six months children in the intervention group were less likely to have a continuous male presence than those in the control group.

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 live in the household, how many siblings the *PFL* child has, and whether or not the child's grandparent lives in the household. Additionally, the participant reported her current relationship status. This information was used to generate two binary indicators, the first denoting whether the participant was currently in a relationship (married, cohabitating, or partner) and the second denoting whether 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 were with when the child was twenty-four 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 twenty-four 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 includes 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 have 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 annual wage of mothers who worked part-time and mothers who worked full time.

FAMILY FINANCES

Participants' perception of financial difficulty was assessed by asking them to rate their level of satisfaction with their financial situation on a five 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 or were not satisfied with their financial situation. Participants were also 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 or not they worried about their financial situation. Participants were 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 indicating; whether anyone in the household received any social welfare payments, whether anyone had a medical card, whether they received one parent benefit or whether they received unemployment benefit. Participants also stated the household's weekly income from all sources, 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 is 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 of the household.

DIFFICULT LIFE CIRCUMSTANCES SCALE

The Difficult Life Circumstances scale (DLC; Johnson et al., 1989) identifies parents' perceptions of the existence of stressors and problems which are caused by factors such as substance, physical or emotional abuse, finances, community support and housing. *PFL* uses 15 items from this measure (α=0.67) and participants responded yes or no as to the presence or absence of each problem. The total score was calculated by summing all items, providing a range of scores from 0 to 15. Difficulties were also grouped into four domains: difficulty with partner, health difficulty, housing difficulty, and financial difficulty. Scores for each domain represent the number of difficulties experienced by participants in that domain.

NEIGHBORHOOD QUALITY EVALUATION SCALE

The Neighborhood Quality Evaluation Scale (NQES; Roosa et al., 2005) was used to assess participants' evaluations about the quality of their neighbourhoods. In response to each of the 11 items (α=.91), mothers were asked to indicate their levels of agreement, ranging from 1 (not at all true) to 4 (very true). Consistent with the scale's coding structure, some items were reverse coded. Items were summed and a mean was computed, with higher scores indicating a more positive attitude towards the neighbourhood.

NEIGHBORHOOD CRIMINAL EVENTS SCALE

Maternal perceptions of criminal events in their neighbourhoods were assessed using the 10 item (α=.92) Neighborhood Criminal Events Scale (NCES; Roosa et al., 2005). In response to each item, mothers were asked to indicate how often they felt each event happened in their area, with responses ranging from 1 (rarely or none of the time) to 4 (most or all of the time). Items were then summed and a mean score was computed, with higher scores indicating a higher occurrence of negative events or high neighbourhood disadvantage.

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.

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

Three of the five measures in the Household Factors category were in the hypothesised direction, however none indicated a statistically significant difference between the high and low treatment groups. In addition, the step-down test showed that the joint effect was not statistically significant.

MATERNAL EMPLOYMENT

All four of the measures in the Maternal Employment category were in the hypothesised direction, and three indicated a statistically significant difference. 29% of mothers in the high treatment group looked after the home/family, compared with 17% of mothers in the low treatment group (p<.05, d=.28). 23% of mothers in the high treatment group were unemployed compared with 34% of the low treatment group (p<.10, d=.25). In addition, 17% of the high treatment group reported an improvement in their work status between twenty-four and thirty-six months, compared with 9% of the low treatment group (p<.10, d=0.23). The step-down test showed that the joint effect of the four measures in the Maternal Employment category was not statistically significant.

PATERNAL EMPLOYMENT

None of the three measures in the Paternal Employment category were in the hypothesised direction, and two indicated a statistically significant difference in the non-hypothesised direction. 36% of the high treatment group fathers were unemployed, compared with 23% of the low treatment group ($p<.10$, $d=.29$). Only 5% of fathers in in the high treatment group, compared with 14% of the low treatment group, were reported to have had an improvement in work status between twenty-four and thirty-six months ($p<.10$, $d=.31$). In addition, the step-down test showed that the joint effect of the three measures in the Paternal Employment category was statistically significant in the non-hypothesised direction ($p<.05$), driven by the significant results for father unemployed and improvement in father's work status.

FINANCES

Six of the twelve measures in the Finances category were in the hypothesised direction, two of these differences indicated a statistically significant difference in the hypothesised direction and one indicated a statistically significant difference in the non-hypothesised direction. Specifically, 82% of the high treatment group anticipated a more favourable financial situation over the next 12 months, compared with 70% of the low treatment group ($p<.10$, $d=.28$). 43% of the high treatment group, compared with 54% of the low treatment group, currently reside in social housing ($p<.10$, $d=.22$). However, the low treatment group were less likely to be in receipt of unemployment benefit, with 24% reporting that they received unemployment benefit compared with 33% in the high treatment group ($p<.10$, $d=.22$) The step-down test showed that the joint effect of the twelve measures in the Finances category was not statistically significant.

NEIGHBOURHOOD

Two of the three measures in the Neighbourhood category were in the hypothesised direction, and one indicated a statistically significant difference. 73% of mothers in the high treatment group indicated that they were satisfied with their neighbourhood, compared with 62% of the low treatment group ($p<.10$, $d=.24$). The step-down test showed that the joint effect of this category was not statistically significant.

DIFFICULT LIFE CIRCUMSTANCES

Two of the five measures in the Difficult Life Circumstances category were in the hypothesised direction. One of the measures was statistically significant in the non-hypothesised direction. 21% of the high treatment group reported that they were having a difficulty with a past or present partner, compared with 12% of the low treatment group ($p<.10$, $d=0.24$). The step-down test showed that the joint effect of the Difficult Life Circumstances category was not statistically significant.

NON STEP-DOWN MEASURES

Four of the ten measures which were not included in the above step-down categories were in the hypothesised direction, however none of the measures were statistically significant.

Table 2.8 - Results for High and Low Treatment Groups: Household Factors and 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										
*	*Number of siblings	136	(68/68)	1.26	(1.18)	1.43	(1.33)	ns	ns	0.13
	Married	150	(74/76)	0.20	(0.40)	0.18	(0.39)	ns	ns	0.05
	Household size	150	(74/76)	4.49	(1.59)	4.47	(1.61)	ns	ns	0.01
	Has a partner	150	(74/76)	0.68	(0.47)	0.72	(0.45)	ns	ns	0.10
*	Resides with grandparent	150	(74/76)	0.24	(0.43)	0.17	(0.38)	ns	ns	0.18
Maternal Employment										
	Looking after the family/home	143	(73/70)	0.29	(0.46)	0.17	(0.38)	*p<.05*	ns	0.28
*	Mother unemployed	143	(73/70)	0.23	(0.43)	0.34	(0.48)	*p<.10*	ns	0.25
	Mother improvement in work status (between 24 and 36 months)	139	(72/67)	0.17	(0.38)	0.09	(0.29)	p<.10	ns	0.23
	Mother in paid employment	143	(73/70)	0.41	(0.50)	0.37	(0.49)	ns	ns	0.08
Paternal Employment										
	Father in paid employment	88	(45/43)	0.56	(0.50)	0.67	(0.47)	ns	ns	0.25
*	Father unemployed	87	(44/43)	0.36	(0.49)	0.23	(0.43)	s~	ns	0.29
	Father improvement in work status (between 24 and 36 months)	86	(42/44)	0.05	(0.22)	0.14	(0.35)	s~	ns	0.31
Finances										
	Household financial situation over the next 12 months	144	(73/71)	0.82	(0.39)	0.70	(0.46)	p<.10	ns	0.28
*	Living in social housing	150	(74/76)	0.43	(0.50)	0.54	(0.50)	p<.10	ns	0.22
*	In receipt of One Parent Family benefit	151	(75/76)	0.37	(0.49)	0.47	(0.50)	ns	ns	0.20
	Owns Home	150	(74/76)	0.31	(0.47)	0.26	(0.44)	ns	ns	0.11
	Saves regularly	150	(74/76)	0.41	(0.49)	0.37	(0.49)	ns	ns	0.08
*	Number of people supported by family income	150	(74/76)	4.45	(1.64)	4.47	(1.61)	ns	ns	0.02
*	Often worried about financial matters	150	(74/76)	0.45	(0.50)	0.45	(0.50)	ns	ns	0.00
	Household current financial situation compared to 12 months ago	150	(74/76)	0.58	(0.50)	0.61	(0.49)	ns	ns	0.05
	Equivalised weekly household income	143	(70/73)	222.11	(103.37)	231.84	(116.46)	ns	ns	0.09
*	Medical Card	151	(75/76)	0.72	(0.45)	0.66	(0.48)	ns	ns	0.13
	Satisfied with household financial situation	150	(74/76)	0.35	(0.48)	0.45	(0.50)	ns	ns	0.20
*	Household Unemployment Benefit	151	(75/76)	0.33	(0.47)	0.24	(0.43)	s~	ns	0.22
Neighbourhood										
	Satisfaction with neighbourhood	150	(74/76)	0.73	(0.45)	0.62	(0.49)	*p<.10*	ns	0.24
	Neighbourhood Quality Evaluation Scale	149	(73/76)	3.58	(0.97)	3.44	(0.94)	ns	ns	0.14
*	Neighbourhood Criminal Events Scale	149	(73/76)	1.03	(0.83)	0.97	(0.78)	ns	ns	0.08

Continued On Next Page.

	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*
Difficult Life Circumstances										
*	Housing difficulty	149	(73/76)	0.34	(0.48)	0.39	(0.49)	ns	ns	0.11
*	Health difficulty	149	(73/76)	0.33	(0.47)	0.37	(0.49)	ns	ns	0.08
*	Financial difficulty	149	(73/76)	0.38	(0.49)	0.33	(0.47)	ns	ns	0.11
*	Difficulty with neighbourhood	149	(73/76)	0.15	(0.36)	0.11	(0.31)	ns	ns	0.14
*	Difficulty with partner (past or present)	149	(73/76)	0.21	(0.41)	0.12	(0.33)	s~	ns	0.24
Non Step-down Measures										
	Partner is the child's father	105	(50/55)	0.96	(0.20)	0.89	(0.31)	*ns*	-	0.26
	Same partner as 12 months ago	103	(49/54)	0.98	(0.14)	0.96	(0.19)	ns	-	0.10
	Mother still in education	150	(74/76)	0.11	(0.31)	0.14	(0.35)	ns	-	0.11
*	Mother long-term unemployed	141	(72/69)	0.11	(0.32)	0.19	(0.39)	ns	-	0.22
	Mother in part-time employment	55	(28/27)	0.68	(0.48)	0.63	(0.49)	ns	-	0.10
	Mother's annual wage (full time only)	19	(9/10)	21801	(6439)	22696	(8222)	ns	-	0.13
	Mother's annual wage (part-time only)	34	(19/15)	11694	(5926)	13387	(7395)	ns	-	0.26
	Father's annual wage	41	(20/21)	24021	(11111)	25314	(13428)	ns	-	0.11
*	Father long-term unemployed	86	(43/43)	0.16	(0.37)	0.12	(0.32)	ns	-	0.13
*	Difficult Life Circumstances Total	149	(73/76)	2.14	(2.54)	1.71	(1.58)	ns	-	0.20

2.11 Main Results Summary

Results at thirty-six months are largely consistent with the home visiting programme evaluation literature. As expected, the positive findings are largely concentrated in the domains of child development, child health, and parenting, with limited significant findings in relation to social support and household factors and SES. However, contrary to expectations we also found a number of positive significant effects in the domains of home environment and maternal health. Most of the outcomes were in the hypothesised direction with the high treatment group reporting somewhat better outcomes than the low treatment group. The significant findings are summarised here.

CHILD DEVELOPMENT

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

- Children in the high treatment group scored higher than children in the low treatment group on the ASQ problem solving scale. The ASQ scores step-down category was significant as a result of this finding.
- Children in the high treatment group were also less at risk of being developmentally delayed on the ASQ problem solving scale.
- Children in the high treatment group scored lower on the CBCL total score and the CBCL external problems score than children in the low treatment group, indicating fewer overall behavioural problems and fewer externalising problems.
- Children in the high treatment group were less likely to score within the cut-off for total CBCL behaviour problems and CBCL external behaviour problems. The CBCL cut-off score step-down category was significant as a result of these findings.
- Children in the high treatment group scored lower on the CBCL somatic complaints, sleep problems, other problems, and aggressive behaviour problems subdomains than children in the low treatment

group. The CBCL subdomains step-down category was significant as a result of these findings.

- Children in the high treatment group scored higher on the DP-3 scale of cognitive development, and were more likely to score above the average DP-3 cut-off than children in the low treatment group.
- Children in the high treatment group were also more likely to receive higher ASQ standardised total scores than those in the low treatment group.
- Children in the high treatment group exhibited a lower standard deviation than the children in the low treatment group on 11 of the 28 continuous child development outcomes.

CHILD HEALTH

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

- Children in the low treatment group were more likely to have had an accident or to have stayed in hospital overnight in the last year than children in the low treatment group.
- Children in the high treatment group were more likely to have been diagnosed with a chronic illness. The long term child health step-down category was significant as a result of this finding.
- Children in the low treatment group were less likely to consume proteins than children in the high treatment group and children in the high treatment group were more likely to be meeting their dietary guidelines than children in the low treatment group.

PARENTING

Mothers in the high treatment group and mothers in the low treatment group did not differ significantly across several of the parenting domains including parenting daily hassles, and maternal attitudes to education. However, the following significant differences were identified:

- Children in the high treatment group spent less time watching TV, DVDs, or videos than those in the low treatment group, in addition they spent less time watching TV alone.
- Of those participants that impose a limit on TV time, mothers in the high treatment group allowed their children to watch less TV than those in the low treatment group.
- Households in the high treatment group reported that the TV is turned on for a shorter duration during the day than households in the low treatment group.
- The TV habits step-down category was significant as a result of the findings in relation to time spent watching TV, DVDs, or videos, time watching TV alone, maximum TV time allowed per day, and time TV is on in the home per day.
- Mothers in the high treatment group were less likely to engage in behaviours associated with either authoritarian or permissive parenting than those in the low treatment group. As a result of these findings the PSDQ step-down category was significant.
- Mothers in the high treatment group were also less likely to engage in punitive or hostile parenting than mothers in the low treatment group. The PSDQ authoritarian parenting step-down category was significant as a result of these findings.
- Mothers in the high treatment group were more likely to display indifference or neglect towards their child. As a result of these findings the PARQ step-down category was significant in the non-hypothesised direction.
- Mothers in the high treatment group scored higher on the PARQ, an indicator of parental coldness and low affection.

HOME ENVIRONMENT AND SAFETY

Households in the high treatment group and households in the low treatment group differed significantly on many of the home environment measures.

- Parents in the high treatment group scored higher on the HOME Involvement subdomain. Thus they are more likely than those in the low treatment group to be involved in the child's learning and promote learning development.
- Parents in the high treatment group also scored higher on the HOME subdomains of Organisation and Acceptance, thus are more likely to demonstrate an established family routine, safe environment, use of community supports, and acceptance of negative behaviours.
- Parents in the high treatment group scored higher, than those in the low treatment group, on a supplementary indicator of organisation of the environment and the ability to provide for the basic needs of the child.
- As a result of the findings in relation to the subscales Organisation, Involvement, and Acceptance, the HOME step-down category was significant.
- Parents in the high treatment group were more likely to report a higher level of organisation within the family as measured by the FES than those in the control group.
- Children in the high treatment group were less likely to be exposed to cigarette smoke.
- Households in the high treatment group were more likely to score higher on observational measures of home environment quality.

MATERNAL HEALTH AND WELLBEING

Mothers in the high treatment group differed significantly from mothers in the low treatment group across several health domains:

- Mothers in the high treatment group were less likely to experience symptoms of depression and reported greater levels of wellbeing. As a result of these findings the maternal mental health step-down category was significant.
- Mothers in the high treatment group were less likely to report drinking alcohol in the previous year.
- Mothers in the high treatment group were more likely to reduce their smoking between twenty-four and thirty-six months than those in the low treatment group.

MATERNAL SOCIAL SUPPORT

- Mothers in the high treatment group did not differ significantly from mothers in the low treatment group in terms of social support across such domains as partner support, support from friends and parents, voting behaviour, or quality of life. However, one significant difference was identified:
- Mothers in the low treatment group were less likely to receive regular child maintenance payments when compared to the high treatment group.

CHILDCARE AND SERVICE USE

Children in the high treatment group did not differ significantly from children in the low treatment group in any of the childcare and service use domains.

HOUSEHOLD FACTORS AND SES

Families in the high treatment group did not differ significantly from families in the low treatment group in terms of household composition, marital status, or income. However, there were a number of significant differences between the high and low treatment group in this domain:

- Mothers in the high treatment group were more likely to classify themselves as looking after the home or family than those in the low treatment group.
- Mothers in the high treatment group were less likely to be unemployed than those in the low treatment group, however someone in their household was more likely to be in receipt of unemployment benefit.
- Mothers in the high treatment group were more likely to have improved their work status in the last year than those in the low treatment group.
- Fathers in the high treatment group were more likely to be unemployed than those in the low treatment group. high treatment fathers were also less likely to have improved their work status in the last year, than those in the low treatment group. As a result of these findings the paternal employment step-down category was significant in the non-hypothesised direction.
- Mothers in the high treatment group were more optimistic than those in the low treatment group about their financial situation over the next twelve months.
- Families in the high treatment group were less likely to be living in social housing.
- Mothers in the high treatment group reported higher levels of satisfaction with their neighbourhood than mothers in the low treatment group.
- Mothers in the high treatment group were more likely to be experiencing difficulty with a past or present partner compared with those in the low treatment group.

SUMMARY

Overall, 204 outcome measures were assessed at thirty-six months. Of these one-tailed tests, 142 (70%) were in the hypothesised direction such that the high treatment group had better outcomes than the low treatment group, and 45 (22%) of these differences were statistically significant. These differences were found across all domains except childcare. Fifty-five (27%) 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 parenting and household factors and SES. Additionally, a number of positive effects were found in the domains of child development, child health, parenting, home environment, and maternal health and wellbeing. Of the 35 step-down categories tested, nine were significant in the hypothesised direction, representing 25% of all categories; age appropriate child development, child behaviour cut-off scores, child behaviour subdomains, long term child health, TV habits, parenting style, subdomains of parenting style, observations of the home environment, and maternal mental health. Two of the step-down categories tested were significant in the non-hypothesised direction, representing 6% of all categories; parental acceptance and rejection, and paternal employment.

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'. The following chapter describes and analyses *PFL* implementation practices regarding attrition, engagement, participant satisfaction, misreporting of participant responses and potential contamination between programme intake and when the *PFL* child was thirty-six months of age.

3.1 Participation Attrition up to Thirty-six 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 them, while for those who enrolled, between 20% and 80% exited them 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 even conflicting (Daro & Harding, 1999; Wagner et al., 2003; Roggman et al., 2008; Holland et al., 2014). In terms of the programme itself, attrition may increase if parental expectations are not met, or if they consider the content irrelevant (Roggman et al., 2008; Holland et al., 2014). This section investigates the level and determinants of attrition in the *PFL* sample between baseline and the thirty-six month survey.

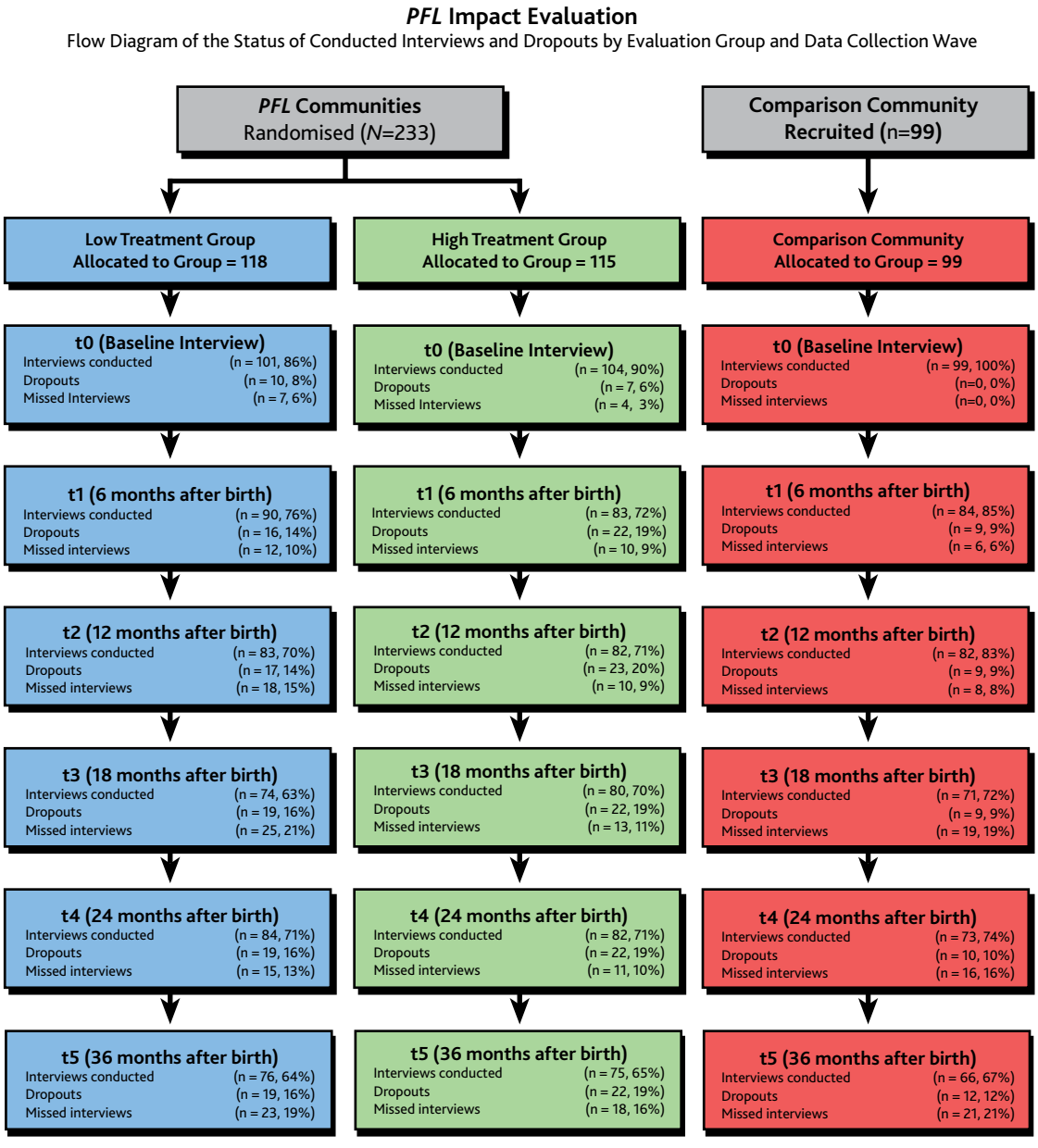
3.1.1 Attrition/Disengagement in *PFL*

The Consort Diagram (Figure 3.1) describes the progression of participants between programme entry and thirty-six months. In total, 217 thirty-six month interviews (n*High* = 75; n*Low* = 76; n*LFP* = 66) 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 thirty-six month completion rate was very similar for the high and (65.2%) low (64.4%) treatment groups and slightly higher for the comparison group (66.7%).

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 thirty-six months, with the highest dropout rate experienced among the high treatment group at 19%, while the low treatment group experienced a dropout rate of 16%, and 12% of the comparison group dropped out before thirty-six months. The dropout rate between twenty-four months and thirty-six months was minimal however. None of the high or low treatment group dropped out during this period, and two comparison group participants dropped out between twenty-four and thirty-six months.

In addition to those who dropped out, 18% of the sample did not complete a thirty-six 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 attempts by the evaluation team to be contacted or declined to be interviewed. At this time point the rates of disengagement across the high and low treatment groups were 16% and 19% respectively, and the corresponding rate for the comparison group was 21%. The overall level of disengagement was similar to that experienced at the eighteen month interview. There was some reengagement at twenty-four months following the eighteen month interview so it is possible that we may see a lower rate of non-completion for the forty-eight month interview as compared with the current wave.

Overall, 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 very similar across both groups (*High* = 36.5%; *Low* = 36.4%) from randomisation to the thirty-six month interview, with the majority of attrition/disengagement occurring prior to the six month interview.



Note: Dropout participants include both voluntary and involuntary dropouts.

Figure 3.1 Thirty-Six Month Consort Diagram

3.1.2 Analysis of Attrition/Disengagement before Thirty-Six Months

It is important to examine whether attrition and disengagement/missed interviews has led to systematic differences between the groups, which may bias the outcome results. The analysis below compares the baseline characteristics of participants who completed a thirty-six month interview to those who did not complete a thirty-six month interview. Thus, the analysis of 'attritors' includes those who have officially dropped out of the programme between baseline and thirty-six months and those who missed the thirty-six 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 two measures. Specifically, high treatment group mothers who did not complete a thirty-six month interview were less likely to be employed and had lower cognitive resources on average.

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	29	22.72	7.94	75	21.15	7.82	ns
Mother's age	29	25.00	6.31	75	25.64	5.69	ns
Partnered	29	0.72	0.45	75	0.80	0.40	ns
Married	29	0.14	0.35	75	0.15	0.36	ns
Living with parent(s)	29	0.59	0.50	75	0.56	0.50	ns
First time mother	29	0.52	0.51	75	0.55	0.50	ns
Low education	29	0.45	0.51	75	0.29	0.46	ns
Mother employed	29	0.21	0.41	75	0.43	0.50	*p<0.05*
Saves regularly	29	0.45	0.51	75	0.48	0.50	ns
Social housing	28	0.61	0.50	75	0.53	0.50	ns
Cognitive Resources (WASI)	29	77.93	11.93	75	83.65	12.17	*p<0.05*
Vulnerable attachment (VASQ)	29	19.03	3.20	75	17.93	3.94	ns
Self-efficacy (Pearlin)	29	2.67	0.72	75	2.81	0.59	ns
Self-esteem (Rosenberg)	29	12.34	2.98	75	13.00	2.57	ns
Knowledge of infant development (KIDI)	29	71.18	8.73	75	72.67	7.13	ns
Positive parenting attitudes (AAPI)	29	5.03	1.73	75	5.34	1.22	ns
Physical Health Condition	29	0.72	0.45	75	0.76	0.43	ns
Mental Health Condition	29	0.24	0.44	75	0.29	0.46	ns
Smoking during pregnancy	29	0.52	0.51	75	0.51	0.50	ns
Drinking during pregnancy	29	0.17	0.38	75	0.28	0.45	ns
Drug ever used	29	0.10	0.31	75	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, six statistically significant differences were found. Specifically, low treatment mothers who did not complete their thirty-six month interview were more likely to be first time mothers, were more likely to have low levels of education and lower levels of knowledge regarding infant development, the attritors were also younger, and were more likely to have taken drugs in the past. However, low treatment mothers who did not complete their thirty-six month interview were less likely to have been diagnosed with a mental health condition at the time of the baseline interview.

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	26	21.27	7.87	75	21.36	6.65	ns
Mother's age	26	23.31	5.78	75	25.99	5.94	$p<.05$
Partnered	26	0.88	0.33	75	0.83	0.38	ns
Married	26	0.19	0.40	75	0.17	0.38	ns
Living with parent(s)	26	0.54	0.51	75	0.44	0.50	ns
First time mother	26	0.65	0.49	75	0.44	0.50	$p<.10$
Low education	26	0.58	0.50	75	0.33	0.47	$p<.05$
Mother employed	26	0.27	0.45	75	0.44	0.50	ns
Saves regularly	26	0.46	0.51	75	0.53	0.50	ns
Social housing	26	0.62	0.50	75	0.53	0.50	ns
Cognitive resources (WASI)	26	79.04	13.91	75	81.56	12.54	ns
Vulnerable attachment (VASQ)	26	18.19	4.00	75	17.69	3.99	ns
Self-efficacy (Pearlin)	26	2.74	0.48	75	2.92	0.63	ns
Self-esteem (Rosenberg)	26	12.88	2.36	75	12.75	3.02	ns
Knowledge of child development (KIDI)	26	66.21	6.95	75	71.07	8.25	$p<.01$
Positive parenting attitudes (AAPI)	26	4.85	1.70	75	5.21	1.31	ns
Physical Health Condition	26	0.58	0.50	75	0.64	0.48	ns
Mental Health Condition	26	0.12	0.33	75	0.28	0.45	$p<.10$
Smoking during pregnancy	26	0.50	0.51	75	0.47	0.50	ns
Drinking during pregnancy	26	0.23	0.43	75	0.28	0.45	ns
Drug ever used	26	0.27	0.45	75	0.11	0.31	$p<.05$

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, five statistically significant differences were found. Specifically, comparison group mothers who did not complete their thirty-six month interview joined the programme later in their pregnancies, were less likely to have a partner at baseline, had lower levels of cognitive resources, lower levels of self-efficacy, and displayed fewer positive parenting attitudes 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	33	27.64	6.11	66	23.92	6.11	*p*<.01
Mother's age	33	26.97	6.28	66	27.44	6.21	ns
Partnered	33	0.79	0.42	66	0.91	0.29	*p*<.10
Married	33	0.21	0.42	66	0.17	0.38	ns
Living with parent(s)	33	0.33	0.48	66	0.32	0.47	ns
First time mother	33	0.33	0.08	66	0.45	0.50	ns
Low education	33	0.30	0.47	66	0.23	0.42	ns
Mother employed	31	0.39	0.50	66	0.45	0.50	ns
Saves regularly	30	0.53	0.51	66	0.58	0.50	ns
Social housing	31	0.47	0.51	66	0.41	0.50	ns
Cognitive resources (WASI)	33	83.97	15.83	66	89.74	12.75	*p*<.10
Vulnerable attachment (VASQ)	33	17.03	3.93	66	16.84	3.44	ns
Self-efficacy (Pearlin)	33	2.58	0.81	66	2.98	0.59	*p*<.01
Self-esteem (Rosenberg)	33	12.45	3.26	66	13.33	2.99	ns
Knowledge of child development (KIDI)	33	72.25	8.98	66	73.25	8.61	ns
Positive parenting attitudes (AAPI)	33	5.28	1.54	66	5.93	1.28	*p*<.05
Physical Health Condition	33	0.61	0.50	66	0.70	0.46	ns
Mental Health Condition	32	0.31	0.47	66	0.39	0.49	ns
Smoking during pregnancy	33	0.42	0.50	66	0.30	0.46	ns
Drinking during pregnancy	33	0.30	0.47	66	0.30	0.46	ns
Drug ever used	33	0.15	0.36	66	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.

3.1.3 Inverse Probability Weighted Results

As the previous analysis highlighted some differences between participants who completed a thirty-six 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 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 was tested using an inverse probability weighting (IPW) technique. This involved three steps. First, 137 permutation tests were estimated to more thoroughly examine the individual baseline factors associated with participation in the thirty-six month interview for the high and low treatment groups separately. Variables which were statistically significant (at the 10% level in a two-tailed test), and had no missing data were retained. If this procedure resulted in two or more highly correlated variables being included only one was kept. For the high treatment group five significant baseline characteristics were used: maternal cognitive resources [as per the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999)], maternal employment, parenting, and child rearing attitudes [as per the Adult Adolescent Parenting Inventory (AAPI; Bavolek & Keene, 2002)], maternal attachment style [as per the Vulnerable Attachment Style Questionnaire (VASQ; Mahon et al., 2003)], and whether or not the mother had a medical card at baseline. For the low treatment group 10 significant baseline variables were retained: maternal satisfaction with neighborhood, knowledge of infant development [as per the Knowledge of Infant Development Inventory (MacPhee, 1981)] maternal openness [as per the Ten Item Personality Index (TIPI; Gosling et al., 2003)], maternal age, maternal use of employment services, an indicator of low maternal education, maternal ethnicity, parenting and child rearing attitudes [as per the AAPI; Bavolek & Keene, 2002)], an indicator of whether the mother has ever had a mental health condition, and whether the mother has ever used drugs.

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

Third, 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 thirty-six month interview yet had baseline characteristics which were similar to those who dropped out. One participant who participated in the thirty-six 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 Unweighted & IPW Results at Thirty-Six Months

PFL Low - *PFL* High	Proportion of Measures Significantly Different at Thirty-Six Months		Proportion of Measures Significantly Different at Thirty-Six Months	
	Unweighted Individual Tests	Unweighted Multiple Hypothesis Tests	IPW Individual Tests	IPW Multiple Hypothesis Tests
Child Development	33% (39)	38% (8)	38% (39)	50% (8)
Child Health	24% (21)	33% (3)	14%(21)	66% (3)
Parenting	26% (34)	43% (7)	26% (34)	43% (7)
Home Environment	40% (15)	50% (2)	40% (15)	0% (2)
Maternal Health & Wellbeing	24% (17)	33% (3)	18% (17)	33% (3)
Social Support	5% (19)	0% (2)	5% (19)	0% (4)
Childcare & Service Use	0% (17)	0% (2)	6%(17)	0% (2)
Household Factors & SES	14% (42)	0% (6)	7% (42)	0% (6)
Total Statistically Different	22% (44/204)	26% (9/35)	20% (41/204)	29% (10/35)

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 twenty-four 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.

In the area of child development there were two 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 preformed significantly better than the low treatment group on the ITSEA Pro-social subdomain and the CBCL Attention subdomain. Additionally, the CBCL Domains multiple hypothesis test was significant in the weighted results but not in the unweighted results.

The unweighted results yielded two significant child health differences, which were not replicated in the IPW results. The unweighted results indicated that the high treatment group were significantly less likely to have an accident and stay overnight in hospital than the low treatment group. However, the difference between the two groups did not remain when the IPW procedure was applied. There was an additional significant multiple hypothesis test in the IPW adjusted results for the Food and Drink subgroup which was not identified in the unweighted results.

The unweighted parenting and home environment results had the same number of significant individual tests between the high and low treatment groups as the IPW results. However, the HOME step-down test was non-significant in the weighted results.

The unweighted results showed two significant differences in the area of maternal health and wellbeing, which were not replicated in the IPW results. Specifically, the unweighted results indicated that high treatment group mothers had higher WHO-5 scores and had a larger reduction in their smoking compared to the low treatment group. According to the weighted results, mothers in the high treatment group were more likely to have been pregnant following the birth of their *PFL* child than mothers in the low treatment group. However, this difference was non-significant in the unweighted results.

There was no overall change in the proportion of significant social support results; however different measures were significant in the weighted and unweighted results. The weighted results indicate that the high treatment group were significantly more likely to vote in local and European elections than the low treatment group. This effect was not replicated in the unweighted results, however, the high treatment group were more likely to receive regular child maintenance payments according to the unweighted findings.

The weighting procedure produced one significant effect in the childcare and service use domain which was non-significant in the unweighted results. The use of community services was significantly higher amongst the high treatment group once the weights were applied.

There were five variables in the household factors and SES domain that showed significant group differences in the unweighted results, which were not replicated in the IPW results. These variables measured whether the respondent was currently unemployed, whether their work status improved over the past twelve months, whether they think their household financial situation will improve over the next twelve months, whether they live in social housing, and finally whether or not the respondent is satisfied with their neighbourhood. Additionally, there were two significant differences in the weighted results which were not present in the unweighted results. Specifically, members of the high treatment group were significantly less likely to receive one parent benefit, and significantly more likely to be partnered with the father of their *PFL* child according to the IPW results.

In total 20% of the individual tests were significantly different in the hypothesised direction when IPW was used, which is slightly less than the 22% of measures on which significant differences were identified when no weighting was applied. However, a slightly higher proportion of the multiple hypothesis tests were significant when IPW was used (29%) versus when no weighting was applied (26%). In general, more significant differences were found in the IPW results for *PFL*'s primary outcome of child development, while the remaining domains had mainly fewer significant findings. However, in both the weighted and unweighted results the same proportion (28%) of multiple hypothesis tests was significant. As mentioned above the unweighted results identified a significant step-down difference for the HOME step-group while the IPW results identified a significant step-down test in the CBCL domains step-group.

3.1.4 Key Findings

The level of official attrition from *PFL* between baseline and thirty-six months was quite low at 16% across the whole sample. Importantly, the level of official attrition was minimal between the twenty-four and thirty-six month interview rounds, with no attrition experienced in the high or low treatment groups and only 2% attrition in the comparison group. Overall official attrition between programme intake and thirty-six months is slightly higher among the high treatment group (19%) than among the low treatment group (16%) 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 slightly greater proportion of the low treatment group (19%) was classified as disengaged or missed their thirty-six month interview when compared with the high treatment group (16%). Thus, it is possible that many of the participants who missed the thirty-six month survey represent participants who are less engaged with the programme and more inclined to dropout. Total non-completion (attrition & disengaged) at thirty-six months was very similar among the high (35%), and the low (36%) treatment groups, and slightly lower among the comparison group (33%). 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 thirty-six months survey to those who did not. Overall, there is some evidence that there are systematic differences between these groups. In general, we found that more disadvantaged participants were difficult to contact to schedule an interview or had dropped out of the programme by thirty-six months. For example, in the low treatment group, those who did not participate in this survey had lower levels of education and were more likely to have taken drugs in the past. Mothers in the high treatment group who did not participate were less likely to have been employed at the time of the baseline interview and had lower levels of cognitive resources. In the comparison group, mothers who did not complete a thirty-six month survey also had lower levels of cognitive resources; additionally they had lower levels of self-efficacy and were less likely to have had a partner at baseline.

To account for potential bias which differential attrition may introduce, the outcome analyses contained in Chapter 2 were re-estimated using an IPW technique. Slightly fewer of the individual tests showed

significant differences between the high and low treatment groups when the weighting was applied, as compared with the unweighted results. However, the same number of multiple hypothesis tests were significant in the weighted and unweighted results.

3.2 Participant Engagement up to Thirty-Six 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 her 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 et al., 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 thirty-six month survey.

3.2.1 INSTRUMENTS

Information on participant engagement within *PFL* was gathered from two sources – the *PFL* database maintained by the *PFL* mentors and survey responses from participants at the thirty-six 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 thirty-six months old, b) the percentage of prescribed home visits delivered between intake and thirty-six 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 thirty-six 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 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 3.5 provides a summary of participant engagement in the *PFL* programme between programme entry and thirty-six months of age for the high treatment group. The analysis is 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, and total engagement up to thirty-six 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 thirty-six 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 for 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 thirty-six months was 88 home visits. Table 3.5 shows that on average, participants in the high treatment group received 45.8 home visits between programme entry and thirty-six months. The minimum number of visits received was 1 and the maximum was 106. 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 12 month period between twenty-four and thirty six months the average number of visits delivered was 13.2. This figure is more than twice the number of home visits delivered between twelve and eighteen months and between eighteen and twenty-four months. Thus the rate at which home visits are occurring has risen. Figure 3.2 displays the variation in the number of home visits over the entire period.

These figures were used to calculate the proportion of prescribed home visits actually delivered. Table 3.5 shows that in total, based on a fortnightly prescribed visit, 51.9% of visits were delivered on average. The proportion is relatively similar in the different phases of the programme; however a larger proportion of visits were delivered in the most recent period, compared to the previous two periods.

Table 3.5 also reports the average and total duration of all home visits. These times are 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 46.8 hours participating in home visits. The minimum duration spent in home visits was 6 minutes and the maximum was 110.8 hours in total. Figure 3.3 displays the variation in the duration of home visits over the entire period.

Table 3.5 - Participant Engagement in Home Visits in *PFL* up to Thirty-six Months of Age

	Prenatal – Birth	Birth - 6 Months	6 Months – 12 Months	12 Months – 18 Months	18 Months – 24 Months	24 Months – 36 Months	Total
Prescribed no. of home visits (based on bi-monthly visits)	10	13	13	13	13	26	88
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	45.8 (25.0) 1-106
% 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	51.9 (28.0) 1-121
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	58.3 (11.5) 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	46.8 (27.4) 0.1-110.8
N	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 twenty-four 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.

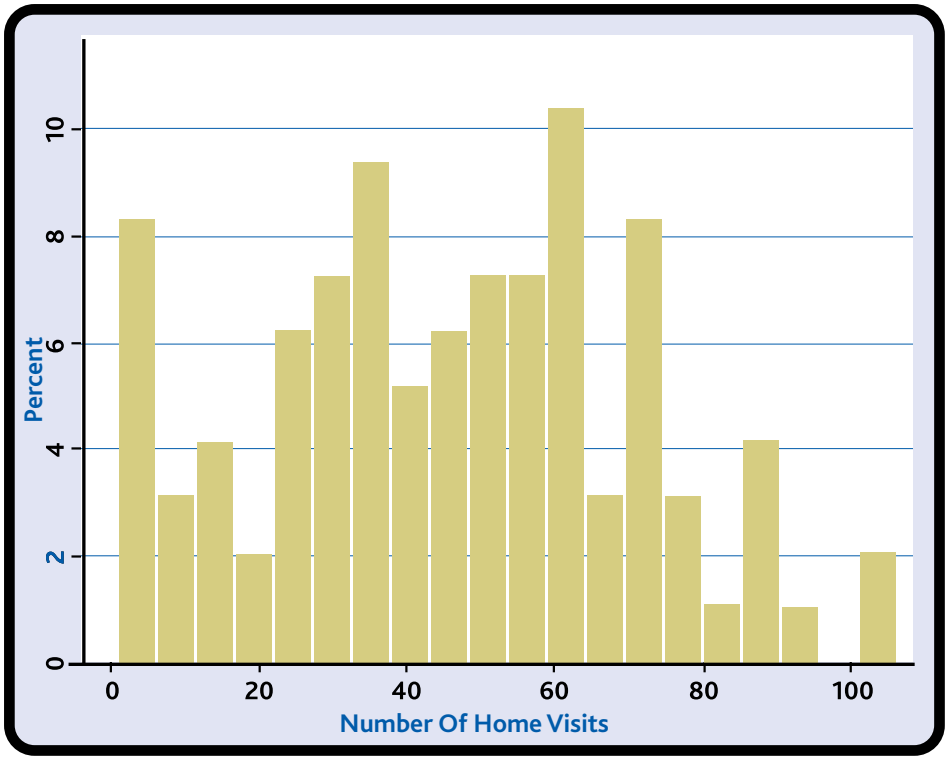


Figure 3.2 Variation in Number of Home Visits from Programme Entry to Thirty-Six Months

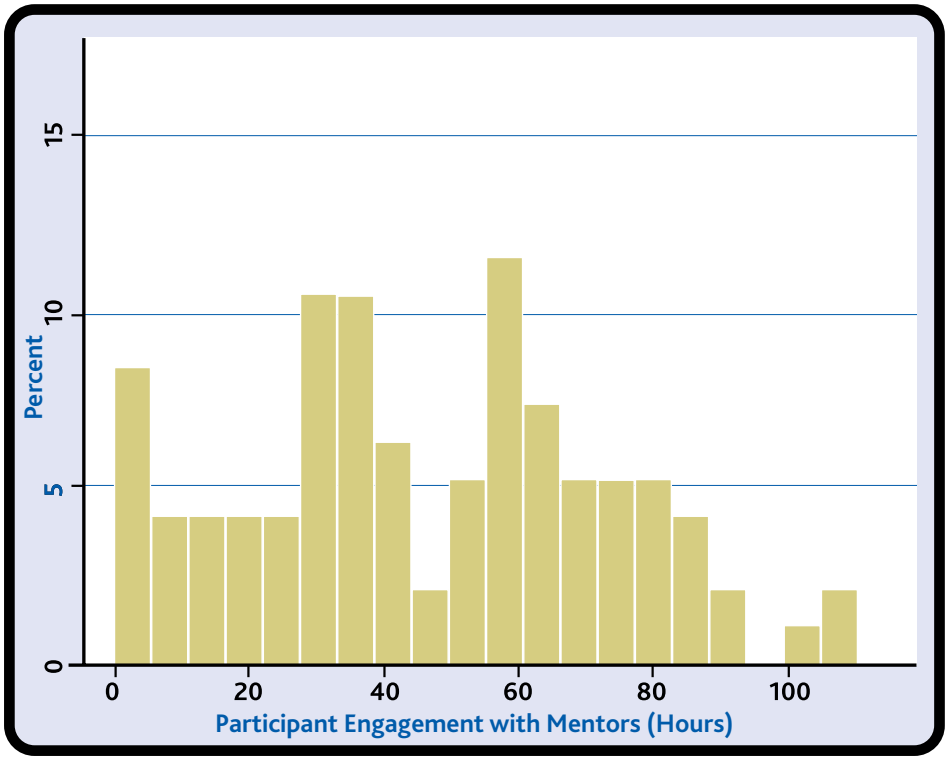


Figure 3.3 Variation in Duration of Home Visits from Programme Entry to Thirty-Six Months

3.2.3 Participant Engagement from Participant Interviews

HIGH TREATMENT GROUP

Based on participant responses to the thirty-six month interview, 5.3% of participants in the high treatment group reported meeting with their mentor once a week, 57.3% reported meeting twice a month, 26.7% reported meeting once a month, and 8.7% 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, 6% received fortnightly visits, 48% received monthly visits and approximately 46% 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 thirty-six month interview, 0% of participants in the low treatment group reported meeting the information officer more than once a month, 4% reported meeting the information officer once a month and 96% 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 information officer at their discretion.

3.2.4 Factors Associated with Engagement in Home Visiting

As described above, 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 thirty-six months. Specifically, we examined the relationships between participant engagement and a range of socio-demographic and maternal psychosocial factors collected at the baseline assessment. 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 3.6 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 total time in visits. Whereas mothers who smoked during their pregnancy had fewer home visits since joining the programme and spent a shorter amount of total time in visits.

Table 3.6 OLS Regression Model of Frequency & Duration of Home Visits Between Programme Entry and Thirty-Six Months

Dependent Variables	Frequency of Visits	Duration of Visits
	Prenatal - 36M	Prenatal - 36M
Weeks in pregnancy at programme entry	-0.56 (0.36)	-0.53 (0.39)
Mother's age	0.54 (0.78)	0.61 (0.83)
Partnered	6.46 (7.57)	9.41 (8.09)
Married	-9.60 (9.34)	-8.00 (9.98)
Living with parent(s)	1.96 (6.90)	2.82 (7.37)
First time mother	0.10 (7.95)	2.26 (8.50)
Low education	2.64 (6.72)	0.69 (7.18)
Mother employed	5.25 (6.29)	6.27 (6.72)
Saves regularly	-7.25 (5.93)	-9.40 (6.34)
Social housing	5.60 (5.91)	6.22 (6.32)
Cognitive resources (WASI)	0.53* (0.27)	0.59** (0.29)
Mental well-being (WHO5)	0.79 (0.65)	0.88 (0.69)
Vulnerable attachment (VASQ)	0.63 (0.84)	0.40 (0.90)
Self-efficacy (Pearlin)	1.23 (6.21)	2.67 (6.64)
Self-esteem (Rosenberg)	0.54 (1.44)	0.45 (1.53)
Knowledge of child development (KIDI)	0.17 (0.41)	0.34 (0.43)
Positive parenting attitudes (AAPI)	-2.07 (2.73)	-3.01 (2.91)
Physical Health Condition	0.01 (7.12)	-2.38 (7.61)
Mental Health Condition	4.57 (6.47)	5.41 (6.91)
Smoking during pregnancy	-13.70** (6.45)	-12.53* (6.89)
Drinking during pregnancy	2.07 (2.73)	1.45 (6.87)
Drug ever used	0.29 (7.72)	1.24 (8.25)
Constant	-35.32 (44.84)	-51.78 (47.89)
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 Key Findings

The analysis of participant engagement found that families in the high treatment group received an average of 45.8 home visits by the *PFL* mentors between programme entry and thirty-six months, representing 51.9% of prescribed home visits when based on fortnightly targets. Thus, over half of all home visits were delivered between programme entry and thirty-six months, which translates into about one home visit per month on average. This is consistent with the majority of home visiting programmes which typically find that over half of prescribed home visits are not delivered (Gomby et al., 1999). The average number of home visits delivered during the first six months (7.6), the second six months (6.9), the third six months (6.3) and the fourth six months (5.8) was broadly similar, suggesting that a regular pattern of visits has been established between the mentors and participants. However, 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 period, which reverses the pattern of reducing engagement. The average duration of home visits was in line with the *PFL* manual which recommended that each visit last between 30 minutes and two hours. The average duration of home visits during the twenty-four to thirty-six month period was just over 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 indicate 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 thirty-six months indicates 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, compared with those mothers who did not smoke during their pregnancy. 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 et al., 1999; Kitzman, Cole et al., 1997). That mothers in the *PFL* sample with higher cognitive resources participated in more home visits and had visits of a longer duration suggests 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 are 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 are consistent with the analysis of engagement reported in the six, twelve, eighteen, and twenty-four month reports, with very few individual characteristics associated with engagement at any time point. The report at forty-eight months will contain a final analysis of engagement patterns.

3.3 Participant Satisfaction up to Thirty-Six Months of Age

Participant satisfaction is an important aspect of any intervention as it can greatly impact commitment and engagement (Rao, 2000). According to previous findings, parents have reported high levels of satisfaction with home visiting programmes (Barth, 1991; Chaffin et al., 2004; McNaughton, 1994). However, there are a number of problems with measures of client satisfaction; for example, high ratings have been provided for programmes generating negative outcomes (Chafin & Friedrich, 2004). It is possible that parents feel indebted to service providers and therefore obliged to give high satisfaction ratings. Alternatively, participants may feel that providing negative ratings of parenting programmes may reflect poorly on their skills as a parent (Wesley et al., 1997). This literature suggests that while client satisfaction may be an important indicator of engagement in intervention programmes, it may not be an accurate reflection of programme efficacy.

3.3.1 Participant Satisfaction Instruments

CLIENT SATISFACTION

Client satisfaction was measured using the Client Satisfaction Questionnaire (CSQ; Turner et al., 1998). The CSQ (α = 0.91) is an adaption of the Therapy Attitude Inventory (Eyberg, 1993) which was developed to measure consumer satisfaction with parent training programmes. The CSQ addresses the clients' perception of the quality of the service that they received, how well the programme met both their needs and their child's needs, and whether the programme increased the parent's skills and reduced the child's problem behaviours. The CSQ contains 14 questions relating to how the participant feels about the programme. Questions 1 to12 are scored on a 7-point scale ranging from negative to positive ratings while items 13 and 14 are open ended questions. The Total Satisfaction measure is the sum of the other 12 measures except the improved relationship with partner measure, as this was a routed question dependent on whether the participant reported having a partner, thus implying a minimum value of 11 and a maximum value of 77 for the total score.

3.3.2 Participant Satisfaction Results

COMPARISON OF CLIENT SATISFACTION IN HIGH AND LOW *PFL* TREATMENT GROUPS

Table 3.7 compares the high and low treatment groups based on their responses to the CSQ. The first row reflects the 'Total Satisfaction' score and the remaining rows report the 12 individual items.

Total satisfaction with the programme is in the hypothesised direction, such that the high treatment group reported greater satisfaction with the programme than the low treatment group (p<.01, d=1.36). Eleven of the individual 12 programme satisfaction measures showed statistically significant differences between the high and low treatment groups, with the high treatment group consistently reporting greater satisfaction than the low treatment group on overall satisfaction with the programme (p<.01, d=1.11); whether the programme helped the mother to deal more effectively with their child's behaviour (p<.01, d=1.20); whether the programme helped to develop skills that can be applied to other family members (p<.01, d=1.31); the amount of help the mother and child received (p<.01, d=1.28); whether the programme met the child's needs (p<.01, d=1.15); whether the programme helped the mother to deal more effectively with family problems (p<.01, d=1.06); whether the programme met the mother's needs (p<.01, d=1.03); quality of the services received (p<.01, d=1.03); whether the participant received the type of help they wanted (p<.01, d=.96); whether the programme improved the mother's relationship with their partner (p<.05, d=.35); and finally mothers in the high treatment reported greater satisfaction with their child's progress (p<.10, d=.25). There was no difference between the high and low treatment group regarding the mothers' reports of improvements in child behaviour (d=.02).

In terms of the areas in which the participants were most satisfied, the high treatment group reported greatest satisfaction with the programme in general, followed by the amount of help they received and feelings about the child's progress. They reported the lowest level of satisfaction with whether the programme had improved their relationship with their partner. The low treatment group reported greatest satisfaction with feelings about the child's progress, followed by feelings about the child's behaviour, and overall satisfaction with the programme. The low treatment group also reported being the least satisfied with whether the programme had improved their relationship with their partner.

Table 3.7 Participant Satisfaction: High and Low Treatment Groups

Variable	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test $p^{\dagger}$	Effect Size *d*
Total Satisfaction	73/75	70.67	(6.62)	56.51	(13.21)	*p*<0.01	*1.36*
Help deal with child's behaviour	73/75	6.55	(0.97)	4.91	(1.47)	*p*<0.01	*1.20*
Helped develop skills	73/75	6.47	(0.85)	4.52	(1.93)	*p*<0.01	*1.31*
Amount of help received	73/75	6.75	(0.62)	5.25	(1.54)	*p*<0.01	*1.28*
Programme met child's needs	73/75	6.36	(1.10)	4.48	(2.04)	*p*<0.01	*1.15*
Overall satisfaction	73/75	6.78	(0.53)	5.53	(1.50)	*p*<0.01	*1.11*
Help deal with family problems	73/75	5.99	(1.14)	4.65	(1.38)	*p*<0.01	*1.06*
Programme met mother's needs	73/75	6.19	(1.25)	4.47	(2.02)	*p*<0.01	*1.03*
Quality of services received	73/75	6.55	(0.82)	5.27	(1.57)	*p*<0.01	*1.03*
Received type of help wanted	73/75	6.53	(0.99)	5.09	(1.88)	*p*<0.01	*0.96*
Improved relationship with partner*	73/75	4.15	(1.85)	3.51	(1.87)	*p*<0.05	*0.35*
Child's progress	73/75	6.74	(0.58)	6.55	(0.93)	*p*<0.10	0.25
Child's behaviour	73/75	5.77	(1.20)	5.79	(1.35)	ns	0.02

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. '*' indicates that this measure was not used to compute 'Total Satisfaction'. † one-tailed (right-sided) *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 Participant Satisfaction Key Findings

Overall, participant satisfaction with the programme between twenty-four and thirty-six months was high. As expected, the high treatment group reported greater satisfaction with the programme than the low treatment group. This likely reflects the greater number of supports and services provided to the high treatment group. However, the low treatment group still reported relatively high levels of satisfaction with the programme despite the minimal supports they receive. In addition, the areas where participants reported the highest and lowest levels of satisfaction are similar across the two groups. In line with previous findings, both groups were generally satisfied with the whole programme, their child's progress and the amount of help they receive from the programme. However, both groups reported being least satisfied with how the programme has improved their relationships with their partner. This is consistent with previous reports and may reflect the goals of the programme which are focused on improving child outcomes rather than family relationships more generally.

3.4 Misreporting

Social desirability refers to an individual's potential 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). 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, as cited in Mummendey, 1981; Edwards, 1953, 1957).

The main outcome 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 is

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 on validity 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.

3.4.1 Misreporting Instruments

The possibility of misreporting was tested in the thirty-six 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 Pyramid of Progression. 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 Pyramid of Progression, relative to the low treatment or comparison groups we may have reason to believe that they are more likely to answer in a manner which presents a favourable image of themselves.

3.4.2 Misreporting Results

Table 3.8 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 Pyramid of Progression. The findings indicate that a similarly low proportion (3%) of the low treatment group and comparison group participants claimed to have heard of the Pyramid of Progression. However, high treatment participants were more likely (9%) to report having heard of the phrase. The difference between the high and low treatment groups was statistically significant ($p<0.10$), however the difference between the high treatment group and the comparison group was not.

Table 3.8 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	*d*	*p*	*d*	*p*	*d*
'Pyramid of Progression'	0.093	(0.29)	0.026	(0.16)	0.031	(0.17)	$p<.10$	0.29	ns	0.26	ns	0.03
N	75		76		65							

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.

3.4.3 Misreporting Key Findings

The high treatment group participants are significantly ($p<0.10$) more likely to report having heard of the 'bogus' phrase Pyramid of Progression than are the low treatment group participants. This provides some evidence that relative to low treatment group participants, members of the high treatment group may be more likely to provide answers which they feel portray a better image of themselves as parents.

3.5 Contamination in *Preparing For Life*

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 treatment 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.

3.5.1 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 thirty-six 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 thirty-six 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. Nevertheless, both indirect and direct measures of contamination are used to provide an indication of whether contamination occurred during this period.

3.5.2 Indirect Measures of Contamination

Information to track contamination indirectly was collected from participants during the thirty-six month interview. Specifically, participants reported how many neighbours they knew personally, how many neighbours they knew who had a child the same age as their own child, and finally how many neighbours they knew who were taking part in the current study. 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 share their *PFL* materials with anyone else. The results of this analysis are presented in Table 3.9 below.

Table 3.9 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	0.92	(0.27)	0.92	(0.27)	0.86	(0.35)	ns	ns	ns
Knows neighbours with children the same age	0.75	(0.43)	0.75	(0.44)	0.77	(0.42)	ns	ns	ns
Knows neighbours taking part in *PFL*	0.49	(0.50)	0.49	(0.50)	0.11	(0.31)	ns	p<.01	p<.01
Shares *PFL* material with others	0.85	(0.36)	0.67	(0.47)	~	~	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 92% of the high treatment group and 92% the low treatment group knew other neighbours in their area personally and there was no statistical difference across the two groups. The second indirect measure of contamination shows that 75% of the

high treatment group and 75% the low treatment group knew other neighbours who have a child around the same age as their own child and again there was no statistical difference across the two groups. The third indirect measure of contamination shows that 49% of the high treatment group and 49% the low treatment group knew other neighbours taking part in the *PFL* programme and there was no statistical difference across the two groups.

A similar proportion of the comparison group knew other neighbours in their area personally (86%) and knew neighbours who have children of a similar age (77%), such that neither the high nor the low treatment groups knew a significantly different number of neighbours or people with similarly aged children, when compared with the comparison group. However, the comparison group were less likely to know a neighbour taking part in the study, with only 11% reporting that they know a neighbour who is participating. 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 interact 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 (85%) and low (67%) 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 is 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.

3.5.3 Direct Measures of Contamination

In order to elicit a direct measure of contamination a 'blue-dye' question was included in the thirty-six 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 'Emotional Literacy', and if they know what this phrase means. The phrase is related to a topic which only participants in the high treatment group should be aware of as the mentors discussed and promoted this behaviour with participants when delivering the programme between twenty-four and thirty-six months. Emotional Literacy is the ability to understand your emotions and empathise with others.

This question can be used to measure contamination as if the participants in the low treatment group state that they knew what the phrase meant and they correctly identified how to engage in this behaviour, it indicates 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 has occurred. 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 state that they are aware of this term or engage in this behaviour compared to the low treatment group and the comparison group. If contamination between the high and low treatment groups is 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 thirty-six month survey had two possible responses – *yes, no* which was used as 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 the following follow-up question: "Can you tell me what emotional literacy is". This was an open ended question which was subsequently recoded based on whether the respondent provided a valid response. Table 3.10 shows that 26% of participants in the high treatment group reported having heard the phrase Emotional Literacy, compared to 14% of the low treatment group and 8% of the

comparison group. The The differences between the high and low treatment groups and between the low treatment and comparison groups were non-significant. The difference between the high treatment group and the comparison group ($p<.01$, $d=.49$) was statistically significant.

Table 3.10 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	*d*	*p*	*d*	*p*	*d*
'Emotional Literacy'	0.26	(0.44)	0.14	(0.35)	0.08	(0.27)	*ns*	0.28	*p<0.01*	0.49	ns	0.22
N	74		76		66							

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.

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. 16% of the high treatment group, 5% of the low treatment group, and 2% of the comparison group gave an incorrect response. The results reporting the proportion of participants who accurately described Emotional Literacy were then re-estimated.

Table 3.11 shows that once we exclude participants who stated that they had heard of the phrase, yet provided incorrect responses there was no statistical difference between the proportion of any of the three groups reporting knowledge of Emotional Literacy.

Table 3.11 Corrected Comparison of Direct Contamination Responses across Groups

Have you heard the phrase ____ ? (corrected)	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	*High – Low*		*High – LFP*		*Low – LFP*	
							p	*d*	*p*	*d*	*p*	*d*
'Emotional Literacy'	0.11	(0.32)	0.10	(0.30)	0.06	(0.24)	*ns*	0.05	*ns*	0.18	ns	0.13
N	62		72		65							

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.

3.5.4 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 statistically significant 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 is as expected as the high treatment group receive 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 did not report a greater knowledge of the phrase than the low treatment group, or the comparison group. This suggests that contamination may have occurred between the high and low treatment groups. However, as a similar proportion of the comparison group, which could not have been contaminated by the high treatment group, also reported knowledge of the phrase, it suggests that the chosen phrase was simply a poor indicator of contamination and the intervention did not materially impact the high treatment group's knowledge of the phrase.

In addition, as the high treatment group had the highest proportion of incorrect responses it provides additional evidence that they may be more likely to misreport, relative to the other two groups.

3.6 Substitution bias

It is possible that members of the low treatment group are aware that other members of their community are receiving a more intensive intervention than they are personally receiving. If this occurs, members of the low treatment group may seek out or substitute similar parenting supports elsewhere. Assuming these 'off-programme' treatments are effective, then some of the treatment effects identified in Chapter 2 may be biased downwards. If this case, we would be comparing the high treatment *PFL* supports to the external supports which some, or all, of the low treatment group were exposed to. Thus, if both treatments were effective we would not identify significant treatment effects for *PFL*.

3.6.1 Instruments

In order to determine whether the low treatment group were accessing substitute treatments, all participants were asked whether they were taking part in any parenting programmes or classes (excluding *PFL*). Additionally, participants were asked whether anyone (again excluding staff from *PFL*) such as a friend, family member, or health care provider had provided them with books, leaflets, or guides about parenting. Finally, participants were asked whether any non-*PFL* staff member had provided them with professional advice about parenting. Table 3.12 tests whether the high, low and comparison groups differed on any of these items.

Table 3.12 Comparison of Parenting Advice/Materials 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* $p^{\dagger}$
Taking part in another parenting programme	0.00	(0.00)	0.05	(0.22)	0.08	(0.26)	ns	$p<.05$	ns
Received parenting books, leaflets, guides	0.26	(0.44)	0.20	(0.40)	0.68	(0.47)	ns	$p<.01$	$p<.01$
Received professional parenting advice	0.07	(0.25)	0.07	(0.25)	0.12	(0.33)	ns	ns	ns

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.

There were no statistically significant differences between the high and low treatment groups regarding the use of alternative parenting supports. None of the high treatment group and only 5% the low treatment group reported accessing another parenting programme, while a greater proportion of the high treatment group (26%) reported receiving parenting books, leaflets or guides from non-*PFL* staff compared to the low treatment group (20%). Finally, only 7% of both the high and low treatment groups reported receiving non-*PFL* professional parenting advice. However the comparison group, who are unaware of the *PFL* programme, reported significantly higher incidences of taking part in parenting programmes and receiving parenting books, leaflets, or guides than the high treatment group.

3.6.2 Key Findings

Overall, the results presented in Table 3.12 provide no evidence that the low treatment group are significantly more likely to seek out alternative treatments to the *PFL* programme than are members of the high treatment group. Thus there is no evidence of substitution bias suggesting that the treatment effects estimated in Chapter 2 are unlikely to be affected by this issue. It is interesting to note that the comparison group, who are generally a higher SES group than the treatment groups, are actively seeking out parenting support in the absence of any *PFL* support in their community.

Chapter Four

Comparison Group and Dynamic Results Summary

This chapter presents summaries of the dynamic analysis which examined changes in child and parent outcomes between baseline and thirty-six months for the high and low treatment group. It also summarises the results comparing the thirty-six 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 Low Treatment and Comparison Group Analyses

A number of standardised instruments used to evaluate the *PFL* programme are 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 and the thirty-six 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
Child Development	Ages & Stages Questionnaire (ASQ)		X	X	X	X	X
	Ages & Stages Questionnaire (ASQ) Social/Emotional		X	X	X	X	X
	Brief Infant Toddler Social and Emotional Assessment (BITSEA)			X	X	X	X
	Developmental Profile 3 (DP-3)			X	X	X	X
	Child Behaviour Check-List (CBCL)					X	X
Maternal Health & Wellbeing	Maternal Psychological Wellbeing (WHO-5)	X	X	X			X
	Edinburgh Postnatal Depression Scale (EPDS)		X		X	X	X
	Maternal Social Support (MSSI)				X		X
Parenting	Parenting Daily Hassles Scale				X		X
	Activities With Child		X		X		X
Home Environment	Home Observation for Measurement of the Environment (HOME)		X		X		X
	Family Environment Scale (FES)			X			X
	Difficult Life Circumstances (DLC)				X		X
Participant Satisfaction	Client Satisfaction Questionnaire (CSQ)		X	X		X	X

Generalizied Estimating Equations (GEE) are used to evaluate changes in parent and child outcomes over time and to determine whether the programme's impact increases or diminishes 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 one unit (i.e. one 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 is assumed, rather it is left as undefined.

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 in nature, 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 the interaction variables between time and treatment.

4.1.1 Hypotheses

In total, 64 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 thirty-six months differed significantly from the size of treatment effects arising from previous waves, on the same measures. The results are summarised below.

CHILD DEVELOPMENT

The ASQ Scores were standardised to a mean of 100 and standard deviation of 15 to ensure comparability amongst the time points. All of the other child development measures were comparable across time and thus did not require standardising. The comparison of the standardised ASQ scores found one statistically significant difference between the high and low treatment groups. On the ASQ Problem Solving subdomain the high treatment group experienced an increase in score between twenty-four and thirty-six months. However, the low treatment group's score fell. This led to a significant widening of the gap between the two groups.

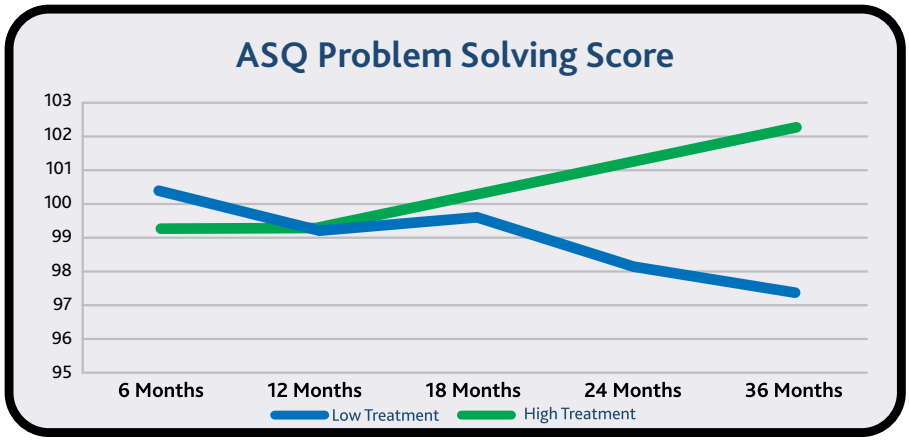


Figure 4.1 High and Low Treatment ASQ Problem Solving Scores over Time

There was also one significant interaction between treatment status and intervention wave for the Somatic Complaints subdomain of the CBCL. Somatic Complaints decreased for the high treatment group between twenty-four and thirty-six months, however it rose for the low treatment group over the same time period.

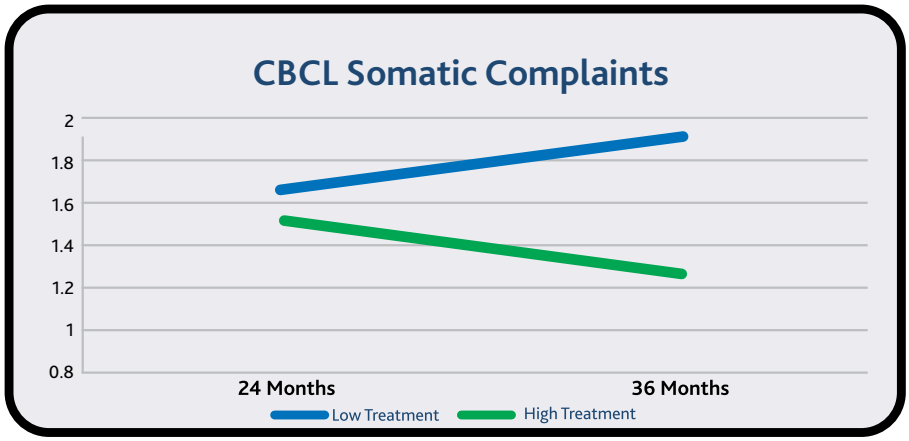


Figure 4.2 High and Low Treatment CBCL Somatic Complaints Scores over Time

MATERNAL HEALTH & WELLBEING

The comparison of the high and low treatment groups on measures of maternal health and wellbeing revealed two significant interactions between treatment status and interview wave. Firstly, on the WHO-5 the high treatment group experienced an increase in their total score between twelve and thirty-six months. However, the low treatment group's score fell slightly, resulting in a significant widening in the gap between the groups.

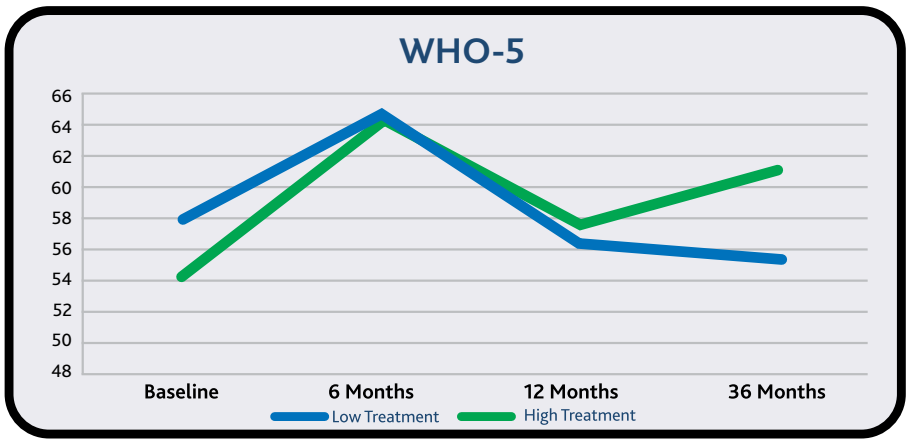


Figure 4.3 High and Low Treatment WHO-5 Total Scores over Time

The other significant interaction was found on the EPDS. The proportion of the low treatment group classified as being at risk of depression has risen significantly faster than the proportion of the high treatment group in the same category.

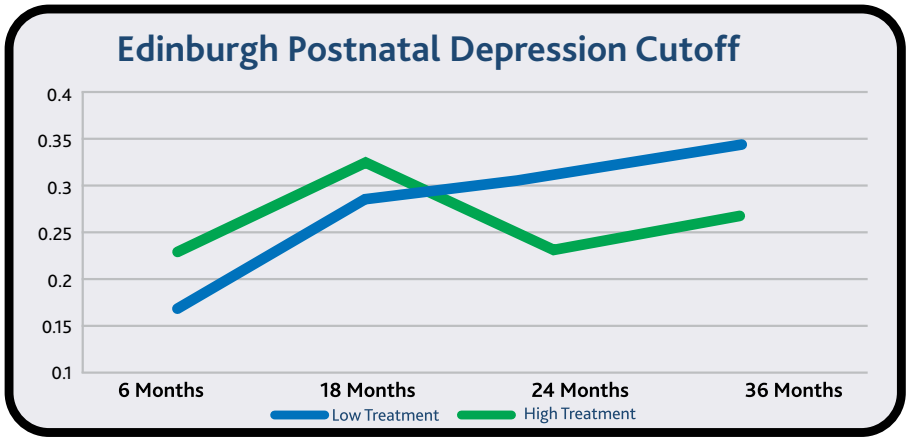


Figure 4.4 High and Low Treatment Edinburgh Postnatal Depression Cut-off Scores over Time

PARENTING

Comparing the high and low treatment groups on measures of parenting revealed no significant differences over time.

HOME ENVIRONMENT

The comparison of the HOME scores over time identified four significant differences between the high and low treatment groups. The first area where a dynamic result was identified was the HOME Organisation subdomain which pertains to the degree of routine in a family's schedule, safety of the environment, and community supports utilised. In this case, the scores for both groups fell between eighteen and thirty-six months, however the rate of decline was significantly greater for the low treatment group.

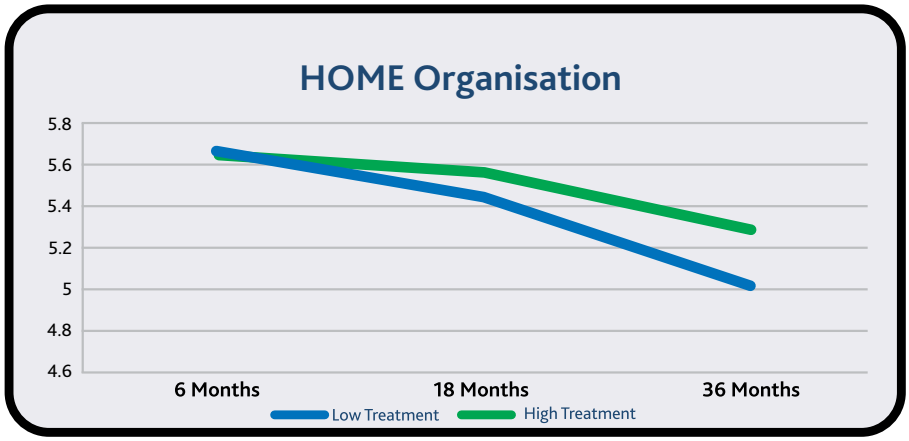


Figure 4.5 High and Low HOME Organisation Scores over Time

The opposite pattern was observed for both the Learning Materials and the Variety subdomains of the HOME measure. The Learning Materials subdomain assesses the appropriateness of play materials for the child, while the Variety subdomain assesses visitation of people and attendance of activities that introduce variety into the child's life. The high and low treatment groups' scores rose between eighteen and thirty-six months on both subdomains, however, the rise experienced by the low treatment group was significantly larger than the equivalent rise for the high treatment group. This means that the scores of both groups are converging and the programme's impact has diminished on these two domains. These are both non-hypothesised findings.

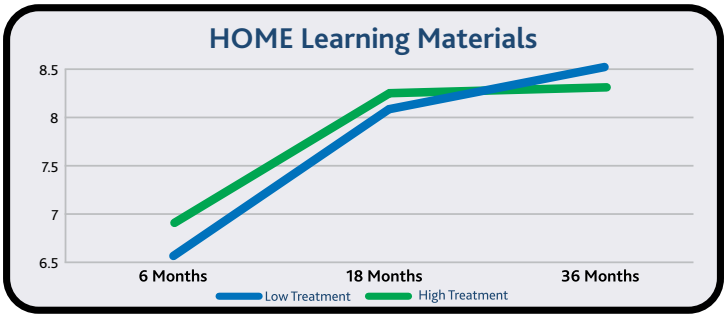


Figure 4.6 High and Low Treatment HOME Learning Materials Scores over Time

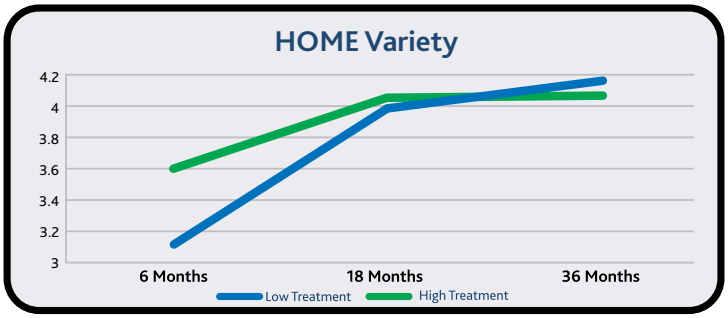


Figure 4.7 High and Low Treatment HOME Variety Scores over Time

PARTICIPANT SATISFACTION

There were significant differences between the high and low treatment group in terms of the participants' satisfaction with the programme over time. Between twenty-four and thirty-six months the low treatment group's overall satisfaction fell, while the high treatment group experienced a slight increase in satisfaction. This resulted in the gap between the two groups widening significantly.

Figure 4.8 High and Low Treatment Client Satisfaction Questionnaire over Time

SUMMARY

Overall nine (14%) of the 64 measures upon which dynamic analyses were conducted, yielded significant findings relating to changes in the size of treatment effects at thirty-six months relative to previous waves. Seven (11%) of these nine findings were in the hypothesised direction implying the strengthening of a positive treatment effect. However, two (3%) were in the non-hypothesised direction, which means the programme's impact is diminishing over time in these two areas. The hypothesised findings were concentrated in four areas: child development, maternal health and wellbeing, the home environment, and participant satisfaction. Whereas the non-hypothesised findings were both found in the home environment.

4.2 Low Treatment and Comparison Group Analyses

This section presents the results comparing the thirty-six month outcomes of the low treatment group to the external, no treatment comparison group. As the *PFL* community is geographically small, the purpose of the comparison group was to have an external community sample that was not at risk of contamination from the high treatment group. If the low treatment group outperforms the comparison group it may be an indication that the low treatment group has been 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.

4.2.1 Hypotheses

We hypothesise that there would be a limited number of statistically significant differences between the *PFL* low treatment group and the LFP community comparison group in both the hypothesised and non-hypothesised direction. Controlling for the baseline differences between the groups, a finding that the low treatment group have outperformed the comparison group at thirty-six months suggests that either the low treatment group may have received some the high treatment supports or that the low treatment supports were effective.

4.2.2 Key Findings

In total, 193 measures were included in the analysis, of which positive significant differences between the low treatment group and the comparison group were found on 19 (10%) of the individual measures. Conversely, of the individual measures 28 (15%) were statistically significant in the non-hypothesised direction such that the comparison group outperformed the low treatment group. The finding that the comparison group outperformed the low treatment group on a higher proportion of measures suggests that contamination did not have an impact, and also suggests that the low treatment supports are having minimal effect.

CHILD DEVELOPMENT

Children in the low treatment group and children in the comparison group differed significantly across many of the child development domains, with the low treatment group generally outperforming the comparison group on this domain:

- Children in the low treatment group had higher scores on the ASQ Fine Motor and Personal Social domains than children in the comparison group. However, children in the comparison group were less likely to fall into the ASQ Gross Motor cut-off category.
- Children in the low treatment group were reported to exhibit less behavioural problems, as measured by the BITSEA Problem score, than children in the comparison group. This result was driven by differences between the groups on the BITSEA Internalising and BITSEA Externalising subdomains. Also, according to the ITSEA measure, children in the low treatment group were more likely to engage in pro-social behaviour than the children in the comparison group.
- Children in the comparison group were also more likely than children in the low treatment group to fall into the clinical cut-off category for CBCL Internalising problems.

CHILD HEALTH

Children in the low treatment group and the comparison group differed significantly across many child health domains, with the comparison group generally outperforming the low treatment group:

- Children in the low treatment group were more likely to have spent a night in hospital during the last twelve months and less likely to have good overall health than children in the comparison group.

- Children in the low treatment group were less likely to have a disability than children in the comparison group, but more likely to suffer from a chronic condition.
- A lower proportion of children in the low treatment group met the daily requirements for consumption of dairy products than comparison children. Children in the comparison group also had higher overall diet quality scores.

PARENTING

Mothers in the low treatment group and the comparison group did not differ significantly across the majority of the parenting domains. However, three significant differences were identified:

- Mothers in the low treatment group were less likely to have had their child on a waiting list for primary school than mothers in the comparison group.
- A higher proportion of comparison mothers believe that education helped them get ahead in their life.
- A higher proportion of comparison mothers talk about television with their child.

HOME ENVIRONMENT AND SAFETY

Family home environments in the low treatment group and the comparison group differed in a number of ways:

- Families in the comparison group scored better than families in the low treatment group on the HOME Organisation domain.
- Low treatment families had higher scores than comparison families on the Control subdomain of the Family Environment Scale, however the comparison group had higher scores on the Organisation subdomain.
- Comparison families were less likely than low treatment families to have someone in the house who smokes other than the mother.

MATERNAL HEALTH AND WELLBEING

Mothers in the low treatment group did not differ significantly from mothers in the comparison group in terms of substance use, however a number of significant differences between the two groups were identified in terms of physical and mental health:

- Mothers in the low treatment group reported fewer visits to their GP compared to the comparison group. The maternal physical health step-down category was significant as a result of this finding.
- Mothers in the comparison group had lower levels of depression relative to the low treatment mothers, as measured by the EPDS. In addition mothers in the comparison group had higher levels of emotional wellbeing as measured by the WHO-5 Index.
- Mothers in the low treatment group that were pregnant were less likely to report that their pregnancy had been planned.
- Mothers in the low treatment group were less likely to have been pregnant since the birth of their child in the programme than comparison group mothers.

MATERNAL SOCIAL SUPPORT

Mothers in the low treatment group did not differ significantly from mothers in the comparison group in terms of social support across the majority of domains such as father support, Family Quality of Life, and voting behaviour. However, two significant differences were identified:

- Mothers in the low treatment group were less likely to receive a lot of support both from their friends and their parents when compared to the comparison group.

CHILDCARE AND SERVICE USE

The low treatment group did not differ significantly from the comparison group in terms of the number of hours per week the child was enrolled in childcare, the age at which child was placed in childcare, the cost per hour of childcare, and satisfaction with childcare. However, there were two significant differences between the low treatment and comparison groups in this domain:

- The low treatment group were more likely to use formal childcare and less likely to use grandmother care than the comparison group.

HOUSEHOLD FACTORS AND SES

Families in the low treatment group did not differ significantly from families in the comparison group on the majority of household factors such as the reporting of difficult life circumstances. However, there were a number of other significant differences, with the comparison group generally outperforming the low treatment group in this domain:

- Mothers in the low treatment group were less likely to be married than comparison group mothers.
- Both fathers and mothers in the comparison group were less likely to be unemployed than parents in the low treatment group. However, mothers in the low treatment group were less likely to receive unemployment benefit than comparison group mothers. Also a higher proportion of mothers in the comparison group experienced a positive change in their work status between twenty-four and thirty-six months.
- Mothers in the low treatment group were also less optimistic about the future financial situation of their household over the next 12 months.
- Comparison group mothers had more positive and less negative feelings regarding the area they live than do mothers in the low treatment group.

Overall, the mixed results of the low treatment group and comparison group analysis support the study design as they suggest that the low treatment group is not performing systematically better than the comparison group across most domains. The low treatment group outperformed the comparison group on 10% of measures, while the comparison group outperformed the low treatment group on 15% of measures. There were positive significant findings for the low treatment group in all domains except for maternal social support. In contrast, there were significant findings in favour of the comparison group in all domains with the exception of the childcare and service use domain. The lack of a coherent pattern within the findings suggests that the low treatment group is not receiving the services and supports designed for the high treatment group.

Chapter Five

Triple P - Positive Parenting Program

Once children in the programme reached twenty-four months an additional treatment the Triple P Positive Parenting Program, was offered to the high treatment families. Triple P is a gold standard parenting intervention, which seeks to promote optimal parenting practices and secure attachment relationships between parents and children. The programme targets the home environment, parental self-awareness, and parenting practices. This chapter provides an overview of Triple P, the international evidence regarding its effectiveness, and the role of Triple P in *PFL*. In addition, it reports on the implementation of Triple P up to thirty-six months and considers the thirty-six month outcome results disaggregated by *PFL* participants' Triple P participation status. Lastly it provides an overview of the key findings from the Triple P analysis.

5.1 Introduction to Triple P

The aim of the Triple P Positive Parenting Program is to promote healthy parenting practices and positive attachment relationships between parents and children (Sanders et al., 2003). Established in Australia in 1981, Triple P utilises social learning principles to prevent problems in children's development by targeting parenting practices. It was initially developed as a small-scale parenting intervention for a specific group of at-risk families (Sanders, 2008) and has since been regularly refined amidst consistently positive evaluation findings (e.g. Sanders et al., 2014). Due to its strong evidence base, it received the number 1 ranking by the United Nations in their compilation of evidence-based family skills training programmes (UN Office on Drugs & Crime, 2009). It receives state funding for national public health implementation in Australia, and is also delivered in 24 further countries. In Ireland, Triple P is currently being offered by the Midlands Area Parenting Partnership in Longford, Westmeath, Laois, and Offaly, and by *PFL* in Dublin.

STRUCTURE

Triple P is a "tiered continuum of interventions of increasing strength but narrowing reach" (Sanders et al., 2003). There are several specialised types of Triple P with different target groups. Specifically, Triple P has evolved to include five programme levels of increasing strength, from a media campaign (Level 1) to intensive individual family interventions (Level 5). More intensive interventions are offered on an individual basis, while those with a broader focus are implemented at group level.

- Level 1: Media and communication strategy on positive parenting
- Level 2: Brief parenting interventions
- Level 3: Narrow focus parenting interventions
- Level 4: Broad focus parenting programmes
- Level 5: Intensive family interventions

(Sanders, 2012)

Within each level, there are a number of different potential delivery methods, enabling the programme to be tailored to its target community .

THEORETICAL FRAMEWORK

Triple P is underpinned by a number of robust psychological theories (for further information, see Sanders et al., 2003). Its guiding principles focus on three areas: the home environment, parent self-awareness, and parenting techniques. It promotes the importance of an environment that is safe, engaging, and facilitates learning. Participating parents are educated about the positive effect of a stable nurturing parent-child relationship on parental wellbeing, and are therefore encouraged to take care of their own mental and physical health needs. In terms of parenting skills, parents learn to use assertive disciplining techniques while having realistic expectations of what their children can achieve (Sanders et al., 2003). These core principles underpin all levels of Triple P. On a practical level, specific areas explored within Triple P include child management strategies, principles of self-regulation, knowledge about parent-child interaction,

and family issues. Strategies for dealing with problems are typically specific and systematic, and focus on positive reinforcement of good behaviour while minimising parental reaction to challenging behaviour.

5.1.1 Effectiveness of Triple P: International Evidence from a Meta-Analysis

Triple P has been extensively evaluated internationally using multiple methods. A recent meta-analysis (Sanders et al., 2014) collated the effects of Triple P from 101 studies. Results were analysed separately for short-term data (pre and post-test) and for follow-up data (up to thirty-six months). Five constructs were investigated: children's social, emotional and behavioural outcomes, parenting practices, parenting satisfaction and efficacy, changes in parental adjustment, and parental relationships. Most of the studies included examined specific Triple P levels (Level 4 in many cases) and involved non-experimental as well as experimental designs.

Results of the short-term impact combining all Triple P intervention levels revealed highly significant differences (p<.001) in a positive direction for children's social, emotional and behavioural outcomes, parenting practices, parenting satisfaction and efficacy, parental adjustment, and parental relationship (Sanders et al., 2014). Overall effect sizes varied from medium (d=0.225, parental relationship) to large (d=0.578, parenting practices). In addition, outcomes varied according to the intensity of intervention provided. There were significant findings in a positive direction for child outcomes, parenting practices, and parenting satisfaction and efficacy on all levels from 2 to 5. Individually, effects on parental adjustment and parental relationships were significant if Level 4 (broad focus parent training) was provided. However, Level 1 (media interventions) exhibited a weaker effect, with moderate significance for parenting satisfaction and efficacy and non-significant results for other examined constructs (Sanders et al., 2014).

The analysis of the follow-up data indicated that there were significant differences for all five constructs if the effect of Triple P intervention Level 4 was measured separately or if data from all Triple P intervention levels were combined (Sanders et al., 2014). Additionally, the effects of Level 2 (brief parenting interventions) on child outcomes and parenting practices were shown to be significant, but non-significant for parenting satisfaction and efficacy and parental relationships. Level 3 (narrow focus parenting interventions) generated weakly significant differences on parenting practices and parental adjustment (p<.05). There were no significant findings for Level 1 (media interventions). In comparison to other levels, the short and long-term effects of Triple P Level 4 interventions are underpinned by the large number of studies in the meta-analysis on this level. A considerably smaller number of studies in the meta-analysis examined other Triple P levels (Sanders et al., 2014).

In comparison to other parenting programmes with a similar theoretical basis (e.g. The Incredible Years; Webster-Stratton, 1998) Triple P is unique in its population approach as the broad focus targets a vast number of parents and its multilevel system meets various needs of programme intensity, delivery modalities and languages.

5.1.2 Participant engagement in Triple P

There is a small body of evidence regarding the predictors of parental participation and engagement in group-based parenting programmes such as Triple P. Low engagement may threaten the programme's efficacy and the validity of the evaluation results. Definitions of 'engagement' are inconsistent and the roll-out of Triple P internationally leads to diverse factors influencing engagement in Triple P depending on the social and cultural context. In line with previous research, Dumas et al. (2006) found that for mothers of pre-schoolers, time constraints were a significant predictor of lower engagement in the PACE parenting programme. Participants who enrolled for the programme tended to be more financially disadvantaged and reported high levels of personal or familial stress.

[1] For the purpose of this report, we describe the delivery methods chosen by *PFL* only. For further information on the range of delivery methods available, please see the Triple P website: http://www.triplep.net/glo-en/the-triple-p-system-at-work/the-system-explained

A randomised trial in Zürich (Eisner & Meidert, 2011) examined the implementation of group-based Level 4 Triple P to parents of first year primary school children. Triple P was offered to 821 parents of which 257 (31.3%) enrolled in the programme and 28.4% participated in the research interviews. Results identified several family characteristics that were associated with the likelihood of enrolment. Families with dural-earners, three or more children, or a low socioeconomic status were less likely to participate in the course. Furthermore, significantly more parents with a high density of neighbourhood networks enrolled (51.6 %) and attended at least one session. Although dual-earner families and large families were less likely to attend at least one Triple P session, these characteristics did not predict a lower likelihood of course completion. In addition, neither the level of perceived problems nor a multicultural background was related to dropouts during the course once participants had attended at least one session. Yet parents with a low socioeconomic status or with poor neighbourhood integration were significantly more likely not to complete the course. Significantly more participants who attended all Triple P sessions were likely to use Triple P techniques in their parenting practice, but it should be noted that fidelity of Triple P techniques was not measured.

Similar findings were reported in a German study by Heinrichs et al. (2005) such that a greater likelihood of programme participation in Triple P was associated with a higher socioeconomic status. In contrast, Bodenmann et al. (2008) did not find differences in demographic factors (except citizenship) regarding attrition within a Swiss sample participating in Group Triple P (50 couples, drop-out rate 4%). Although the literature on engagement in Triple P is limited, most of the available studies found socially disadvantaged participants were less likely to engage and drop-out rates of around one third were frequently reported.

5.2 Triple P in *PFL*

Triple P was offered to all high treatment group participants when the *PFL* child was at least two years old. The programme was delivered by the mentors, who received direct training from core Triple P staff in the delivery of group, primary care, seminar series, and lifestyle Triple P. While a number of parenting programmes were considered by *PFL*, Triple P was selected as it was evidence-based, offered shorter courses and had an option of one-to-one delivery. Initially offered to *PFL* families in Group format only, a decision was taken to deliver it at the Primary Care level also, when considered appropriate by the mentor. In these cases, salient aspects of Triple P, selected by the mentor specifically for that family, were delivered on an individual basis in place of the usual mentor visit.

As both group and primary care attendance levels have declined since the initial introduction of Triple P in 2011, *PFL* added individual seminar and discussion group options in an effort to maximise uptake of Triple P with minimal parental commitment required. The seminars cover individual broad topics, each of which is based on a specific Triple P strategy. The discussion groups are each focused on one key behavioural area. While there are a number of delivery methods available within each Triple P level, as per the programme's flexible design, *PFL* selected the most appropriate methods for their target families. Thus, as illustrated in Table 5.1, 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). Through a combination of these sessions and follow-up individual contact, it was anticipated that parents would learn how to manage and address their children's behaviour, while promoting competence and development.

Table 5.1 Triple P Delivery in *PFL*

Level	*PFL* delivery method	Number of sessions	Session duration	Format	Session location	Programme focus
2	Seminar Series	3 individual sessions, delivered separately in a group format	90 mins	Group	*PFL* premises	Each seminar provides a general introduction to one of the established Triple P strategies: e.g. The Power of Positive Parenting, Raising Confident, Competent Children.
3	Primary Care	4 weekly sessions with mentor	30-60 mins	Individual	Participant's home	These consultations target a particular aspect of child's behaviour or development e.g. tantrums, toilet training, aggression
3	Discussion Groups	4 individual sessions	90 mins	Group	*PFL* premises	Each group has a specific topic, e.g. developing bedtime routines
4	Group Triple-P	4-5 weekly group classes & 2-3 individual phone calls	120 mins	Group	*PFL* premises	Positive parenting skills for multiple child behaviour issues

5.3 Implementation of Triple P at Thirty-Six Months

CHARACTERISTICS OF TRIPLE P ATTENDEES

Of the families who completed the *PFL* twenty-four month interview (n=82), 59% (n=48) took part in some form of Triple P. Table 5.2 shows a comparison of the baseline characteristics of the parents who took part in Triple P (n=48) versus those who did not (n=34). Those who took part in Triple P were more likely to be living with their own parents at baseline ($p<.10$), and the mother of the *PFL* child was more likely to be in paid employment during pregnancy ($p<.05$). They typically had higher cognitive resources ($p<.10$), yet were more likely to have a mental health condition ($p<.01$) than non-participants. Finally, those who took part were less likely to have smoked during pregnancy ($p<.10$). There were no other significant differences between the two groups.

Table 5.2 Comparison of Baseline Characteristics between Triple P Participants and Non-Participants: High Treatment Group

Variables	Triple-P Participants			Non-Participants			Individual Test
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*p*
Weeks in pregnancy at programme entry	48	21.17	8.23	34	22.88	7.51	ns
Mother's age	48	26.17	5.67	34	25.41	6.17	ns
Partnered	48	0.83	0.38	34	0.74	0.45	ns
Married	48	0.21	0.41	34	0.09	0.29	ns
Living with parent(s)	48	0.63	0.49	34	0.44	0.50	$p<.10$
First time mother	48	0.52	0.50	34	0.53	0.50	ns
Low education	48	0.29	0.46	34	0.29	0.46	ns
Mother employed	48	0.52	0.50	34	0.29	0.46	$p<.05$
Saves regularly	48	0.48	0.50	34	0.50	0.51	ns
Social housing	48	0.52	0.50	34	0.56	0.50	ns
Cognitive Resources (WASI)	48	84.71	11.60	34	81.03	13.06	$p<.10$
Vulnerable attachment (VASQ)	48	18.27	3.83	34	17.62	3.96	ns
Self-efficacy (Pearlin)	48	2.79	0.63	34	2.84	0.56	ns
Self-esteem (Rosenberg)	48	12.83	2.56	34	13.24	2.71	ns
Knowledge of infant development (KIDI)	48	73.42	7.11	34	71.18	6.90	ns
Positive parenting attitudes (AAPI)	48	5.17	1.32	34	5.36	1.09	ns
Physical Health Condition	48	0.71	0.46	34	0.82	0.39	ns
Mental Health Condition	48	0.38	0.49	34	0.12	0.33	$p<.01$
Smoking during pregnancy	48	0.42	0.50	34	0.62	0.49	$p<.10$
Drinking during pregnancy	48	0.31	0.47	34	0.24	0.43	ns
Drug ever used	48	0.15	0.36	34	0.15	0.36	ns
Psychological Wellbeing (WHO-5) Index	48	13.67	4.94	34	14.32	4.98	ns
Baby Gender	48	0.52	0.50	34	0.44	0.50	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.

ATTENDANCE

Between their child's second and third birthdays, 50 *PFL* families availed of some form of Triple P regardless of whether the completed a twenty-four month interview or not. Figure 5.1 presents the number of participants who attended at least one Triple P session at any level. It should be noted that some families attended more than one type of Triple P.

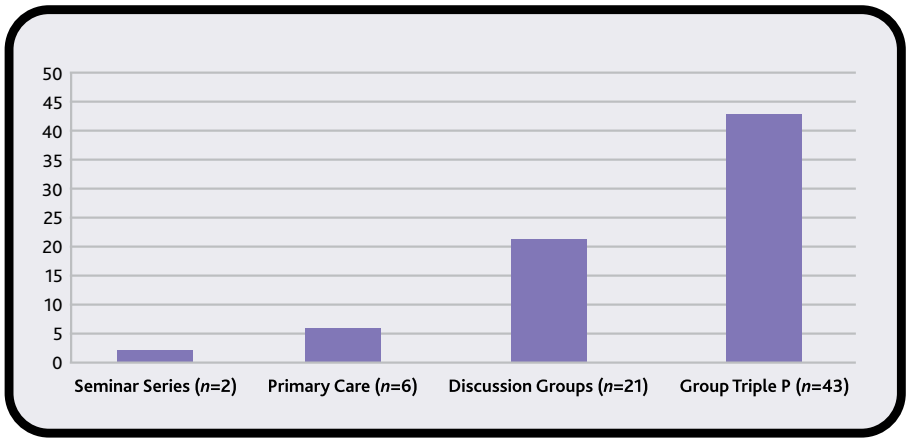


Figure 5.1 Triple P Attendance by Delivery Method

While Triple P was available to both the child's mother and her partner, mothers constituted three-quarters of attendees (n=50 out of a total of 67 attendees). Of those mothers who attended Triple P, one-third of their partners also attended some form of Triple P (n=17). There were no cases of partners attending Triple P without mothers attending. The number of partners who attended each type of Triple P is presented in Figure 5.2, which shows that partners were most likely to attend Group Triple P (Level 4).

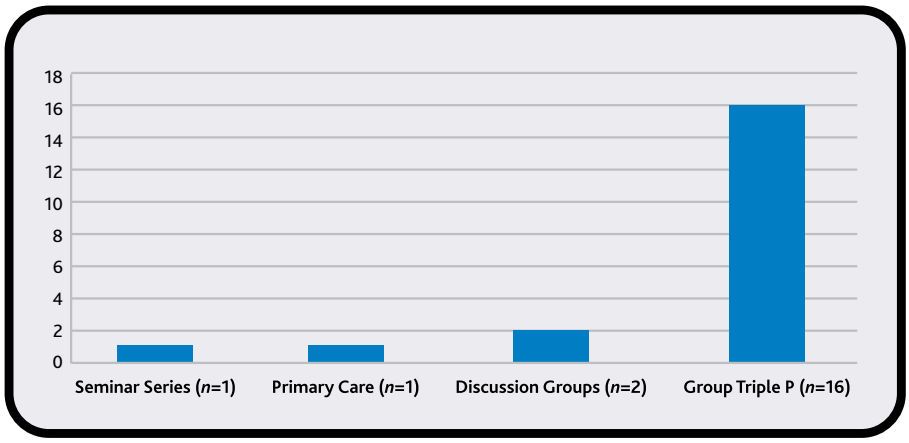


Figure 5.2 Triple P Attendance by Partner

The Seminar Series (Level 2) and Discussion Groups (Level 3) comprised individual sessions only. Completion of Primary Care and Group Triple P required attendance at a number of sessions.

[2] This number included two former partners of *PFL* mothers.

DISCUSSION GROUPS

Each discussion group was offered on two occasions to maximise participant attendance. The 21 participants who took part in discussion groups included two participants who attended discussions on all four topics, five participants who attended two each and 14 who attended only one. Figure 5.3 illustrates the number of participants who attended each discussion group. Of the four options, bedtime routine was the most popular, followed by fighting and aggression. Dealing with disobedience and hassle-free shopping had fewer attendees.

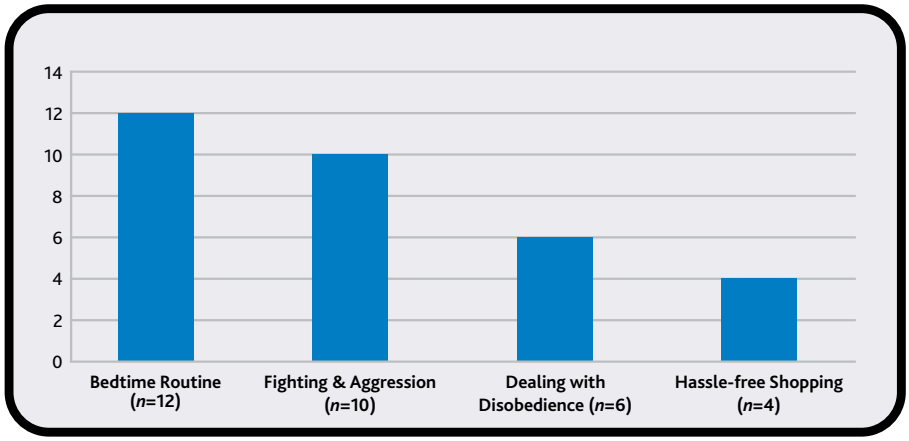


Figure 5.3 Discussion Group Attendance

SEMINARS

One Triple P seminar was held in June 2013 and was attended by two *PFL* mothers and one partner.

PRIMARY CARE

To complete Primary Care Triple P, a participant was required to attend all four sessions. Using this definition, of the six families who took part in primary care Triple P, four completed it, while two did not. The average number of sessions attended by those who did not complete Primary Care Triple P was 2.75.

GROUP

Defining completion of Group Triple P was more complex. While the programme in full consisted of four to five group sessions in addition to two to three individual phone calls, there was some variation in perceived requirements among participants. Therefore, *PFL* mentors made a decision about completion on a case-by-case basis, depending on how much they felt the participant had committed to and gained from Triple P. Participants were usually defined as completers if they attended at least three group sessions. Using this definition, of the 43 participants who took part in group Triple P, 84% (n=36) completed. The average number of sessions attended by those who did not complete Group Triple P was 1.55.

Of the 16 partners who took part in group Triple P, 50% (n=8) completed. The average number of sessions attended by partners who did not complete Group Triple P was 1.75.

5.4 Thirty-Six Month Outcome Results by Triple P Status

The results presented in Chapter 2 comparing the outcomes of the high and low treatment groups were based on the accumulated impact of the home visiting sessions and Triple P Program. In this section we test whether the high treatment participants who received Triple P had different outcomes to the high treatment participants who did not receive Triple P. This was analysed by re-estimating the tests presented in Chapter 2 separately for both subgroups. Table 5.3 summarises the main outcome results for the entire high treatment group versus the low treatment group, and includes separate results for the high treatment Triple P participants versus the low treatment group, and the high treatment non-Triple P participants versus the low treatment group. Relevant differences between the high treatment group who participated in some form of Triple P and the low treatment group are controlled for in the permutation tests, as are differences between the high treatment group who did not receive Triple P and the low treatment group. As in Chapter 2 the tests comparing the entire high treatment group versus the low treatment group are unconditional.

Given that the sample size is lower when we only consider a subset of the high treatment group we would expect that a lower percentage of the subgroup tests to reach significance, relative to the main analysis that included the entire high treatment group. In the main analysis we found that the whole high treatment group outperformed the low treatment group on 22% of the individual tests. Table 5.3 shows that the Triple P participants outperformed the low treatment group on 18% of the individual tests, while the Triple P non-participants outperformed the low treatment group on 19% of the individual tests. This could suggest that the addition of Triple P between twenty-four and thirty-six months had little impact on outcomes; however, the pattern of findings was decidedly different. First, a higher percentage of multiple hypothesis tests were significant when only Triple P participants were considered (23%) versus the non-participants (9%). In addition, the domains on which significant differences were observed were not consistent across both high treatment subgroups. High treatment non-Triple P participants outperformed the low treatment group on 38% of the child development measures and 33% of the child health measures, while the equivalent figures for high treatment Triple P participants were only 13% and 14% respectively, suggesting that Triple P did not have an impact on child health and development. However, the opposite pattern was observed in the parenting and home environment domains. The Triple P participants outperformed the low treatment group on 29% of the parenting measures and 33% of the home environment measures, while the proportion of significant differences between those who did not receive Triple P and the low treatment group was only 18% and 7% respectively on those domains. Both subgroups differed from the low treatment group on 6% of the maternal health and wellbeing measures. On the remaining three subdomains - social support, childcare and service use and, household factors and SES - the Triple P participants were significantly different to the low treatment group on a higher proportion of measures (16%, 18%, 17%) than when the non-Triple P participants were compared to the low treatment group (11%, 6%, 12%).

Table 5.3 Comparison of Thirty-Six Month High and Low Treatment Group Results for Full High Treatment Sample, Triple P Only Participants and Non-Participants Only

PFL Low - *PFL* High	Proportion of Measures Significantly Different at Thirty-Six Months					
	Full Sample		Triple P Participants		Triple P Non-Participants	
	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests
Child Development	33% (39)	38% (8)	13% (39)	13% (8)	38% (39)	13% (8)
Child Health	24% (21)	33% (3)	14%(21)	0% (3)	33%(21)	0% (3)
Parenting	26% (34)	43% (7)	29% (34)	57% (7)	18% (34)	0% (7)
Home Environment	40% (15)	50% (2)	33% (15)	50% (2)	7% (15)	0% (2)
Maternal Health & Wellbeing	24% (17)	33% (3)	6% (17)	0% (3)	6% (17)	0% (3)
Social Support	5% (19)	0% (2)	16% (19)	25% (4)	11% (19)	0% (4)
Childcare & Service Use	0% (17)	0% (2)	18%(17)	0% (2)	6%(17)	0% (2)
Household Factors & SES	14% (42)	0% (6)	17% (42)	17% (6)	12% (42)	33% (6)
Total Statistically Different	22% (44/204)	26% (9/35)	18% (37/204)	23% (8/35)	19% (38/204)	9% (3/35)

KEY FINDINGS

Triple P received a relatively good uptake by *PFL* families, with 50 mothers and 17 partners receiving at least one Triple P session. This take-up rate is largely consistent with the take-up rate for prescribed home visiting sessions. Group Triple P was the most popular type of Triple P offered by *PFL*. Of the 43 mothers and 16 partners who attended Group Triple P, 84% and 50% respectively completed the prescribed dosage. Of the six *PFL* families who took part in Primary Care Triple P, four completed the allocated series. The Discussion Groups, which required less time commitment on the part of families, were attended by 21 mothers and two partners. Finally, two families attended the *PFL* seminar series.

In total, the Triple P participants and the non-participants outperformed the low treatment group on a similar proportion of the individual tests, however the Triple P participants outperformed the low treatment group on a greater proportion of multiple hypothesis tests. The Triple P participants outperformed the low treatment group on a higher proportion of measures in five of the eight domains, most notably in the areas of parenting and the home environment. However, Triple P non-participants outperformed the low treatment group on a greater proportion of measures on the child development and child health subdomains. This suggests that Triple P may have impacted on some dimension of parenting and the home environment, however, these effects have not yet translated into impacts on children's health and development.

Chapter Six

The Day in the Life of a *PFL* Parent Study

6.1 Overview

The Day in the Life of a *Preparing For Life* Parent study (DTL) ran from October 2012 – November 2013 through funding awarded to Dr Orla Doyle by the Irish Research Council with the assistance of project consultants Professor Liam Delaney and Dr Michael Daly (University of Stirling). The study sought to complement the main *PFL* evaluation by providing a targeted investigation of the impact of the *PFL* programme on parental wellbeing. A multi method approach was employed utilising a unique combination of measurement techniques including a day reconstruction method, which records participants' activities and emotional/affective states over the course of the day (Day Reconstruction Method; DRM, Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004), global questions about mood and life satisfaction, and a standardised measure of parenting stress (Parenting Stress Index Short Form; PSI; Abidin, 1995). We also examined biomarkers across the study day including electrodermal activity (EDA), which acts as a marker for emotional arousal, and accelerometer data, which indicates participants' level of movement. This chapter focuses on the results arising from the DRM, global mood and life satisfaction questions and the PSI[1]. It details the background to the study, the methodology employed, the results arising from the study, and a summary of the study findings. The results are discussed in combination with the thirty-six month impact evaluation findings in Chapter 8.

6.2 Background

Improving maternal wellbeing is one goal of home visiting programmes. Home visiting programmes tend to target families at risk of low wellbeing (in terms of income, education, mental health, and or parenting resources), in order to mitigate intergenerational transmission to children's wellbeing. To achieve this, home visiting programmes work directly with parents, thus, according to Sweet and Appelbaum (2004), in order to be considered effective, home visiting programmes must demonstrate effects both for parents and their children. While child outcomes are widely studied, there is less evidence on the impact of home visiting programmes on parent outcomes. The studies that are available show that programmes are effective for certain parent outcomes but not others. The prevailing pattern, based on meta-analytic findings, suggest that home visiting programmes are more effective in improving parenting outcomes (behaviour and attitudes), than maternal life course outcomes (employment, self-sufficiency, reliance on public assistance), with the exception of effects for mothers seeking education (Sweet & Appelbaum, 2004). This differential impact is intuitive as home visiting programmes specifically target parenting skills.

In regards to maternal psychological wellbeing, research has focused predominantly on home visiting programmes' impact on negative aspects of wellbeing. This is driven by a substantial literature illustrating the deleterious effects of stress and depression on parent functioning and child wellbeing (e.g., Crnic & Low, 2002; Murray, Fiori-Cowley, Hooper, & Cooper 1996). Depression, in particular, affects a considerable proportion of low SES mothers enrolled in home visiting programmes and undermines the impact of these interventions (Ammerman, Putman, Bosse, Teeters, & Van Ginkel, 2010). However, Ammerman and colleagues' (2010) systematic review found that home visiting programmes are not sufficiently powerful, in and of themselves, to substantially mitigate depression. In addition, families enrolled in home visiting programmes, who typically face considerable socioeconomic risk, are exposed to pervasive stressors that diminish their parenting abilities (McLeod & Shanahan, 1993). As with depression, home visiting programmes tend not to be effective in reducing parental stress (Sweet & Appelbaum, 2004).

Comparatively fewer studies have examined the impact of home visiting programmes on positive aspects of wellbeing, such as self-efficacy and self-esteem. This is despite theories of self-efficacy (Bandura, 1977), which link people's beliefs about their capabilities to their subsequent motivation, behaviour, and wellbeing, being central to many home visiting programmes. Studies that have examined these aspects of wellbeing are inconclusive, and have yet to be subject to systematic review. While programmes such as ProKind (Jungman et al., 2012), and the Nurse Family Partnership (Kitzman et al., 1997) have demonstrated

[1] Information on biomarkers collected in the study will be included in the forty-eight month report.

treatment effects for self-efficacy, no effects were observed on measures of self-efficacy and self-esteem employed in the Healthy Families America (Mitchell-Herzfeld et al., 2005), Early Intervention Program for Adolescent Mothers (Koniak-Griffin et al., 2002), and the Family Partnership Model (Barlow et al., 2007) evaluations.

Collectively, this evidence supports Brooks-Gunn and Markman's (2005) inference that it may be easier to alter parenting behaviour than emotional states. However, there are important methodological issues that preclude definitive conclusions on this issue. Evaluations of maternal wellbeing in home visiting programmes have relied almost exclusively on standardised measures of wellbeing and single item questions relating to life satisfaction. While such measures typically offer strong psychometric properties and cross-study comparability, they also necessitate retrospective and aggregated appraisals of a participant's circumstances and feelings about them. As such, they may overlook nuanced aspects of wellbeing that relate to participants' daily life experiences. To address these issues in the wellbeing literature Kahneman and colleagues (2004) proposed the use of the Day Reconstruction Method (DRM) to elicit variations in wellbeing over the course of the day. The DRM obtains participants' ratings of a series of episodes across their previous day on several dimensions of positive and negative emotional states.

As Dolan and Metcalfe (2012) point out, global evaluations of wellbeing (e.g., questions on life satisfaction) and experience-based measures sometimes converge, but often do not. For example, Knabe et al. (2010) found that the employed reported more global life satisfaction than the unemployed, however affective responses to daily activities as measured by the DRM, did not differ by employment status. While the application of experience-based measures in maternal wellbeing is in its infancy, studies have also shown discordant results related to different wellbeing measures. Hoffenarr et al. (2010), who investigated maternal wellbeing across the transition to motherhood, found changes in episodic assessments of affect for certain daily activities which could not be identified by global measures of life satisfaction or depression. These results are in keeping with Newham and Martin's (2013) assertion that experience-based measures may outperform traditional instruments in their ability to capture dynamic changes in maternal wellbeing, whilst maximising ecological validity and limiting recall bias. These benefits have also created interest in the use of experience-based measures to evaluate health interventions (Dolan & White, 2007; Kahneman & Sugden, 2005), although such measures have yet to be integrated into the formal evaluation of a home visiting programme.

6.3 The Day in the Life of a *PFL* parent study

The Day in the Life (DTL) of a *PFL* Parent study sought to conduct a detailed examination of the potential programme impact on maternal wellbeing, by utilising a selection of measurement techniques. The study design allowed us to provide a detailed comparison of the programme impact across a number of wellbeing measures. To our knowledge this is the first attempt to compare experience-based wellbeing measures to more traditional measures of wellbeing in the context of a social intervention. In doing so, the study provides the most comprehensive study of home visiting programmes' impact on maternal wellbeing yet to be documented in the literature.

6.4 Method

PARTICIPANTS

Of the 233 *PFL* participants who were originally recruited into the programme during pregnancy, 192 (high treatment = 93; low treatment = 99) were eligible to participate in the DTL study as they had not voluntary or involuntarily (e.g., moved out of the catchment area) dropped out of the programme and/or evaluation. Participants who were scheduled to complete an interview during 2013 (n = 79) were invited to participate at the time of their interview and the remaining participants (n = 113) were invited separately by telephone. Of the 192 target participants, 102 mothers (high treatment = 46; low treatment = 56) participated in the DTL study. Thirty-four participants chose not to take part (the leading reported reason for refusal was a lack of time), two agreed but did not participate, and 54 could not be reached by telephone, letter, or

text. The DTL participants were at various stages in the programme when they completed the study, the youngest *PFL* child was 24.6 months and the oldest *PFL* child was 62.5 months old.

Participants who chose to take part in the DTL study did not differ from those who chose not to participate on 93% of the baseline characteristics collected (106/114). Significant differences on 7% of measures indicated that the DTL subsample was somewhat more disadvantaged than those *PFL* participants who did not take part. Specifically, a higher percentage of participating mothers reported being in receipt of social welfare payments, having their activity impaired by illness, consuming more drinks per week, availing of a greater number of certain services, meeting the risk cutoff for lack of empathy towards their child's needs [as per the AAPI (Bavolek & Keene, 2002)] and being more open [as per the TIPI (Gosling, Rentfrow, & Swann, 2003)].

Within the DTL sample, the high treatment group (n = 46) did not differ from the low treatment group (n = 56) on 87% (99/114) of baseline characteristics. To account for these differences, we control for six of the 15 significant variables in analyses examining treatment effects (the control set is composed of an emotional attachment score, a self-efficacy score, the number of neighbours known by the participant, whether or not the participant exercises at least three times per week, a community service use variable, and whether or not the participants' pregnancy was planned). In addition we control for the infant's gender, which was significantly different between the two groups at the twenty-four month interview and the length of time spent by participants in the programme by the time of the DTL study.

Table 6.1 presents descriptive information on the DTL sample based on information collected during the baseline interviews disaggregated by treatment status.

Table 6.1 Descriptive statistics regarding DTL participants' characteristics

	N^a_{HIGH}	N^a_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1
Age	101	(46/55)	26.00	(5.45)	25.35	(5.75)	ns
Child gender							
Male	101	(46/55)	0.48	(0.51)	0.31	(0.47)	$p<0.10$
Number of non-*PFL* children	101	(46/55)	1.00	(1.32)	1.05	(1.25)	ns
First time mother	101	(46/55)	0.50	(0.51)	0.47	(0.50)	ns
Weekly household income	86	(39/47)	705.13	(392.75)	735.64	(407.63)	ns
Married	101	(46/55)	0.17	(0.38)	0.16	(0.37)	ns
Work Status							
Employed	101	(46/55)	0.39	(0.49)	0.36	(0.49)	ns
Looking after family	101	(46/55)	0.13	(0.34)	0.13	(0.34)	ns
Unemployed	101	(46/55)	0.43	(0.50)	0.40	(0.50)	ns
Other	101	(46/55)	0.04	(0.21)	0.11	(0.31)	ns
Education							
Lower than second level education	101	(46/55)	0.41	(0.50)	0.44	(0.50)	ns
Second level education	101	(46/55)	0.20	(0.40)	0.25	(0.44)	ns
Primary degree/non-degree qualification	101	(46/55)	0.39	(0.49)	0.31	(0.47)	ns

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p-value* from an individual permutation test with 100,000 replications. [a] One participant did not complete a baseline interview.

These data show that high and low treatment mothers were largely equivalent on the majority of demographic indicators at baseline.

PROCEDURE

Data collection commenced in February 2013 and ended in November 2013. Participants were visited by a researcher on two occasions over a three day weekday period. At the first visit participants were given diaries and asked to record the next day's activities (study day). An interview was completed on the second visit which occurred immediately following the study day. The interview consisted of four parts: an adapted version of the DRM, global questions about life satisfaction and mood, and the PSI.

MEASURES

Adapted Day Reconstruction Method (DRM; Kahneman et al., 2004). The DRM was adapted for the DTL study based on the research question, literature review, and piloting. To assist the completion of the DRM, participants were asked to keep a diary of the study day broken down into episodes across the morning, afternoon, and evening. Participants used their diary as a prompt to describe each of the day's episodes in terms of the time it began and ended, the activity they were participating in, where they were, and who they were interacting with based on the categories displayed in Table 6.2. Participants were also asked to rate each episode in terms of 12 emotion states, displayed in Table 6.2, on a Likert scale ranging from 0-6 (not at all, very slightly, somewhat, moderately, quite strongly, strongly, and very strongly). Participants recorded an average of 11 episodes per day each which lasted 80 minutes on average.

Table 6.2 Day Reconstruction Method Episode Descriptors

Activity[a]			
grooming/care	exercising	attending training	paid work
preparing food	eating	housework	computer/email/internet
socialising	on the phone/skype	watching TV	relaxing
sleeping	commuting	shopping	taking care of child(ren)
playing with child(ren)	putting child(ren) to bed	feeding child(ren)	getting child(ren) dressed
other			
Location[a]			
home	work	on the road	elsewhere
Interaction[a]			
alone	*PFL* child	other child(ren)	spouse/partner
own parent(s)	partner's parent(s)	partner's child(ren)	partner's relatives
other people's child(ren)	friends	clients/customers	health professional
work colleagues	other relatives	other	
Positive emotion states			
happy	affectionate	competent	relaxed
in control			
Negative emotion states			
depressed	impatient	criticised	angry
frustrated	irritated	stressed	

Note. [a] participants could select as many activities/locations/interactions as applied.

The emotion scores can be analysed in a number of different ways. Individual emotion scores can be examined separately across the entire day and can also be averaged to create overall positive and negative scores, these are known as positive and negative affect respectively. The difference between positive and negative affect is also calculated to provide an overall measure of utility, this is known as net affect. Affect scores, as well as the individual emotions were weighted by episode length. This means that longer episodes count more towards an individual's overall emotional score than shorter episodes. In the DTL study, positive and negative affect and individual emotions were considered for the entire day and for episodes when the participant is with and without their *PFL* child only. We also generated a U-Index score (Kahneman & Krueger, 2006) to account for the likelihood that people will anchor themselves at different points along the Likert scale (e.g., an emotion of a given intensity may be scored as a 4 by one participant and as a 6 by another participant). The U-Index captures the proportion of time a participant spends in an unpleasant state. An episode was categorised as unpleasant if the highest rated emotional score was a negative one (i.e. impatient, frustrated, depressed, irritated, angry, stressed, criticised). The U-Index for a particular participant was calculated by weighting the U-Score for each episode (either 0 or 1) by the fraction of time that episode accounted for. This resulting figure represents the proportion of time during the day where a participant's strongest emotion was a negative one.

Global mood. Participants were also asked to indicate the percentage of time they spent in a bad mood, a little low or irritable, in a mildly pleasant mood, and in a very good mood in relation to the day overall and separately in terms of the time they spent with their child(ren). Binary mood variables (positive/negative) were created for the overall day and the time spent with child(ren).

Life satisfaction. Participants were asked to indicate the degree to which they were satisfied with their "life as a whole", "life at home", and their "life as a parent" on a 4-point Likert scale from very unsatisfied to very satisfied. Three binary satisfaction variables (satisfied/unsatisfied) were created.

Parenting Stress Index Short Form (PSI; Abidin, 1995). Participants self-completed a paper version of the PSI. The PSI consists of 36 items rated on a five-point Likert scale ranging from strongly disagree to strongly agree. The scale yields a total stress score (36 items, α=.94) and three subscale scores relating to parental stress – Parental Distress (12 items, α=.90), Parent-Child Dysfunctional Interaction (12 items, α=.90), and Difficult Child (12 items, α=.89). The Parental Distress subscale measures a parent's level of distress regarding her perception of her parenting role. The Parent-Child Dysfunctional Interaction subscale measures a parent's perception of the quality of her interactions with the child. The Difficult Child subscale measures a parent's perception of her child's behavioural characteristics. Responses were summed to generate scores for each of the subscales (scoring range 12 – 60) and the total stress score (scoring range 36 – 180). A binary variable was also created to represent mothers scoring above a cut-off of 90, indicating a high level of stress. In accordance with the manual, subdomain and total scores were not computed for participants who were missing data on more than one item on a given subscale.

6.5 Results

DESCRIPTIVE STATISTICS ON AFFECT MEASURES

The positive affect values were standardised for the entire sample to have a mean of zero and a standard deviation of one. Figure 6.2 illustrates the pattern of average positive affect over the course of the day for the high and low treatment groups. The high treatment group report higher positive affect scores at every hour on the graph, compared to the low treatment group. Figure 6.3 illustrates the pattern of average negative affect scores over the course of the day for the two groups. There is no clear difference in negative affect between the two groups - both the high and low treatment groups display a similar pattern over the course of the day. Potential group differences in positive and negative affect, averaged over the course of the day, are tested below.

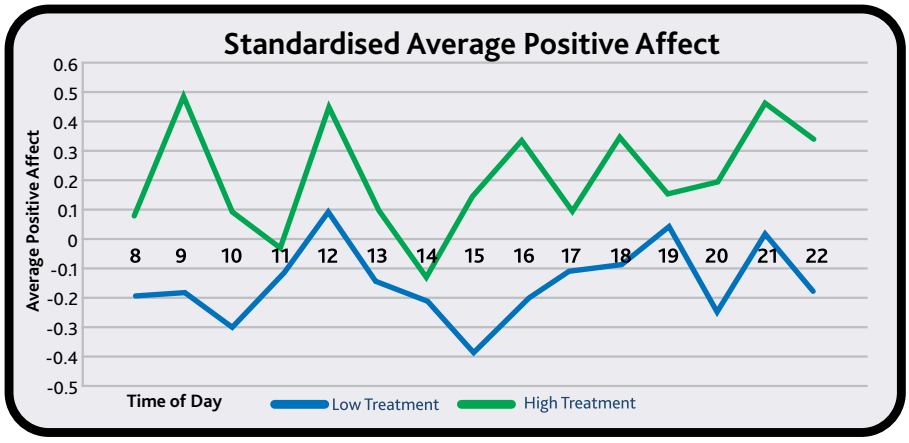


Figure 6.1 DTL Participants' Average Positive Affect across the Study Day

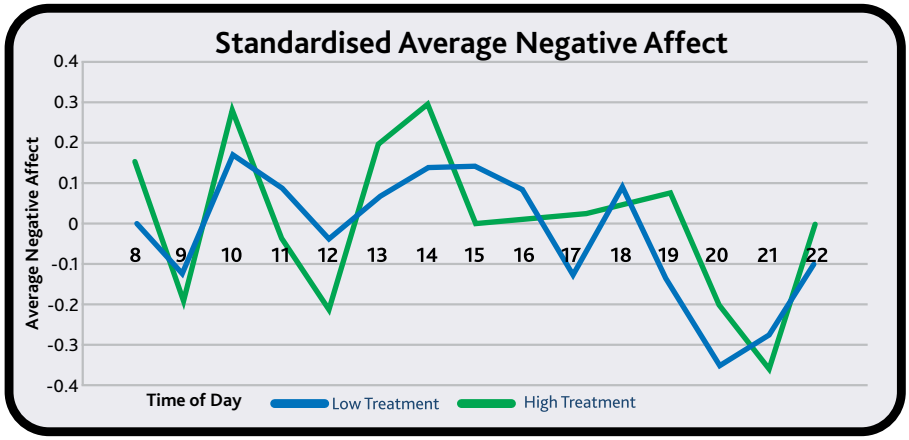


Figure 6.2 DTL Participants' Average Negative Affect across the Study Day

TESTING OF TREATMENT EFFECTS

All scores contained in Tables 6.3, 6.4, and 6.5 are weighted by episode length and encompass all episodes recorded. The p-values reported are based on two-tailed tests and are conditional upon the group differences discussed above. Table 6.3 compares the high and low treatment groups in terms of their positive affect and individual positive emotions. The day is considered as a whole and also disaggregated in terms of the time participants spent with and without their *PFL* child.

Table 6.3 compares the treatment and control groups in terms of their net affect and U-Index for the day as a whole and also time spent with and without the *PFL* child. On average, participants in both groups report a net affect score of approximately 3 over the course of the whole day. This implies that on average, participants reported experiencing positive emotions more strongly than negative emotions. Therefore, it is unsurprising that, both groups spend approximately only 10% of their day in an episode where the strongest experienced emotion is a negative one, as shown by the U-Index. Both groups experience a slight decline in net affect and a corresponding slight rise in the U-Index in episodes when they are without their *PFL* child. No significant treatment effects are identified for the three net affect or three U-Index measures.

Table 6.3 Net affect and U-Index Results for High and Low Treatment Groups

	N_{HIGH}	N_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1	*Step-down Test* p^2
Net Affect								
Net Affect	101	(46/55)	3.03	(1.41)	2.84	(1.37)	ns	-
Net Affect during time spent with *PFL* Child	101	(46/55)	2.98	(1.58)	2.95	(1.38)	ns	-
Net Affect during time spent without *PFL* child	96	(46/50)	3.00	(1.78)	2.68	(1.59)	ns	-
U-Index (range 0-1)								
* U-Index	101	(46/55)	0.10	(0.14)	0.09	(0.18)	ns	ns
* U-Index during time spend with *PFL* Child	101	(46/55)	0.10	(0.16)	0.08	(0.18)	ns	ns
* U-Index during time spend without *PFL* Child	96	(46/50)	0.11	(0.24)	0.12	(0.27)	ns	ns

Note. '*N*' indicates the sample size. '*M*' indicates the mean. '*SD*' indicates the standard deviation. [1] two-tailed p-value from an individual permutation test with 100,000 replications. [2] two-tailed p-value from a step-down permutation test with 100,000 replications. * indicates the variable was reverse coded for the testing procedure. * indicates that higher scores are worse. '*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.

Table 6.4 compares the high and low treatment groups in terms of their positive affect and individual positive emotions both for the entire day and the episodes when participants were with and without their *PFL* child. A treatment effect is identified for overall positive affect; however it is only significant for the time spent without the *PFL* child. The two groups do not significantly differ in terms of positive affect over the entire day or during episodes spent with their *PFL* child. The significant group difference is primarily driven by a decline in the low treatment group's positive affect during episodes in which they are not with their *PFL* child, while the high treatment group is slightly more stable in terms of positive affect during episodes with or without their *PFL* child. In terms of the individual positive affect states we find that high treatment participants report higher levels of happiness for the day overall and during times spent without the *PFL* child when compared with the low treatment group. The groups do not significantly differ on the remaining four positive affect states for the day overall or the time spent with the *PFL* child, compared to the low treatment group.

Table 6.4 Positive Affect Results for High and Low Treatment Groups

Positive Affect (range 0-6)	N_{HIGH}	N_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1	*Step-down Test* p^2
Overall								
Positive Affect	101	(46/55)	3.94	(0.96)	3.66	(0.95)	ns	ns
Positive Affect during time spend with *PFL* Child	101	(46/55)	3.97	(1.02)	3.77	(1.00)	ns	ns
Positive Affect during time spend without *PFL* child	96	(46/50)	3.84	(1.13)	3.48	(0.92)	*p<0.10*	ns
Positive Emotions								
Happy	101	(46/55)	4.03	(1.00)	3.59	(1.12)	*p<0.10*	ns
Affectionate	101	(46/55)	3.75	(1.49)	3.43	(1.38)	ns	ns
Competent	101	(46/55)	4.40	(1.04)	4.18	(1.12)	ns	ns
In Control	101	(46/55)	4.25	(1.16)	4.04	(1.19)	ns	ns
Relaxed	101	(46/55)	3.24	(1.16)	3.04	(1.16)	ns	ns
Positive Emotions during time spent with *PFL* child								
Happy	101	(46/55)	3.99	(1.22)	3.59	(1.17)	ns	ns
Affectionate	101	(46/55)	4.25	(1.42)	3.98	(1.40)	ns	ns
Competent	101	(46/55)	4.34	(1.09)	4.13	(1.22)	ns	ns
In Control	101	(46/55)	4.25	(1.20)	4.13	(1.17)	ns	ns
Relaxed	101	(46/55)	2.94	(1.34)	3.00	(1.21)	ns	ns
Positive Emotions during time spent without *PFL* child								
Happy	96	(46/50)	3.98	(1.07)	3.50	(1.25)	*p<0.10*	ns
Affectionate	96	(46/50)	3.08	(1.89)	2.57	(1.59)	ns	ns
Competent	96	(46/50)	4.31	(1.40)	4.16	(1.15)	ns	ns
In Control	96	(46/50)	4.17	(1.44)	4.00	(1.29)	ns	ns
Relaxed	96	(46/50)	3.67	(1.59)	3.18	(1.27)	ns	ns

Note. *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p-value* from an individual permutation test with 100,000 replications. [2] two-tailed *p-value* from a step-down permutation test with 100,000 replications.* indicates that higher scores are worse. *'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.

Table 6.5 compares the high and low treatment groups in terms of their negative affect and individual negative emotions both for the entire day and the episodes when participants were with and without their *PFL* child. Table 6.5 shows that there were no significant differences found in either the individual tests or the step-down tests.

Table 6.5 Negative Affect Results for High and Low Treatment Groups

Positive Affect (range 0-6)	N_{HIGH}	N_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1	*Step-down Test* p^2
Overall								
* Negative Affect	101	(46/55)	0.91	(0.79)	0.82	(0.76)	ns	ns
* Negative Affect during time spent with *PFL* child	101	(46/55)	0.98	(0.88)	0.82	(0.73)	ns	ns
* Negative Affect during time spent without *PFL* child	96	(46/50)	0.84	(0.97)	0.80	(0.92)	ns	ns
Negative Emotions								
* Stressed	101	(46/55)	1.47	(1.25)	1.24	(1.08)	ns	ns
* Irritated	101	(46/55)	1.29	(1.12)	1.08	(1.05)	ns	ns
* Frustrated	101	(46/55)	1.26	(1.02)	1.10	(1.00)	ns	ns
* Angry	101	(46/55)	0.66	(0.84)	0.55	(0.84)	ns	ns
* Impatient	101	(46/55)	1.27	(1.15)	1.32	(1.02)	ns	ns
* Depressed	101	(46/55)	0.23	(0.37)	0.28	(0.50)	ns	ns
* Criticized	101	(46/55)	0.18	(0.40)	0.16	(0.36)	ns	ns
Negative Emotions during time spent with *PFL* child								
* Stressed	101	(46/55)	1.61	(1.45)	1.25	(1.08)	ns	ns
* Irritated	101	(46/55)	1.36	(1.22)	1.04	(0.98)	ns	ns
* Frustrated	101	(46/55)	1.37	(1.19)	1.11	(1.00)	ns	ns
* Angry	101	(46/55)	0.66	(0.87)	0.56	(0.85)	ns	ns
* Impatient	101	(46/55)	1.43	(1.26)	1.36	(1.09)	ns	ns
* Depressed	101	(46/55)	0.24	(0.53)	0.24	(0.49)	ns	ns
* Criticised	101	(46/55)	0.22	(0.49)	0.17	(0.39)	ns	ns
Negative Emotions during time spent without *PFL* child								
* Stressed	96	(46/50)	1.36	(1.61)	1.23	(1.31)	ns	ns
* Irritated	96	(46/50)	1.16	(1.38)	1.03	(1.33)	ns	ns
* Frustrated	96	(46/50)	1.10	(1.31)	1.07	(1.29)	ns	ns
* Angry	96	(46/50)	0.70	(1.21)	0.58	(1.15)	ns	ns
* Impatient	96	(46/50)	1.15	(1.46)	1.12	(1.29)	ns	ns
* Depressed	96	(46/50)	0.26	(0.57)	0.44	(0.91)	ns	ns
* Criticised	96	(46/50)	0.14	(0.58)	0.13	(0.34)	ns	ns

Note. *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p-value* from an individual permutation test with 100,000 replications. [2] two-tailed *p-value* from a step-down permutation test with 100,000 replications.* indicates that higher scores are worse. *'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.

Table 6.6 compares the high and low treatment group in terms of their global mood and life satisfaction. It shows that members of the high treatment group reported spending a higher proportion of the DTL day in a positive mood. This result is significant in the individual and step-down test. There was no significant difference for the portion of time spent with children in a positive mood. Also, there were no significant differences found in terms of life satisfaction, either overall, at home, or as a parent.

Table 6.6 Global Mood and Life Satisfaction for the High and Low Treatment Groups

Positive Affect (range 0-6)	N_{HIGH}	N_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1	*Step-down Test* p^2
Mood								
Portion of Day Spent in a Positive Mood	99	(45/54)	0.76	(0.18)	0.71	(0.25)	*p<0.05*	*p<0.10*
Portion of Time Spent with Children in a Positive Mood	101	(46/55)	0.83	(0.21)	0.84	(0.19)	ns	ns
Life Satisfaction								
Satisfaction with Life as a Parent	100	(45/55)	0.98	(0.15)	0.89	(0.31)	ns	ns
Satisfaction with Home Life	100	(45/55)	0.96	(0.21)	0.89	(0.31)	ns	ns
Satisfaction with Life Overall	100	(45/55)	0.93	(0.25)	0.89	(0.31)	ns	ns

Note. *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p-value* from an individual permutation test with 100,000 replications. [2] two-tailed *p-value* from a step-down permutation test with 100,000 replications.* indicates that higher scores are worse. *'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.

Finally, Table 6.7 compares the high and low treatment group in terms their PSI scores and shows that there were no significant differences for any of the five PSI scores.

Table 6.7 Parenting Stress Index Results for the High and Low Treatment Groups

Parenting Stress Index	N_{HIGH}	N_{LOW}	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*Individual Test* p^1	*Step-down Test* p^2
PSI subdomains								
* Parent-Child Dysfunctional Interactions[3]	99	(45/54)	18.04	(5.44)	17.23	(5.40)	ns	ns
* Difficult Child[4]	94	(43/51)	22.42	(8.34)	22.18	(7.03)	ns	ns
* Parental Distress[5]	100	(45/55)	24.82	(8.39)	24.67	(8.50)	ns	ns
Non-Step-down								
* Total Parental Stress[6]	93	(42/51)	64.52	(18.17)	64.02	(17.95)	ns	-
* Stress Cut-off[6]	93	(42/51)	0.10	(0.30)	0.08	(0.27)	ns	-

Note. *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. [1] two-tailed *p-value* from an individual permutation test with 100,000 replications. [2] two-tailed *p-value* from a step-down permutation test with 100,000 replications. [3] One participant was excluded due to missing data. [4] Two participants were excluded due to missing data. [5] Seven participants were excluded due to missing data. [6] Eight participants were excluded due to missing data. * indicates that higher scores are worse. *'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.

6.6 Summary of Results

The DTL study shows that high treatment mothers reported higher levels of experienced positive emotion than the low treatment group, for times when they were without their *PFL* child. This was broadly consistent with mothers' global judgements for their overall levels of positive mood on the study day, with a positive treatment effect observed for the overall day, yet not time participants spent with their children. In keeping with the existing home visiting literature there were no treatment effects for mothers' negative emotions, in terms of both their experienced negative emotions across the study day or their general ratings of parenting stress as measured by a standardised instrument. Lastly, there were no treatment effects in terms of participants' life satisfaction. Taken together these results suggest that the DRM provides divergent results to those observed by traditional techniques. In this way the DRM appears to yield more nuanced insight into participants' day-to-day emotional wellbeing, which is different to their judgements about their general life circumstances. In our study this allows us to detect a treatment effect for mothers' positive affect, which is restricted to the time they spend without their *PFL* child. Overall, *PFL* appears to be generating an impact on some aspects of participants' wellbeing, however, not others. These results will be further interpreted in Chapter 8.

Chapter Seven

Father Figure Focus Groups

7.1 Introduction

Fathers play a crucial role in child development (Paquette, 2004; Caldera & Lindsay, 2006). Similar to mothers, they can directly and indirectly influence their children's development, through facilitating social engagement and promoting healthy psychological and cognitive development (Caldera & Lindsay, 2006). Yet they also provide a unique style of parenting, separate to mothers, which can foster development. Specifically, through encouraging curiosity, exploration and risk-taking, while ensuring the child feels safe and secure, fathers can encourage children's "openness to the world" (Paquette, 2004). In addition to biological fathers, support from non-biological father figures can yield similar positive effects, facilitating the development of social competence, improved behaviour and academic achievement, and helping to reduce depressive symptoms in children (Coley, 1998; Dubowitz et al., 2001). Father-child interactions are influenced by a multitude of factors, including but not limited to, their relationship with the child's mother, the family context, human and financial resources, and work stress (Bonney et al, 1999; Lamb, 2004; Repetti, 1994).

Father engagement in home visiting programmes is typically low, and is affected by child, father and mother characteristics (Roggman et al, 2002; Raikes et al, 2005). Indeed, there is evidence to suggest that although programmes may aim to adopt a family approach, home visitors often naturally target mothers, excluding fathers even when they are living in the home (Duggan, Fuddy et al., 2004). Nonetheless, there is evidence to suggest that, when encouraged, fathers can engage with parenting programmes to the same extent as mothers (Bagner & Eyberg, 2003). To this end, a number of home visiting programmes in the US have begun to develop strategies specifically to engage fathers. These strategies are mostly practical, such as using the father's name on correspondence, arranging visits in the evenings and at weekends, employing male staff to deliver the programme, and meeting the fathers outside of the home. They also include explaining the importance of father engagement to mothers and tailoring the programme to suit fathers' unique parenting skills (Circle of Parents, 2011). However, to date there is still a relative dearth of literature on father participation in home visiting programmes (Duggan, Fuddy et al., 2004). Thus, the aim of this qualitative element of the *PFL* evaluation was to investigate the experience of fatherhood among the *PFL* community, and to determine factors influencing fathers involvement in the *PFL* programme.

7.2 Method

PARTICIPANTS

In total, three focus groups and two semi-structured interviews took place with high and low treatment fathers and father figures between November 2013 and January 2014 (nHigh=6/nLow=4). In consultation with the *PFL* programme implementation team, the study included father figures in addition to biological fathers, due to the demographic profile of the community. At thirty-six months, roughly 70% of *PFL* mothers were in partnered relationships, 19% of whom were married. In most cases, the partner was the child's biological father (High=96%/Low=89%). Father figures were described to participants as "male role models", and could include any male over the age of 18 who had a father-like relationship with the child, such as a stepfather, grandfather, uncle, cousin, brother, or any other adult male known to the mother and child. For the purpose of this chapter, all participants will be referred to as "father figures" hereafter, regardless of their biological connection to the child.

In consultation with the *PFL* implementation team and other local service providers, the decision was taken to recruit the father figure participants through the *PFL* mothers. Mothers in the high and low treatment groups (n=182) were sent a flyer about the father figure study and were asked to pass on another enclosed flyer to any father figures in the child's life. This was followed up with a phone call to the *PFL* mothers whose families had not chosen to opt out of the father figure study by text (n=181). 72 mothers opted out over the phone on behalf of the father figures. Typical reasons given included that there was no father figure in the child's life, that the father figure would not be interested in taking part, or that he was busy

with work and would not have time to attend. Through the mothers, 16 father figures agreed to attend, and the remainder of mothers were either unreachable or did not respond to repeated contact attempts. Of the 16 who agreed to attend, 10 presented at the agreed times. In cases where only two participants presented, they were offered a choice of cancelling, proceeding with the focus group, or taking part in individual semi-structured interviews. Regardless of which option they selected, participants received a €20 shopping voucher as compensation for their time.

The 10 participants consisted of six fathers and four father figures from differing generations, who were all related to the children either biologically or through marriage. There were six participants from the high treatment group and four from the low treatment group. The majority of participants were of Irish nationality, with two international participants. Participants ranged in age from 19-63 (n=2, age missing) and the average age of the *PFL* children to whom the participants were linked was 44 months.

All focus groups were recorded digitally, transcribed verbatim and analysed thematically by a team of three researchers using the guidelines set out by Braun & Clarke (2006). Initial coding was conducted by hand, before the transcripts were uploaded to nVivo10 for more in-depth analysis. Data for the high and low treatment groups were first analysed independently, however, as no clear differences emerged between the groups, the data were combined and a set of overarching themes was developed.

7.3 Results

The data analysis resulted in six main themes, which are presented in Table 7.1 and are outlined in more detail hereunder.

Participants' overall perception of an involved father figure was a strong, positive male influence that nurtures and helps in the raising of the child, regardless of the biological relationship. This role was seen as complex and multifaceted, and both treatment groups agreed that the ideal father figure should be present, understanding, supportive, caring, a good listener and a moral guide who will teach their child necessary life skills. Knowledge of, and involvement in the *PFL* programme was low among both treatment groups, yet the father figures were supportive of *PFL* and desired closer involvement with the programme and parenting support more generally.

Table 7.1 Focus Group Themes & Sub-Themes

	THEME	SUB-THEMES
1.	A familial learning curve	• Father figure growth and development • Child growth and development • Father figure-assisted child development
2.	Role embodiment	• Multi-faceted role • How to be a father figure • Perceptions of father figures
3.	The ecology of the father figure role	• Personal influences on fatherhood • Mother-father figure relationship • Combination families • Society and societal change
4.	Barriers to being a father figure	• Work • Self-care • Mothers as gatekeepers
5.	Future Focus	• Hopes for child • Challenges to expectations • Achieving aspirations
6.	*PFL*	• Knowledge and involvement • *PFL* Perceptions • *PFL* Impact • *PFL* Improvements

THEME 1: A FAMILIAL LEARNING CURVE

This theme concerned the transition to becoming a father figure, and how the dynamics of the father figure-child relationship changed according to the child's age and stage of development.

SUB-THEME A) FATHER FIGURE GROWTH AND DEVELOPMENT

The first subtheme focused on the growth process father figures experience, from learning about their new role to adapting to their responsibilities and overcoming any challenges that may arise. Participants emphasised how becoming a father figure cannot simply be 'learned from a book' but that it is a long-term learning process:

> *You just pick it up…there's no guide anyway there's no, no college to help you become a dad so you just do, do as you do and hope for the best*
>
> Low treatment participant

One participant from the high treatment group stated that in order for some father figures to settle into their roles, they felt that they had to perform to the best of their ability without doing or saying anything wrong. Similarly, he suggested that they did not want to be told if they made a mistake. He, on the other hand, saw making mistakes as a positive joint growth process whereby 'you learn and the child learns':

> *People don't wanna know that, if you're doing the wrong thing or not… it's like a macho thing that you're doing things right and not things wrong and you don't wanna say the wrong things*
>
> High treatment participant

Since becoming father figures, participants reported that a shift in priorities had led to the adoption of a less self-centred role, whereby their needs became secondary to those of their children:

> *In the past I was just looking after myself, but now it's it's the child first[…]I look after her first*
>
> High treatment participant

> *It is hard[…]for 15/20 years you're just looking after you, you know? And then all of a sudden someone comes along and it's totally them for the next 20 years so, it's a tough mental thing to em regard that this is your child, you have to look after your child and that's it*
>
> High treatment participant

Another high treatment participant focused on the communication difficulties he was now facing as his children were growing up in terms of how much information to vet and protect them against. He had developed quite a protective role as his children had grown older and found it difficult to set boundaries with them.

> *Certain ages it's very hard to communicate with them you see because when they start asking questions and, 'why not why, I can go on my own'…if you explain too much for them as well you start to be worried…they kidnap kids and also you can't go when it be dark, you have to stay close to the house like inside the gate not outside[…]honestly I feel difficult like to explain exactly what's going on*
>
> High treatment participant

SUB-THEME B) CHILD GROWTH AND DEVELOPMENT

The next sub-theme involved participants' understanding of child growth and development. The main focus in both treatment groups was on their children's academic progress. Participants discussed the transition period their children had to go through in beginning crèche and their academic education. They also described the enjoyment their children got from school, especially the recognition and rewards they received for good behaviour and meeting targets. Coming to terms with the rules, discipline and expectations placed upon them by their teachers was seen as a developmental milestone for children:

> *School is a big, it's the next step[…]it's gonna be a big jolt for her so, but she loves school… which is a positive*
>
> Low treatment participant

> *They'll be delighted to explain to [mentor] what's happened them in the school and what they get like, you know, stickers or whatever*
>
> High treatment participant

High treatment participants also commented on their children's intelligence levels and their ability to absorb information from different sources of technology:

> *His brain is like a sponge, he's a very, very intelligent child, he takes everything in…it's brilliant for him to have all this knowledge*
>
> High treatment participant

> *My young one has more or less learned everything off Mickey Mouse… she's learned everything, she counts in [language] she does… she's gone beyond her age, yeah she has more brains for a 4 year old*
>
> High treatment participant

SUB-THEME C) FATHER FIGURE ASSISTED CHILD GROWTH AND DEVELOPMENT

When asked how they assisted child growth and development, the majority of participants stressed that they were very much concerned with helping their children in the adjustment process during different transition periods such as beginning school. One participant from the high treatment group explained that if his child was not progressing developmentally then the blame would lie with him. Seeing her enjoy and progress through school was validation for him that he was 'doing a good job' as a father figure:

> *She is at the stage where she is progressing at a steady rate and not being held back by me or by my partner, because if she's not ready for school, that's down to me, that's my fault*
>
> High treatment participant

He also explained how he was able to assist his child in *PFL* by adapting and tailoring the tools, knowledge and skillset the programme had provided his family with to specifically suit the needs of his child:

> *You just have to take your bits out that will work for you... because eh, every household is different so we do have to tailor to what suits [child]*
>
> High treatment participant

In addition, father figures from both treatment groups described how a more family-oriented role model approach was adopted in their household when raising their children. Older siblings were seen to actively enhance the contribution of parent–child relationship processes and mitigate various burdens and stressors placed on the parents by contributing to some of the parenting responsibilities and caregiving duties:

> *I think when you have more kids like in the family you know, the youngest always should […] be taking from the older sisters or brothers, it will be easy for you*
>
> High treatment participant

> *They all take part in the child's upbringing, in a sense, when they're there*
>
> Low treatment participant

Participants from the high treatment group also talked about the challenges they faced at different stages of child development. One participant from the high treatment group emphasised difficulties with his older children. He described the effect of living with teenagers as 'carnage' and how a breakdown in communication with them had caused a tense environment for the whole family:

> *I've no problems with my four year old whatsoever, it sounds stupid but I haven't...it's just the teenagers[…] there's loads of friction going on in the house with the whole family that starts to get hard, you don't get any time together*
>
> High treatment participant

THEME 2: ROLE EMBODIMENT

The second theme concerned the multidimensionality of the father figure role, how participants viewed themselves and felt they were viewed by society.

SUB-THEME A) MULTIFACETED ROLE

It was evident from the narratives that being a father figure required the embodiment of a number of roles simultaneously. Both treatment groups spoke about a number of key roles included being a playmate, disciplinarian, carer, supporter, protector, positive re-inforcer, and a teacher of life lessons. In addition, the low treatment group identified novel roles of story teller, listener and moral guide, while the high treatment group highlighted their roles as a reasoner, educator, motivator, and encourager.

SUB-THEME B) HOW TO BE A FATHER FIGURE

When asked about how to be a father figure, participants typically described themselves as being involved in the routine, care, and nurturing of their child, and as being supportive, a good listener and approachable. Many participants felt that creating structure and routine was a crucial part of fostering child development, and they were very much involved in this process:

> *Just being there for her in general…I'm always there for her if she needs me and I'm just there for a smile and open arms for her*
>
> High treatment participant

> *It's all about routine. If we have a routine in place, [child] will benefit from the routine because she knows step by step what way, how the day goes and so she's not, you know, all flustered or confused*
>
> High treatment participant

All participants identified play as a major part of their role, with one commenting that:

> *Every single evening is, I play with the kids from the second I'm home til they go to bed and the weekends*
>
> High treatment participant

One participant from the high treatment group also discussed at length the importance of different types of play used to forge bonds of trust within their relationship. In particular he focused upon 'rough and tumble play' and the numerous benefits generated by this type of play interaction:

> *Well I'm a I'm a big thing of rough, rough play…the tumbling and the wrestling and all, because she needs that kind of, that bonding em with the father[…] dads have more of a risk factor than the mother. I think that they [mothers] don't want to see their child hurting themself but, ye do have to hurt yourself to know how to do things and how to fall properly[…] that's what your dad can teach you*
>
> High treatment participant

Activities, such as outdoor trips, represented an important part of time the father figures spent with their children:

> *I just normally take them to the park or bring them to the shops or, eh, go on the bike while my daughter is trying to cycle*
>
> Low treatment participant

There was an appreciation within both treatment groups of the many benefits that positive role play scenarios and play facilitation can bring. They emphasised the importance of letting the child take the lead in these situations and highlighted the significant implications this can have on child development:

> *If you're doing something [with] them get on their level…have an interest in what they're doing, let them come up with ideas of what needs to be played or something like that and not what you want to do*
>
> High treatment participant

I would role play that in a sense in a positive way…I say well "[child] is not to be sayin' this you know because that's not nice"…you're picking up on this then you know, and as I say this is an opportunity

Low treatment participant

Although they may not necessarily be living in the family home, participants from both treatment groups described how a father figure should still be present and involved in their child's life and sensitive to their needs, which was why the dedication of time was important:

I think time you have to give the time[…] I have other children as well […] when I was working, the difference of time you have with [child] now compared to the others so I think time is a big thing

High treatment participant

I'm mad into fishing and sports, but I'm also mad into wildlife, I'll sit there and watch documentaries on wildlife, and he'll sit there and watch them, and what I love about him is he'll ask questions

High treatment participant

Encouragement and support of extra-curricular activities was another way participants from both treatment groups felt they could assist in child development. They actively sought to get their children involved in a hobby which would facilitate opportunities for peer interaction and the development of their social skills:

Could we get them into some sport or, you know, the scouts where they can continue socialising[…] with the development of social skills

Low treatment participant

Participants said they felt it was their duty to be teachers and moral guides for their children. Father figures in both treatment groups made reference to the fact that rewarding children's bad behaviour and appeasing their demands instead of taking the time to explain the consequences of their actions was unacceptable. They recognised that, at times, enforcement of discipline was required and they cited a number of strategies which included withdrawal of privileges and time outs:

You just have to be firm and say "listen", explain to her again, get down on her level and explain why she can't have it and give her a reason so she understands

High treatment participant

One high treatment participant discussed how being a disciplinarian was not a role he had necessarily chosen but one he had to assume, and at times did not want:

He'll cry and call for me, and like I have to be the bad guy, and I'd have to say no you can't, listen to [mother] and that like kind of hurts a bit, cause I want that freedom to do what I want to do

High treatment participant

SUB-THEME C) PERCEPTIONS OF FATHER FIGURES

Father figures' perceived obligations and role expectations reflected their self-initiated efforts to meaningfully contribute to family life and the running of the household, and to complement the caregiving role of the child's mother. Reference was made to the sociocultural transformations and revised expectations of them as father figures, expanding their roles from that of a breadwinner to providing love and support:

> *Management of the house and the cleaning of it[…] there should be a natural awareness kind of, of you can be a role model*
>
> Low treatment participant

The same participant related this back to a change in societal perceptions, remarking that in the past:

> *You were the father figure in that sense, and you provided, and you didn't provide care or […] emotional support or anything for the kids it was just get out and work, bring the money in […] now in this generation, I feel that the males have more of a role in the child's development*
>
> Low treatment participant

Despite this, participants from both treatment groups stressed the misconceptions and underestimations that still surround their role. Participants from the low treatment group emphasised the need for more help and societal support:

> *So many men are pushed out of the way so to speak […] 'how could he cook, how could he clean?*
>
> Low treatment participant

One high treatment participant noted that while some things are beginning to change, the attitude of society has not:

> *Even though you know in the modern world and, we all try to be equal, but it's never gonna happen*
>
> High treatment participant

THEME 3: ECOLOGY OF THE ROLE

This theme encompassed the greater familial, relational and societal subsystems that impact the father figure role, including their own father figure influences growing up, the mother-father figure relationship, combination families, and society and societal change:

SUB-THEME A) PERSONAL INFLUENCES

Participants described the mechanisms through which they developed expectations about their roles within society and family. A number made reference to their own fathers' practices as a basis for much of their own father figure behaviours:

> *I always thought I was going to be different than my own dad when I was younger…I didn't really like him that much. Now when I look back I kinda think oh, he done an alright job and I'm turning into my dad, which is not necessarily a bad thing*
>
> Low treatment participant

Another participant mentioned choosing to be involved in the upbringing of the child as when he was young he 'didn't really have a [father figure] to look up to'. References were also made to polarised feminine and masculine identities and clearly defined gender roles that most of the participants had experienced growing up. Participants referred to their mothers as their main point of contact as a child and used words such as 'breadwinner' and 'provider' to describe their fathers:

> *My father wasn't really a role model in that way because he was more work, home, work […] things have changed now[…] you have to be there for your kids*
>
> High treatment participant

One low treatment participant recalled how whenever he asked his father's permission for something, he was told 'ask your mother'. He disagreed with this style of parenting:

> *The da is the parent as well, he should have the exact same authority or or veto, as no is a no and yes is a yes*
>
> Low treatment participant

SUB-THEME B) MOTHER-FATHER FIGURE RELATIONSHIP

The second subtheme concerned the mother-father figure relationship and the perceived relevance, need for, and importance of cohesive parenting practices. Participants from both treatment groups valued a united style of parenting, where the mother and father figure were equal partners in the upbringing of their children:

> *The da does the fun stuff as well as the bad stuff, the ma does the fun stuff as well and if we discipline, we both discipline*
>
> Low treatment participant

> *They'll probably get a better understanding of what a collective support and unit basically, kind of the, the parents are*
>
> High treatment participant

Participants described believing in routine and a consistency in their discipline approach as they and the child's mother worked together as a united front, referring to phrases such as 'set the same structure', 'stay in the same cycle' and 'do the same things'. One participant from the low treatment group stated that although he and the child's mother often had differing views and sometimes argued when it came to matters concerning the children, they ultimately compromised. He recognised the emotional implications that fighting in front of children could have, and so any disputes that did occur were settled in private to avoid children sensing the conflict or using the arguments they heard against one or other parent. Tantrums were identified as one of the challenges of being a father figure, and participants agreed that not giving in to children's demands and both parents adopting a clear and consistent approach was central to handling this behaviour effectively:

> *[child] will be on basically on the straight, straight road instead of deviating from one parent to the next parent to try and get what they want*
>
> High treatment participant

> *If [child] knows we're both doing the exact same...[child] will know if we say no it means no or yes it means yes*
>
> High treatment participant

Father figure involvement was seen as a benefit for the child by the majority of participants. They felt that closer involvement in the child's development led to more joint decision making regarding how best to help the child progress. This, in turn, would be beneficial to the child in the long run:

> *I feel that now, there's more [fathers] that are involved in their child's growing up, so you'll probably see a better developed child in the long run*
>
> Low treatment participant

One participant from the high treatment group highlighted how conflicting parenting styles can be detrimental to child development and emotional outcomes. He felt that inconsistency in discipline gave conflicting messages to his child and as a result the child was acting up and hitting out. Participants from both groups agreed it was important for both the mother and father figure to be clear in their approach to the child, using phrases such as 'mixed up', 'picking up different vibes' and 'not really knowing which one is right' to demonstrate the ramifications of children receiving mixed messages:

> *Everything I've taught him when he's up with us, goes out the window on the way home in the car cos everything at home is completely different […] like you're chastising or trying to correct him for what he's done wrong, [mother]'s stepping in, so he doesn't know which is right and which is wrong*
>
> High treatment participant

SUB-THEME C) COMBINATION FAMILIES

The next sub-theme involved combination families, Having a number of people from different families living together in the one home was believed to influence the family dynamic and psychological wellbeing of those in the house, which itself impacted the father figure role. A low treatment participant discussed the emotional implications of the challenges of raising teenagers, while a high treatment participant focused on the difficulties that arose when integrating into a new family compared to his previous relationship:

> *[referring to a previous situation] it was only me and their mother that was involved, that was very very easy, now that I've come into [wife] and [wife's children]'s lives right, I'm finding it hard to teach them*
>
> High treatment participant

SUB-THEME D) SOCIETY AND SOCIETAL CHANGE

The final subtheme concerned society's influence on the father figure role. While low treatment participants described societal perceptions of father figures, the high treatment group explored the impact societal change has had upon children. They discussed the positive and negative effects of recent technological advances on their children. One participant mentioned how his child had learned a new language from watching television, while another commented on how the electronics of today prevented his child from participating in the outdoor play that he had experienced when growing up. High treatment participants described perceiving society as less safe than in the past, pinpointing the many dangers facing their children whilst referring back to their need to assume a protective role.

> *She does watch far too much TV in fairness because where where we live…if I let my young one out and she'd be around the corner and, I'd be constantly out, you'd have to be with her*
>
> High treatment participant

> *She wants to explore new things, even if there's a danger involved and you just have to explain to her that no it is dangerous and explain what could happen*
>
> High treatment participant

THEME 4: BARRIERS TO BEING A FATHER FIGURE

This theme examined the barriers that prevented or restricted father figures from being more involved with their children. The main barriers included work, time, gatekeeping and self-care. The words/themes that emerged most frequently within the transcripts regarding the barriers to be a father figure were work and time.

SUB-THEME A) WORK

A number of father figures in the high treatment group discussed the challenges that arose from being involved with the children and participating in the *PFL* programme whilst continuing to meet work and family demands. One father provided an example of this perspective:

> *I'm full time in work so now... I have the kinda four hour window where I can kinda go home and play with her*
>
> High treatment participant

> *I'm not really that involved because I'm in work...the mentor comes most times when I'm in work […] in the past year [mentor] kind of waited til I came home to actually to come round, just so I could have more of an input*
>
> High treatment participant

The narratives appeared to suggest that there was a vicious circle facing low-income father figures, whereby their perceived need to provide financially for their families meant that working long hours was a necessity, however, these long hours precluded them from spending time with their children. A lack of time appeared to be a major impediment to child involvement, and some low treatment participants compared their current status as father figures to their experiences with previous children:

> *I worked nights and days and I mightn't see the kids ... I got home, they'd be in bed maybe*
>
> Low treatment participant

> *I'd be getting up, going out before they get up for school and then coming home in the evening, they'd be going to bed then, I wouldn't see them then*
>
> Low treatment participant

Additionally, low treatment participants highlighted household management and the stressors of adult daily life as barriers to child involvement. They placed particular emphasis on financially providing for their families as a major inhibitor to child involvement and almost all made reference to the obligation they felt to provide financial support. Changing circumstances in the form of unemployment and retirement afforded some father figures greater opportunities to spend more quality time with their children than they had previously, which to them was a positive development:

> *Past two years, I really enjoyed being at home with you know, the[children]... and really getting to know them you know, over and over again*
>
> Low treatment participant

SUB-THEME B) SELF-CARE

Participants from both treatment groups referred to the fact that due to family demands they had little time for self-care. One participant from the high treatment group commented that his desire for a little bit of 'me time' was not often met. At times they felt a sense of inescapability from their role as their family's requirements came to the forefront, superseding their own needs.

> *You always have to be a role model […]sometimes you come home from work and you just want to lie down or relax[…]have that hour for yourself but[…] it's the kids always first and then it's [Mother] and then it's me so you just have to put it back to the background*
>
> Low treatment participant

One participant from the high treatment group described feeling guilty for taking some time to put his needs ahead of his children's:

> *Sometimes you could really just do with a bit more time if you promise them something…then you don't get to do it…then you're feeling, aw Jesus if I had of had a half an hour later on I could have done it with them…instead of sitting there having a cup of tea*
>
> High treatment participant

SUB-THEME C) MOTHERS AS GATEKEEPERS

Another barrier to father figures' involvement with children was mothers acting as gatekeepers, whereby father figure involvement with the child and visitation privileges were restricted or completely withheld by the child's mother. Depending on the status of their relationship with the mother, some participants only had access to their children at weekends, while others had unrestricted access. Participants highlighted the lack of rights father figures currently have within the judicial system, and described how children were often used as 'ammunition' and 'emotional blackmail' by their mothers:

> *We don't really have any rights or anything like that so as we progress, we'll probably get more rights to the child and stuff like that, which probably most fathers want in a way*
>
> High treatment participant

> *I don't think it's perceived as really important…I'm having huge difficulties in, you know, just with court access an' all that*
>
> Low treatment participant

Low treatment participants discussed the struggle which many father figures in society face in trying to gain access to their children after separating from the children's mothers. Some were only permitted to see their children under supervision, and this was perceived as having detrimental effects on their relationship with the child. One participant described the anguish and emotional ramifications that a mother's refusal to maintain the visitation agreement had, not only on the child, but also on the wider family:

> *It's emotionally impacting on the child…it involves everybody because it is real hard on the emotion*
>
> Low treatment participant

THEME 5: FUTURE FOCUS

The father figures exhibited strong, well thought-out feelings about the future, including long term goals for their children, potential barriers to their children achieving these goals and how these challenges could be overcome.

SUB-THEME A) HOPES FOR THE CHILD

The unanimous consensus among all father figures was that they wanted their children to succeed and have a better life than they did, both academically and vocationally. They wanted them to pursue their dreams without doubt or fear and to hold onto the belief that anything was possible:

> *I'd like them to have a better life and better education and …get a good job out of it…and be happy all the time in what they're doing*
>
> Low treatment participant

Tell her nothing, nothing is gonna get in the way of her dreams…keep going, you know? You'll get there eventually

High treatment participant

Other hopes included the children being able to afford a comfortable lifestyle and having relationships and children of their own. While some participants focused on the importance of third-level education, one participant explained how he had spent a long time studying, only to discover upon graduating that there was no longer a need for his profession. Thus, his hopes for his children were focused on them pursuing and developing their hobbies and being flexible enough to adapt to and thrive in an ever-changing society. He felt it was best not to make a set long term plan as it is difficult to predict what the future will hold:

I hope she grows up to be a successful girl…have a good strong relationship with whoever she meets in life and, and has kids of her own one day if possible and gets married. Anything's possible… you have to keep the ability for the change and for the new IT or the new courses coming up… so you can't have long term plan

High treatment participant

Participants also described goals for the emotional wellbeing of their children using words such as 'happy', 'healthy' and 'confident' to illustrate this. Interestingly, one participant made reference that the fact that even attempting to chase your dream is, in itself, a challenge for children:

Just live your dream, or chase your dream at least. An that's, of course that's challenging. Cause hardly anyone does that really

Low treatment participant

All participants stressed that whatever their children decided to do, they should be fully committed and give it their all. They wanted them to take ownership of their lives and have the drive to get what they want out of life, knowing that anything is possible if they put their mind to it.

SUB-THEME B) CHALLENGES TO EXPECTATIONS

When asked about obstacles that may impede children from achieving their dreams, financial means was the main challenge identified. However, participants did not want this to deter their children from pursuing their dreams and their narratives portrayed them as having a positive 'can do' attitude:

The only thing is, if my daughter is gonna say "oh I want to be an open heart surgeon" that's what, twenty grand a year or something? That's not the kinda cash that I have in my pocket but surely if there's a will there's a way

High treatment participant

We have to make it clear to to to our kids… that they are allowed to have a dream and that they can chase that dream and don't be restricted by, by any financial means

Low treatment participant

Other challenges identified included outside influences such as peer pressure, bullying and friendship groups, a changing environment and limited job market, and clashing parenting styles which may cause a lack of continuity in the child's life.

SUB-THEME C) ACHIEVING ASPIRATIONS

Participants were asked about how these challenges could be overcome in order to help their children succeed. Money, health and stability were words used recurrently by participants from both treatment groups. One high treatment father figure stressed that he himself would play an important role in helping his child:

> *I'll do anything[…] I'll move house to try and get her the[…]education she wants so if she wants to go further or, I'll just be there to help her*
>
> High treatment participant

Another participant felt that stability in the home could be created by establishing and adhering to a routine whereby quality time was dedicated to both play and studies. Low treatment participants also suggested that when dealing with bullying, for example, keeping in regular touch with children's teachers and providing children with extra-curricular activities as an alternative social outlet from school were the best possible solutions. As well as financial stability, emotional stability was also seen as extremely important by many participants:

> *We'll always be there for her, you know? Just so long as she knows that […] she'll have that stability growing up and she can always come and talk to us*
>
> High treatment participant

> *Spend the time with him right……then you stop, you do his homework with him no matter what…at a certain time he's to be in bed. You have a routine with him, he'll get used to this, and you'll have that right through his school years.*
>
> High treatment participant

Ultimately, participants were in agreement that despite what they wanted, it would be their children who made the final decision about their own lives. Therefore they felt the best they could do for their children was to give them the tools and life skills to help them help themselves. One participant summed up this sentiment by saying that they as role models were not planning their children's futures, but were laying out the path for them to make their own choices.

THEME 6: *PFL*

Participants were asked directly about their awareness and perceptions of the *PFL* programme, alongside the perceived impact that *PFL* had had on them and their child and any suggested recommendations. Their responses are gathered within the different sub-themes described below.

SUB-THEME A) *PFL* KNOWLEDGE AND INVOLVEMENT

The majority of participants in both groups had little knowledge of the programme; however had some idea as to why their children were participating in it. A few participants from the low treatment group said that this focus group was the first contact the programme had made with them. Some participants stated that they preferred not to be involved with the programme and were happy to let the child's mother take the lead:

> *I know nothing about it to be honest… I never really wanted to… all the letters em and my wife opens the post and, "It's only that that Preparing for Life thing"…and I forget about it*
>
> Low treatment participant

In contrast, participants from both treatment groups, although not currently very involved with the programme, expressed interest in playing a more active role in *PFL* once they had more knowledge and a better understanding of the programme's purpose and objectives:

I want to get involved and find out what's happening […] I'd be interested in doin it...for the little lad you know? For him, I'd do anything for him

Low treatment participant

I want to get more involved, so if there was any more discussion groups, to involve me in them

Low treatment participant

It appeared that any contact participants did have with *PFL* was largely controlled by the children's mothers, who they depended on for knowledge and updated information:

I have the contact through [partner] and that, and it's only for when the letter came out to me, I didn't realise cause it was addressed to [partner]...this is only the first contact I'd had with yourselves

Low treatment participant

I wouldn't say it would really have a big impact on me because everything is really focused on my wife.... I would have to just have to pick up the slack when I come home I'd have to read the notes myself and if I'm not sure of things I'd have to get the wife to explain it to me ...and maybe she could get it wrong

High treatment participant

Much of the participants' lack of knowledge and involvement in *PFL* was also due to the barriers to child involvement mentioned previously, such as work or time limitations. Despite this, participants in general had quite a positive perception of *PFL*, and could see the difference it had made in their children as explained below.

SUB-THEME B) *PFL* PERCEPTIONS

Participants were asked about their perceptions of *PFL*. In general, they had a very positive view of the programme, with some stating that it should have been brought into the community years ago and that they would like to see the programme continue into the future. One high treatment participant claimed that despite initial scepticism, he would like to see it continue:

I had me sceptics at the start but I have to say it's an absolutely brilliant programme... again I em I'm not really that involved because I'm in work

High treatment participant

I'd like to see it goin on all the time so, and it helps people and parents...it probably should have been here years ago in the area, in different areas

Low treatment participant

Another participant commented on how he appreciated the personal touches of the programme, such as the contact made on his child's birthday. Almost all participants said they would recommend the programme, and one high treatment participant said he photocopied the programme's tip sheets and passed these and his knowledge onto friends.

SUB-THEME C) *PFL* IMPACT

Participants were asked about the impact the programme has had on their child, their family and themselves. Despite having limited knowledge of *PFL*, they felt that there was value in the programme and identified many benefits of an individual, relational and contextual nature.

One participant discussed what a difference the programme had made in his family's life and in particular his wife's life since they moved into the community. A lack of family support combined with work demands had made it difficult for them raising their child and so they saw *PFL* as offering a supportive setting where they could ask questions, get ideas and receive direction and parenting tips.

Mentor feedback was appreciated and in some cases was a positive reinforcement of the parents' parenting skills:

> *We know from our mentor that she is past the age of where she's meant to be, then we know that we're doing a good job and we just keep doing what we're doing*
>
> High treatment participant

> *It gives me a boost to know that...we're doing the things right*
>
> High treatment participant

One participant described the programme as a 'father figure' for single mothers who had no father figures in their children's lives, while another referred to it as 'like a mother you don't have'.

> *It's a great support system for, for even single mothers ...because the PFL is kind of a father figure for the child*
>
> High treatment participant

Participants from the high treatment group made reference to the variety of different help methods they were provided with, the good relationships they had established with their mentors and how it was nice to have 'another person behind the scenes' when they needed additional support.

SUB-THEME D) *PFL* IMPROVEMENTS

When asked about improvements that could be made, participants from both treatment groups called for a programme that was more accommodating and accessible to their needs, and one which gave due recognition to the importance of father figures in child development. They also called for *PFL* to facilitate more peer support and organise evening group meetings:

> *I would do more of these discussion groups if they were in the evening time, for especially for [father figures] and stuff like that as I work [during the day]*
>
> High treatment participant

Father figures wanted somewhere where they could ask questions, offer tips and get reassurance that they were 'doing things right' and that certain child behaviours and issues relating to that child were 'normal'. One participant commented that unlike mothers, he felt he could not ring other father figures to discuss parenting styles and tips. He also mentioned that it would be useful to hold meetings which facilitated engagement with other father figures in order to discuss fathering perceptions and experiences:

> *I know mothers...they're always on the phone to each other but it's harder for [father figures] to actually pick up a phone and say 'are you doing this with your child?'*
>
> High treatment participant

> *There'd be more information thrown about...you could take that feedback and go back and try and implement it with your child...it's just, you know, different ideas cause you know you can't come up with the ideas all by yourself*
>
> High treatment participant

One participant from the high treatment group felt that he would benefit from being present when the mentor visited, thus acquiring more first-hand knowledge which would better equip him in helping his child progress at each stage of her development. He felt this would be beneficial to both the child and himself in the long-run and would allow him to feel more included and involved:

> *I'd like to spend more time with her, with the mentor, so we can see what's going on*
>
> High treatment participant

> *I just want to have that connection with the mentor as well...to be seen as part of the collective group of the household*
>
> High treatment participant

7.4 Summary of focus group findings

A key finding that emerged from this study was the similarity between the high and low treatment participants on their thoughts about being father figures to young children. They saw themselves as being on a learning curve with their families, such that they were learning, their children were developing, and the rest of the family was also adapting to the child and to the father figure's role. Overall, they found the father figure role multi-faceted and challenging, but enjoyable and rewarding. Their father figure role was influenced by a number of external factors including their own father, their relationship with the child's mother, the wider family environment, and society in general. A number of barriers to interaction with their children were identified such as time, work, and the mother working as a gatekeeper in cases of relationship breakdown. Future plans for children included instilling confidence to follow their dreams, and having more opportunities in life than the father figures themselves.

Overall, the focus groups produced very rich data concerning the experiences of fatherhood among the *PFL* community, but relatively less data on father involvement in the *PFL* programme. Thus a key finding was the lack of knowledge about and involvement of the father figures in *PFL*. Father figures were generally supportive of *PFL*, but felt disconnected from it. High treatment group father figures would like to be more involved with mentors, and participants from both treatment groups would like the opportunity to meet other father figures through discussion groups. Interestingly, there were little differences between the high and low treatment group concerning knowledge or involvement in the programme. This is surprising as the Tip Sheets used by the mentors specifically address how 'Dad can be involved'. Yet, it is possible that the mothers are not passing the Tip Sheets on to father figures, or that the father figures are reading the Tip Sheets but they are missing the mentors discussing the issues with them in person. As experienced by many other home visiting programmes, these results suggest that *PFL* is not effectively engaging father figures in the programme. However, an important limitation of this study, which may have influenced this finding, concerns the sample size and the recruitment process. The sample size was relatively low (n=10), thus self-selection may be an issue if the father figures who agreed to participate in the study were not representatives of the 'typical' *PFL* father figures. In addition, the sample was recruited through the mothers, thus the potential for 'filtering' by the mothers was possible. More motivated families from the low treatment group may have participated as they perceived this study as an opportunity to engage with the programme, while participants in the high treatment group are regularly asked to engage in different dimensions of the programme such as the mentoring sessions and Triple P, and thus may have been less likely to engage. However, these hypotheses are conjecture and do not diminish the key findings from the study regarding the role of father figures in the *PFL* community.

Chapter Eight

Report Summary & Conclusion

8.1 Overview

This report presented the results of the effectiveness of the *Preparing For Life* programme between programme entry and when the *PFL* child was approximately thirty-six months of age. It is important to note that in addition to the home visiting programme, high treatment mothers and their partners were also offered a second treatment, the Triple P Positive Program, between twenty-four and thirty-six months. Triple P is a gold standard parenting programme which seeks to promote healthy parenting practices and positive attachment relationships by focusing on the home environment, parent self-awareness, and parenting techniques. Thus, the results of the thirty-six month evaluation should be interpreted in light of this additional treatment. Comparable with previous reports, the current report included an analysis of the quantitative information derived from the interviews with *PFL* participants and implementation data from *PFL*'s database. In addition, the results of two further studies were presented; a qualitative study with fathers and father figures regarding their experiences of fatherhood and the *PFL* programme; and the Day in the Life of a *Preparing For Life* Parent (DTL) study, an intensive study of maternal wellbeing using a unique combination of measurement techniques. The thirty-six month interview included the largest number of outcome measures employed at any wave of the evaluation to date. Overall, the results are consistent with other evaluations of home visiting programmes, which typically identify moderate treatment effects in areas such as child development, child health, and parenting and limited effects on secondary outcomes such as the home environment, maternal health and wellbeing, maternal social support, childcare and service use, and household factors and SES.

CAVEATS

The results of the thirty-six month evaluation should be considered in light of a number of caveats. Firstly, in addition to an unprecedented number of outcome measures included in the current wave of the evaluation, several of these measures are different to those used at previous time points. Consequently, it is inappropriate to make a direct comparison between the results for the eight outcome domains at different time points. Instead, it is more instructive to consider the proportion of tests in which significant positive treatment effects were observed.

Secondly, while the official level of attrition from *PFL* between baseline and thirty-six months was low, the total rate of non-completion[1] - approximately one third in each of the three groups - was the highest experienced in the evaluation to date. Thus, the thirty-six month analysis was based on 65% of the originally recruited sample. Consequently, the sample is not identical to samples involved at previous time points. Importantly, participants who dropped out of the programme or who were difficult to contact, tended to be more disadvantaged and there was evidence of differential attrition across the high treatment, low treatment, and comparison groups. To account for any potential bias arising from systematic group differences, the robustness of the outcome analyses was tested using an inverse probability weighting technique. Overall, there were few differences between the weighted and unweighted results, and the same number of multiple hypothesis tests were significant in both the original and the re-estimated results. This convergence provides a strong level of confidence in the results.

Thirdly, between twenty-four and thirty-six months, the average number of visits was 13 (min = 0; max = 39) which represented 50% of prescribed visits. This level of engagement increased slightly from the level of participation observed previously. In terms of engagement with Triple P, 59% of high treatment participants, who completed the twenty-four month interview, took part in some form of Triple P, although the level of treatment received was variable - ranging from attendance at a seminar to full participation in an eight week Group Triple P course. In addition, of the participants who took part in Triple P, one-third of their partners also received some form of Triple P.

Finally, there are a number of considerations that should be noted when interpreting the results of the DTL and father figures sub-studies. Unlike the thirty-six month interview, which took place over a two

[1]Which accounts for participants who attrited from the programme and disengaged from the evaluation.

year period in order to interview participants when they were at the same stage of the programme, the sub-studies recruited participants at various stages in the programme ranging from two and a half to five years. In addition, in the case of the father figures study, characteristics of the sample are pertinent to any interpretation of the results, namely the small sample size, the issue of selection, and the decision to analyse the high and low treatment samples in combination. Thus, it is inappropriate to generalise the findings to wider populations including other father figures in the *PFL* community. Nonetheless, it is appropriate to generalise the findings to a theory of *PFL* fathers' experiences of fatherhood and the programme (see Yin, 1989), which can be evaluated in light of other findings. A crucial point to note is that the experiences of fathers in home visiting programmes are typically inferred from the perspectives of mothers or home visitors (e.g., Thullen et al., 2014). Consequently, this study is the first to elucidate insight into the expressed experiences of father figures whose families are participating in home visiting. Regarding the DTL sample, fewer mothers took part in this sub-study than typically participate in the main evaluation. Furthermore, the DTL mothers tended to be more disadvantaged than mothers who did not participate. Thus, salient group differences between the high and low treatment groups were accounted for in the analysis. The central strength of this study is the unprecedented and fine grained understanding it provides regarding the day-to-day wellbeing of mothers participating in a home visiting programme.

OVERVIEW OF TREATMENT EFFECTS AT THIRTY-SIX MONTHS

Overall, 204 outcomes were tested in the main analysis contained in Chapter 2 and the majority of the relationships were in the hypothesised direction, with the high treatment group reporting better outcomes than the low treatment group. The high treatment group had better scores than the low treatment group on 142 (70%) of these measures. In this case, 'better' does not refer to statistical significance but to any positive mean difference. According to a binomial test of distributions, this proportion (142/204) is significantly different to 50% (which is what one would expect if the programme was having no impact). It is also the highest proportion of outcomes in the hypothesised direction observed in any of the reports to date. Of these 142 differences, 44 were statistically significant, which represents 22% of the total outcomes considered. In addition, 9 (26%) of the 35 step-down families were significant. In terms of effect sizes, small to medium effect sizes, of between 0.20 and 0.40, were identified for the majority of significant results.

As outlined in Chapter 1, we hypothesised that the highest proportion of positive treatment effects would be found in the areas of child development, child health, and parenting, with a limited number of positive treatment effects expected in the areas of the home environment, maternal health and wellbeing, social support, childcare and service use, and household factors and SES. The findings in relation to child development, child health, and parenting, conform with our hypothesis, yet contrary to our expectations, a considerable number of positive effects were also observed in the areas of the home environment and maternal health and wellbeing. Additionally, as predicted, few positive treatment effects were observed in the domains of social support.

The possible impact of the additional Triple P treatment on these outcomes is difficult to disentangle. Nonetheless, we can identify that for those mothers who did not participate in Triple P, the thirty-six month treatment effects were mostly concentrated in the areas of child health and child development. One possible interpretation is that these mothers were less motivated to take part in an additional treatment as their children's development was already benefiting from the programme. Equally, it is noteworthy that for mothers who did participate in Triple P, treatment effects were concentrated in the domains of parenting and the home environment. Given the focus of the Triple P Program on improved parenting, it is conceivable that Triple P participants experienced more direct benefits in these areas, although it appears that these gains did not translate into immediate effects for children's wellbeing.

In the context of the broader evaluation, a consistent pattern of treatment effects in the primary outcome of children's development and health has been observed between twenty-four and thirty-six months. There is also reasonable evidence to suggest that these gains may be accumulating through improvements in parenting, although the pattern of treatment effects is less consistent in this domain. Lastly, it is noteworthy that the largest concentration of treatment effects for maternal wellbeing was observed at

thirty-six months, in comparison to any previous wave of the evaluation. Collectively, the thirty-six month results indicate that the *PFL* programme has become well-embedded in the community and is generating real changes for participating families.

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 implementation, dynamic, comparison group, and Triple-P analyses, and the results arising from the father figures and DTL sub-studies. In addition, the results are contextualised within the relevant research literature, with particular consideration given to the Growing up in Ireland report on the development of the infant cohort at three years. The remainder of this chapter is structured by each of the eight outcome domains.

8.2 Child Development

Of the 39 child development measures assessed, 13 (33%) were statistically significant in the hypothesised direction, such that the children in the high treatment group outperformed children in the low treatment group. This is consistent with our hypothesis, based on the literature, that there would be moderate findings in the area of child development at thirty-six months. A number of home visiting programmes have found similar results at the same time point (e.g. Early Head Start, Early Start, Family Check-Up, Healthy Families America, Parents as Teachers). These findings were in the areas of cognitive and language development, social-emotional development, and child behaviour (e.g. Grant & Ridder, 2005). However, others have found no significant treatment effects on any of these domains (Schwarz et al., 2012; Roggman & Cook, 2010; Connell et al., 2008; Shaw et al., 2006; King et al., 2005; Olds et al., 1994). As discussed above, in addition to their regular mentor visits, high treatment families were offered the Triple P Parenting Program at twenty-four months which aims to teach parents how to manage and address their children's behaviour, while promoting competence and development. Although Triple P does not focus on child development specifically, it is possible that, through providing parents with strategies for dealing with difficult child behaviour, it also facilitates opportunities for improved child development as parents and children move beyond child behavioural issues and work together on promoting their developmental skills. Given this additional treatment and the potential cumulative effect of *PFL* over time, we hypothesised that there would be an increase in findings on child development from twenty-four to thirty-six months, however this was not supported as a similar number of effects were identified at both time points.

Significant findings emerged relating to children's cognitive development, as measured by the DP-3 (cognitive) scale. Children in the high treatment group had higher overall scores on this measure, indicating a higher level of cognitive abilities. They also scored higher than low treatment children on the proportion scoring above average on this measure. These results are consistent with those found at eighteen and twenty-four months. Similarly, the high treatment children outperformed the low treatment children on the total score for the ASQ, and specifically on the problem solving domain, which measures the child's ability to solve problems, to follow instruction, and to engage in pretend play. This is consistent with results at twenty-four months whereby no significant differences were observed in the areas of communication, personal social, gross motor or fine motor development. In addition, fewer high treatment children were below the cut-off score for problem solving, indicating that they were less at risk of developmental delay in this domain. These results are also consistent with those found, for the first time, at twenty-four months. Taken together, the results from the DP-3 and ASQ over the last two time points provides evidence that the *PFL* programme is having a positive effect on children's cognitive development.

Between the ages of two and four, the mentors delivered a number of Tip Sheets which provide advice on key developmental issues which may have contributed to these treatment effects. Indeed, 22 Tip Sheets focus on the area of cognitive development alone – covering school readiness factors including aspects of basic maths, scribbling, art, and using scissors, as well as basic self-care skills such as dressing oneself and tying shoelaces. Interestingly, results from the focus groups showed that fathers from both the high

and low treatment groups placed a strong emphasis on their children's academic progress and felt it was their role to be teachers and moral guides for their children. The area of play was also strongly emphasised in the Tip Sheets, with suggestions for encouraging pretend play or facilitating play through providing appropriate materials. Parents were encouraged to aid children's cognitive development through reading with them and helping to develop their speech and language skills.

A programme effect was also found in the area of child behaviour, as measured by the Child Behaviour Checklist (CBCL). Children in the high treatment group were less likely to exhibit somatic complaints, sleep problems, or aggressive behaviour. They also had lower total externalising problems and total problems scores, and were less likely to score above the cut off on these two domains. However, there was no significant difference between the high and low treatment groups on the internalising problems total score, or its cut off. There have now been consistent, positive programme effects as measured by the CBCL at twenty-four and thirty-six months. These results are noteworthy as the Growing Up in Ireland (GUI) national representative study found a clear linear trend between household social class and the probability that a child will be classified as having a problematic behavioural profile as measured by the SDQ (Strengths and Difficulties Questionnaire; Goodman, 1997). When asked as part of focus groups, fathers from both the high and low treatment groups emphasised emotional well-being as one of their future goals for their children. They also highlighted aspects of the child's life that can be detrimental to their developmental and emotional outcomes, such as conflicting parenting styles. Interestingly, there were no significant effects on the BITSEA measure of social and emotional problems or the ITSEA measure of pro-social and aggressive peer behaviours. However, it should be noted that the BITSEA and ITSEA measures are designed to be used up until thirty-six months and many of the interviews were conducted when the children had already turned three. Thus, the results may not be comparable to other studies. To our knowledge, only three other home visiting programmes have reported an impact on social-emotional development at thirty-six months (Early Start: New Zealand; Family Check-up; Healthy Families America), using the ITSEA and CBCL respectively. However, other evaluations of these programmes and evaluations of other home visiting programmes have not observed effects on these outcomes (Family Check-Up; Hawaii Healthy Start; Love et al., 2002).

Similar to cognitive development, 22 Tip Sheets delivered during this period targeted the area of social and emotional development. They provided advice on dealing with challenging behaviours such as lying and whining, as well as teaching parents how to speak to their children about their feelings and interactions with others. Meanwhile, Triple P focuses specifically on strategies for dealing with negative child behaviour while encouraging positive behaviour. Given Triple P's emphasis on parent and child behaviour, it was hypothesised that we would identify more results in terms of child behaviour at this time point than at twenty-four months. Although largely consistent, the thirty-six month results do slightly exceed those found previously. However, as these findings pre-date the introduction of Triple P, it is difficult to ascertain whether the finding is more strongly influenced by the mentoring, Triple P, or a combination of both. As the evaluation includes the CBCL measure at forty-eight months, this will allow us to examine any possible long term effects on socio-emotional wellbeing. The positive findings in the area of child behaviour at thirty-six months are noteworthy as child behaviour can be particularly challenging between the ages of two and three as children can become overwhelmed by their strong feelings and have trouble putting them into words (Parlakian & Lerner, 2008).

Additionally tests were conducted to examine the equality of standard deviations across the high and low treatment groups. The high treatment group were found to exhibit a lower degree of dispersion around the mean on 11 of the 28 continuous child development outcomes, relative to the low treatment group. This could be due to the *PFL* programme moving children in the high treatment group that would otherwise have very poor scores on these outcomes, closer to the mean.

The dynamic analysis, which tracks changes in child development across multiple time points, revealed some interesting findings. The results suggested that the high treatment group experienced an increase in total ASQ problem solving scores at thirty-six months, while the low treatment group's scores fell. This led to a significant widening of the gap in problem solving scores between the two groups. There was also a significant difference between the groups in the somatic complaints subdomain for the CBCL. While the

number of complaints reported by mothers in high treatment group decreased between twenty-four and thirty-six months, it rose for the low treatment group over the same period. These positive dynamic results suggest that in terms of problem solving and somatic complains, children in the high treatment group are performing better, relative to the low treatment group, than was previously the case.

Adjusting for differential attrition, using IPW, resulted in two additional significant findings in the individual tests and one additional finding in the stepdown tests. Specifically, once the IPW-weights were applied to the stepdown tests, the CBCL domain was significant, and in the individual tests the high treatment group performed significantly better than the low treatment group on the ITSEA pro-social and the CBCL attention subdomains. This suggests that the unweighted results may be slightly under-estimating the true impact of the programme on child development.

Analysis was also carried out to investigate differences between the low treatment and the comparison groups. Overall, the results suggest that the low treatment group generally outperformed the comparison group in the child development domain. These findings may suggest that the low level supports provided to the low treatment group (e.g. development packs) may have some impact and will need to be investigated further at the next time point.

8.3 Child Health

Of the 21 measures considered, 16 were in the hypothesised direction, and five (24%) were significant. This is in keeping with our hypothesis of a moderate treatment effect on child health at thirty-six months and is consistent with the literature which typically reports mixed evidence. However, it also reflects a lower proportion of significant findings than identified at twenty-four months. Positive programme effects were found in relation to the number of accidents and hospital stays reported, the proportion of children who had a chronic illness, and the number of children meeting the appropriate protein intake guidelines and overall dietary guidelines. Yet contrary to previous reports, there were no effects regarding overall child health, asthma, chest infections, and the number of times medical attention was sought in the previous twelve months.

In terms of physical health at thirty-six months, high treatment children were less likely to have had an accident or hospital stay over the past 12 months, and were less likely to have been diagnosed with a chronic illness, such as asthma or eczema, than low treatment children. This is in line with findings from other home visiting programmes which have reported favourable effects on accidents and hospital stays (Fergusson et al., 2005). While some of the Tip Sheets delivered earlier in the programme made reference to accident or illness prevention, there were no new Tip Sheets delivered between twenty-four and thirty-six months on this topic. In addition, one would not expect the Triple P Program, which mainly focuses on behavioural issues, to have a direct impact on accidents or hospitalisation. It is possible however, that improved child behaviour as described above, coupled with favourable parenting styles and a greater overall awareness of child health and safety from previous Tip Sheets have combined to create a safer overall environment for children. The GUI report indicates that approximately 16% of three year olds in Ireland have a longstanding illness, disability, or other ongoing health condition, with respiratory illness being the most commonly reported illness type (Williams et al., 2013). Findings from the current evaluation indicated that rates of chronic illnesses among children in the *PFL* community were 15% and 24% for the high and low treatment groups respectively, while 1% of the high treatment group and none of the low treatment group were reported to have a physical disability. While a favourable programme effect was previously observed in relation to wheezing and asthma symptoms, this effect was no longer observed at thirty-six months. At twenty-four months the low treatment group were more likely to have presented to the GP, health centre, or casualty in the preceding 12 months for asthma symptoms. At thirty-six months there is no significant difference for this variable; however, more mothers in the low treatment group reported that their child had an ongoing diagnosed chronic illness, with the majority of responses referring to asthma. Thus, it is possible that by thirty-six months, mothers in the low treatment group were managing their children's wheezing or asthma symptoms more effectively.

The programme appears to have a consistent and clear impact on children's nutrition. Results from previous waves have indicated that children in the high treatment group consumed more grains and dairy at twelve months, more protein and dairy at eighteen months, and more protein, fruits, and vegetables at twenty-four months. *PFL* continues to have a positive effect on nutrition at thirty-six months. Children in the high treatment group were more likely to eat the recommended daily allowance of protein and were more likely to have a healthy overall diet. There were a number of Tip Sheets delivered in this period relating to diet covering issues such as food groups, healthy eating, portion control, and serving size recommendations. There were also recommendations for making healthy choices while grocery shopping, alongside recipes and tips for meal planning. Therefore, information about nutrition provided by the mentors via the Tip Sheets may have influenced mothers in the high treatment group to make healthier food choices. It has been suggested that dietary quality is influenced by socioeconomic status (Kelleher et al., 2008). Using a semi-quantitative food frequency questionnaire, GUI reported that for three year olds, caregivers education was strongly and positively associated with fruit and vegetable consumption, and negatively associated with the consumption of energy-dense foods and non-diet fizzy drinks.

Given the difference in diet-related outcomes reported here, it is noteworthy that there was no significant difference between the high and low treatment groups in terms of their weight or BMI, with approximately one quarter of each group being classed as overweight. This is comparable to findings reported by Williams et al. (2013) in the GUI study, thus it is possibly representative of a problem at a national level rather than a local level. Note that the number of children for whom weight and height measurements were recorded was low for both the high and low treatment groups, with measurements taken for approximately half of the sample only. It was not always possible for the evaluation team to record these measurements as the *PFL* children were not always present while the interviews took place and some children refused to be measured. Other home visiting programmes focus on general health and illness related child health indicators and do not report child nutritional or weight outcomes beyond birth and infancy. However, given the attention to nutrition in the Tip Sheets, these are important outcomes for the *PFL* intervention.

The re-analysis of the child health findings using IPW resulted in fewer individual significant findings (14%), such that the results regarding accidents and hospital stays were no longer significant. This suggests that differential attrition across the sample may have affected the outcomes. Note. However, that the number of significant stepdown families actually increased in the IPW-weighted analysis such that both the long-term child health stepdown family and the meeting dietary recommendations stepdown family were significant.

Results from the comparison group indicate that this group had better health overall than the low treatment group. For example, they were less likely to have spent a night in hospital and more likely to have good self-reported health. While children in the low treatment group were less likely to have a physical disability, they were more likely to suffer from a chronic condition than those in the comparison group. In addition, the comparison group had higher overall diet quality scores and more of the comparison group met the daily requirement for the consumption of dairy products compared to the low treatment group. Given that the comparison group are from a community which was matched with the *PFL* community on a number of sociodemographic variables, it is noteworthy that the low treatment group had poorer health on a number of outcomes. This indicates that the extra supports provided to the high treatment group may be necessary to improve child health.

The overall trend suggests consistent gains in the area of child health, particularly with respect to nutrition, accidents, and hospital stays. In line with previous *PFL* reports, the programme continues to have a positive impact in relation to healthy eating. As the children's weight and height measurements are taken again at forty-eight months, it is possible that the sample size will increase given the age of the children and increased familiarity with weight measurement.

8.4 Parenting

Overall, significant effects in the hypothesised direction were found on nine (26%) of the 34 parenting measures. This is the largest proportion of significant findings on this domain to date. The favourable treatment effects were concentrated on two specific aspects of parenting: parenting styles and parental monitoring of child television watching. Group differences were also observed in relation to maternal attitudes towards education. This supports our hypothesis of moderate programme effects on parenting at thirty-six months. The number of parenting outcomes assessed at this time point was increased due to the implementation of the Triple P Program after twenty-four months. Thus, given the Triple P focus of improved parenting behaviour, we hypothesised a larger proportion of significant findings on some of the standardised measures of parenting. Indeed, we found that those who participated in some aspect of the Triple P Program had a higher proportion of favourable significant outcomes for the parenting domain than those who did not. This result was found for both the individual and stepdown tests. This provides some evidence that Triple P may have led to improvements in parenting behaviour.

High treatment parents were less likely than low treatment parents to utilise an authoritarian or permissive parenting style. More specifically, they were less likely to exhibit punitive or hostile parenting behaviours towards their children. This is a positive finding as the literature indicates that authoritarian or permissive parenting styles can negatively impact on child development. Authoritarian parenting, which is associated with demandingness and unresponsiveness, can lead to difficulties with social competence and self-esteem, while permissive parenting, which is associated with high responsiveness and low demandingness, is more commonly associated with problem behaviour and academic difficulties (Aunola & Nurmi, 2005). The GUI survey found that at three years, most Irish parents engage in warm and consistent parenting with low levels of hostility. In addition, they found that parents with low levels of education were more likely to demonstrate lower levels of consistent parenting (Williams et al., 2013). The finding that mothers in the *PFL* sample use less punitive or hostile parenting behaviours is particularly significant given the prominence of authoritarian parenting within low SES communities (Hoff et al., 2002). This result also supports findings from other home visiting programmes which were similarly effective in increasing non-punitive parenting attitudes, and reinforcement of positive child behaviour (e.g. Ferguson et al, 2005; Dishion et al, 2008). It is interesting to note, however, that no significant difference was observed regarding the use of the authoritative parenting style, which is widely considered to be the most appropriate style of parenting, and has been associated with positive child outcomes such as better school achievement and more psychosocial maturity (Kordi & Baharudin, 2010). Between the ages of twenty-four and thirty-six months, high treatment parents received a number of Tip Sheets on child behaviour which explained how to cope with disobedience, lying and difficult behaviours through techniques such as descriptive praise, encouraging children to discuss feelings, and affirmation of good child behaviour. The messages delivered through the Tip Sheets are congruent with the elements of an authoritative parenting style, as are the techniques taught in the Triple P Program where positive child behaviour is rewarded and encouraged, while negative behaviour is given less attention. Similarly, neither the mentors nor Triple P encourages or teaches behaviours that are in line with more permissive or authoritarian parenting. Thus, the clear reduction in negative parenting styles is an encouraging finding for *PFL*. However, as the high treatment parents received a focused parenting intervention via mentoring and Triple P, one may have anticipated a greater proportion of high treatment parents to report an authoritative parenting style compared to the low treatment group.

While the findings in relation to parenting style are encouraging, they are somewhat in conflict with the result in relation to parenting acceptance and rejection. Mothers in the high treatment group reported a significantly greater likelihood of engaging in behaviours associated with indifference and neglect. They also scored less favourably on an overall indicator of parental acceptance and rejection. Parental indifference and neglect refers to behaviours that indicate the physical or psychological unavailability of the parent (Rohner et al., 2005). These behaviours are directly in conflict with behaviours and strategies recommended on the Tip Sheets which encourage positive parenting practices, and thus this result is surprising. The measure used to assess parental acceptance and rejection, the Parental Acceptance and Rejection Questionnaire (PARQ) short form, has an established reliability and validity in multiple other

studies with a median reported Cronbach's alpha level of 0.82. However, when used with the *PFL* cohort the overall Cronbach's alpha level was 0.59, and Cronbach's alphas for the subdomains of hostility, indifference, and rejection were all less than 0.4. Thus, caution should be used when interpreting the results for this sample.

It is notable that there were no significant differences between the high and low treatment groups on the parenting daily hassles scale. This suggests that, while high treatment parents are exhibiting less hostile or punitive behaviours, they are nonetheless faced with the same challenges and difficulties, and experience them to the same degree as low treatment parents. The GUI survey found that the use of less favourable parenting strategies was more common in parents who were experiencing high levels of stress (Williams et al., 2013). At twenty-four months, high treatment parents reported lower stress levels which may suggest they are at lower risk of poor parenting, however, this outcome was not assessed at thirty-six months. Thus, the association between stress and parenting within the current sample may be explored in future analyses.

There were multiple significant differences between the high and low treatment groups in relation to television habits. High treatment children watched significantly less television than those in the low treatment group, specifically, the TV was switched on for fewer hours per day, and the children spent less time watching TV, DVDs, and videos. They were also less likely to watch TV alone. This is a positive finding as the literature indicates that watching too much television can negatively impact on child development (Pagani, Fitzpatrick, Barnett, & Dubow, 2009). In keeping with this, the GUI survey reported that at three years watching television for long periods of time had a negative impact on gross motor development. Additionally, few home visiting programmes have monitored child TV watching, thus in this respect, *PFL* has a relatively novel contribution to make to the literature on home visiting programmes. The results are comparable to the findings from Healthy Steps, which is the only other home visiting programme to examine television use (Johnston et al., 2006). The average time limit for TV watching was 1.95 hours in the high treatment group, which was within *PFL*'s recommended time limit of no more than two hours for children over twenty-four months as stated in the Tip Sheets (they advise against any TV viewing for children under the age of twenty-four months), and was also within the limit recommended by the American Academy of Pediatrics. Interestingly, only a certain proportion of families reported placing a daily limit on TV watching and this proportion was similar in both high and low treatment groups (High= 59%, *Low* = 58%) and relatively comparable to the UK finding of 51% (Ofcom, 2013). Three separate Tip Sheets discussed television viewing. One explained the long term negative effects of television on children including reduced health, increased aggression and imitation of TV characters' risky behaviours. The second Tip Sheet explained the concept of the age ratings applied to films, while the third advised on 'good' TV habits – not having TVs in children's bedrooms, turning the TV off during mealtimes, and watching TV with children etc. Finally, they explain that educational programmes, the news and wildlife programmes, in moderation, can have a positive impact on children. Thus, there is a clear link between *PFL*'s mentoring and improved child TV habits. However, given the clarity of the Tip Sheets around this issue, a higher proportion of high treatment families reporting imposing a limit on child TV watching may have been expected.

There were limited significant differences between high and low treatment mothers' attitudes to education, with those in the high treatment group expressing somewhat more favourable attitudes towards the education they themselves had received. While there were no significant differences on the other indicators of maternal attitudes to education, this remains a noteworthy finding as children whose parents have a positive attitude towards education are more likely to take an interest in their child's education and also to promote positive attitudes towards education in their children (Hoover-Dempsey & Sandler, 1995). *PFL* mentors support parents interested in accessing education, thus it is possible that this result reflects the facilitation of access to education through engagement with the programme. At baseline, mothers in the high and low treatment groups were equivalent in relation to their educational levels, however, at twenty-four months more mothers in the high treatment group reported having post-secondary school education than those in the low treatment group. It is possible that this increase in education may have resulted in the more favourable attitudes towards education reported by high treatment group mothers. The absence

of findings on the other indicators of maternal attitudes to education is unsurprising as the Tip Sheets do not explicitly focus on school preparation skills until after the children are three years old. In addition, the lack of an effect on the number of children on school waiting lists may be explained by the waiting list procedure in the schools. In some schools in the community, parents cannot add their child's name to a waiting list before a certain date for the upcoming year. Furthermore, results from the father focus groups would suggest that families in both the high and low treatment groups have similar educational aspirations for their children.

The thirty-six month evaluation is the first time that we do not observe a significant difference between the groups regarding mother's interaction with her child. It is possible that the mentors provided mothers in the high treatment group with strategies that facilitated interactions with their children at an early age when interactions are more parent-led. Now that the children are older, it is likely that low treatment mothers have naturally progressed to having the same level of interactions with their children, as children typically initiate interactions with their parents at this age.

An interesting finding to arise from the father focus groups was the emerging role of fathers within the community and the sense of the importance of coparenting. Fathers who participated in the focus groups mentioned the importance of consistency and referred to behaviours that reflect Triple P practices. While one third of those who engaged in Triple P participated with their partner, and partners sometimes engage in the mentoring sessions, we do not have father reports on parenting measures thus we cannot explore the influence of the programme on fathers' parenting practices.

When we explored the impact of controlling for differential attrition using the IPW method, we found no differences on any of the parenting measures in this domain in either the individual or stepdown tests. Another positive finding to emerge within this domain was the equivalence in outcomes between the low treatment group and the comparison community. Mothers in the low treatment group and the comparison group did not differ significantly across the majority of parenting domains, which suggests that contamination between the high and low treatment groups is low in this domain.

8.5 Home Environment

Overall, significant effects in the hypothesised direction were found on 6 (40%) of the 15 home environment measures. The results show that the high treatment families were more likely to have an organised, child-friendly home environment, where the parent is actively involved in the child's learning and development. These results were echoed in the father figure focus groups (high treatment group) which referred to the importance of stability in the home, and establishing and adhering to a routine, whereby quality time was dedicated to both play and studies. The current findings are encouraging as many studies in the home visiting literature do not find effects at thirty-six months (Duggan, McFarlane et al., 2004, Love et al., 2002; Jones Harden et al., 2012, Landsverk et al., 2002). Of those that do find effects in relation to the HOME measure, the subdomains on which treatment effects are observed differ across programme (Shaw et al., 2006; Wagner & Clayton, 1999). Given the implementation of Triple P from twenty-four months onwards, it was hypothesised that more significant findings would emerge at thirty-six months than in previous waves, as one of Triple P's focus areas is the home environment. However, the results in terms of the HOME measure were largely consistent with those found at the six and eighteen-month time points, with the high treatment group scoring significantly higher than the low treatment group on a number of subdomains. While there was some overlap, the subdomains on which the groups differed have changed over time. At six months, the high and low treatment groups differed in the areas of variety and learning materials, at eighteen months, they differed in the areas of acceptance and learning materials, at thirty-six months they differed in the areas of organisation, involvement, and acceptance. This inconsistency across time points is difficult to interpret, however, it is in keeping with results from other home visiting programmes using this measure.

These differing findings may reflect aspects of the home environment which are more salient at certain developmental stages. For instance, the acceptance subdomain measures parental acceptance of less than

optimal behaviour and undue restriction and punishment. The lack of an effect in this subdomain at six months may reflect children's mobility at this age and fewer opportunities to display less than optimal behaviours (Sheridan, 2004). Yet a significant effect on this subdomain was observed at eighteen and thirty-six months when children are more mobile and more likely to show irritability when parents place limits on the expression of their needs for autonomy and exploration (Carr, 2006). The *PFL* programme, through the mentoring visits and Triple P, has provided the high treatment group with parenting strategies which have enabled them to be better prepared and equipped with skills and knowledge, making them more able to accept less than optimal behaviours than the low treatment parents.

The dynamic analysis identified three significant differences on the HOME measure between the high and the low treatment groups over time. In terms of programme effectiveness, the results are mixed. While both groups' HOME scores decreased between eighteen and thirty-six months in terms of organisation of the home environment, the rate of decline was significantly greater for the low treatment group. It is possible that this decline is related to the changing developmental needs of children. This positive result may have been influenced by the mentors via the Tip Sheets which cover topics such as rest and routine and safety and supervision. However, the opposite pattern was observed for the learning materials and variety subdomains. On these subdomains both groups experienced a rise in scores, however, the increase experienced by the low treatment group was significantly larger than the equivalent increase for the high treatment group. Given the high treatment group scored significantly higher than the low treatment group on variety and learning materials at six and eighteen months, these dynamic results suggest that the groups' scores are converging over time. While these results are in the non-hypothesised direction, it suggests that the programme had an effect on these areas mainly during the critical period of early infancy.

Some limitations regarding the HOME measure should be noted. The subdomains on some of the HOME measures had very low Cronbach's alpha scores; suggesting low internal consistency. The authors of the HOME measure report coefficients of between .30 - .80 on the subscales (Bradley, 1994) and the alpha levels from the current sample are below these (.12 - .66). Thus the results presented here must be interpreted with caution. Furthermore, the HOME measure requires interviews to be conducted in the home with an awake and alert child. As this was not always possible in the *PFL* sample due to participant request or scheduling issues, the number of participants across subdomains differ. This should be considered when interpreting the results as self-selection may have occurred.

In terms of being involved with a social worker, no significant difference was found between the high and low treatment families. This result is similar to that found by the Healthy Families America programme which reported no differences in terms of interactions with child protective services. At twenty-four months, the *PFL* results showed that a lower proportion of high treatment families were involved with a social worker than the low treatment families. However, the small number of families reporting social worker involvement may account for these differing results.

Our results also found that children living in high treatment families were less exposed to cigarette smoke than those in low treatment families. While this finding is consistent with the eighteen month result, at twenty-four months the groups did not significantly differ. As part of the programme, two Tip Sheets were delivered specifically addressing smoking and health (between pre-birth and six months) and the effects of passive smoking (between twenty-four and thirty-six months). To our knowledge, only one other home visiting programme reported on children's exposure to smoke in the home at thirty-six months. In contrast with our results, the Early Start programme found no effects in the percentage of children living in a smoke free home (Fergussson et al., 2005). The results reported here are particularly noteworthy as research conducted by the Health Service Executive (HSE) noted higher rates of cigarette smoking among lower socioeconomic groups in Ireland, indicating that children living in disadvantaged areas may be particularly vulnerable to increased risk of exposure to second hand smoke (HSE, 2013b).

When IPW was explored to control for differential attrition there was no difference in the number of significant results observed in relation to individual home environment tests. However, once the weighting was applied to the stepdown tests, the HOME stepdown family was no longer significant. In addition,

there were a number of differences observed between the low treatment group and the comparison group across a number of measures in the home domain. The pattern of differences between the groups was inconsistent, for example, low treatment families demonstrated higher levels of organisation on the HOME measure, but not on the family environment scale.

8.6 Maternal Health and Wellbeing

Of the 17 maternal health and wellbeing measures considered at thirty-six months, 14 were in the hypothesised direction and four (24%) were statistically significant. This represents the highest proportion of significant findings in this domain across all five waves of the evaluation to date. These results are consistent with findings from other programmes which found that home visiting has limited effects on maternal health and wellbeing at thirty-six months. Positive programme effects were observed on measures of maternal psychological wellbeing, consumption of alcohol, and cigarette smoking.

The high treatment group mothers were less likely to report symptoms of postnatal depression, and were more likely to experience positive subjective wellbeing as measured by the WHO-5 scale. This is in line with findings reported by Shaw and colleagues (2009) from the Family Check-Up programme in which mothers in the intervention group reported a significantly greater decrease in depressive symptoms. This is the first time that significant differences on measures of maternal psychological health have been observed for the *PFL* programme, despite measuring mental health in the three previous waves using the same instruments. Overall, the current findings are surprising as HVPs tend not to produce meaningful effects on maternal depression based on evidence from a systematic review (Ammerman et al., 2010). It is possible that these results are related to the reduction in child externalising problems and lower total CBCL scores within the high treatment group. Mothers in the high treatment group reported that their children were less likely to exhibit somatic complaints, sleep problems, or aggressive behaviour than mothers in the low treatment group. As a result, it could be argued that mothers in the high treatment group have higher scores on measures of psychological wellbeing as a result of their children displaying less behavioural problems and difficulties. The GUI study recently reported that increases in parental stress between nine months and three years of age were associated with an increased likelihood of behavioural problems (Williams et al., 2013).

In relation to maternal psychological wellbeing, changes were also observed over time. The high treatment group experienced an increase in their total scores on the WHO-5 between twelve and thirty-six months, revealing improved subjective wellbeing, yet the low treatment group's scores declined slightly, indicating a significant strengthening of programme effect on mental health. Dynamic analysis of the EPDS data also revealed that the proportion of the low treatment group classified as being at risk of depression rose significantly faster than the high treatment group. This indicates that the high treatment group's psychological health has improved over time and this may be as a result of the mentoring and Triple P supports provided to this group. Tip Sheets provided to mothers remind them of the importance of self-care and taking time for themselves. These Tip Sheets were distributed before the child's second birthday, thus it is possible that this advice, as well as discussions with mentors, may have had an impact on long term health and wellbeing. The importance of self-care was also recognised by fathers who attended the focus groups. One participant from the high treatment group mentioned a need for "me time" which was not often met, while a participant from the low treatment group mentioned how family requirements often came before his own needs.

Data from the DTL sub-study sought to provide a more intensive investigation of maternal wellbeing in the *PFL* programme by using a multi-method approach to examine experienced affect using a day reconstruction method, global mood and life satisfaction, and parenting stress. The study incorporated more experience-based measures of wellbeing than have been used in the literature to date and report on both positive and negative affect. In relation to experienced affect, mothers in the high treatment group reported significantly higher positive affect during time spent without the *PFL* child. Mothers in the high treatment group also reported spending a significantly greater proportion of the day in a positive mood than the low treatment group. As stated previously, this may indicate that mothers in the high treatment

group recognise the importance of self-care and taking time out as a result of the information given to them by their mentors via Tip Sheets. Alternatively, due to the increased energy and effort exerted during their time spent with children, high treatment mothers may experience greater wellbeing in their time spent without children. In keeping with previous literature, there were no differences in negative affect using either experienced or standardised measures (Sweet & Appelbaum, 2004). In addition, there were no differences in life satisfaction. It is possible that consistent reductions in maternal stress and depression are only achievable via more targeted therapeutic interventions (Ammerman et al., 2010) which are beyond the scope of the *PFL* programme.

In relation to alcohol intake and cigarette consumption, mothers in the high treatment group were less likely to have drunk alcohol in the previous 12 months and they reported a reduction in the number of cigarettes smoked between twenty-four and thirty-six months. Previously, reduced alcohol intake has been reported by mothers in the high treatment group at twelve months and fewer mothers in the high treatment group reported binge drinking at eighteen months. This shows some consistency in attitudes to alcohol intake over time. This is the first time that differences in maternal reported smoking have been observed in the *PFL* evaluation.

Overall, the re-analysis of the maternal health and wellbeing results using IPW resulted in fewer individual significant findings, yet the same number of significant stepdown findings. Specifically, the between group difference observed on the WHO-5 did not remain significant once differential attrition was adjusted for, however, the overall maternal mental health stepdown family remained significant. When compared with the low treatment group, the comparison group reported lower levels of depression on the EPDS and higher levels of emotional wellbeing on the WHO-5. This suggests that high treatment supports may be particularly beneficial for improving maternal psychological wellbeing within the *PFL* community which typically demonstrates low level of wellbeing without treatment. Mothers in the low treatment group also reported fewer trips to the GP than mothers in the comparison group, again indicating better health within the comparison group.

Significant programme effects in the domain of maternal health and wellbeing have been consistently low in previous waves. *PFL* is a programme which is designed to promote children's school readiness skills. Improving maternal health and wellbeing is a secondary outcome and therefore it is unsurprising that few significant effects have been observed in this domain to date. However, improvements in maternal psychological wellbeing at the thirty-six month evaluation are promising and may be a result of improvements in the primary outcomes targeted by the programme.

8.7 Maternal Social Support

Of the 19 measures included in the maternal social support domain, all were in the hypothesised direction and one (5%) measure was significant. Few programmes find that home visiting has a significant effect on social support at thirty-six months. One study reported differences at thirty-three months, however, social support was measured as the time to referral for early intervention and the time to receipt of early intervention services (Schwarz et al., 2012). No differences were observed in the present study on the Maternal Social Support Index (MSSI) or maternal reports of support from their partner, relatives, friends, or neighbours. The same proportion of significant findings remained when IPW was used to adjust for differential attrition from the programme.

Findings in relation to social support have changed over the course of the *PFL* evaluation. At six months there were differences on 38% of measures, increasing to 43% at twelve months. However, there was a sharp reduction in the percentage of significant findings at eighteen months to just 8%, with a small increase at twenty-four months to 11%. It is possible that these results may be linked to changes in parenting over time. Mothers may feel more isolated when they have a new-born baby, thus mothers in the high treatment group may have sought more social support at this time as a result of advice from their mentors. However, as infants get older, it is likely that a wider social network is available to all parents as children begin to explore their environments and take part in group-based activities.

Mothers in the low treatment group were less likely to receive a lot of support from their friends and their parents when compared to the comparison group. No differences were observed between the high and low treatment groups, or the low treatment and comparison groups in relation to family quality of life. This is the first time that Family Quality of Life (FQOL) scale has been included in the evaluation and limited effects were hypothesised for this measure. In addition, a measure of relationship romantic quality (Relationship Quality Index) was included for the first time, and again no differences were observed on this measure. Mothers in both groups reported being least satisfied with how the programme had improved their relationships with their partner according to the Client Satisfaction Questionnaire (CSQ) scale. Tip Sheets which contain information relating to the relationship between the *PFL* child's mother and father/ partner are distributed by mentors earlier in the programme, typically between pregnancy and twenty-four months. The focus of the Tip Sheets moves more specifically towards the primary outcomes of the programme and factors which affect school readiness between twenty-four and thirty-six months. This may have had an impact on findings in the social support domain at this time point.

Fathers from both treatment groups, who took part in the father figures focus groups, described how a family-oriented role model approach was adopted in their household when raising their children. Fathers emphasised the need for more help and societal support and spoke about expanding their roles from that of a breadwinner to providing love and support. Participating fathers also highlighted the importance of the mother-father figure relationship and the importance of cohesive parenting practices. Participants from both treatment groups valued a united style of parenting, where the mother and father figure were equal partners in the upbringing of their children. The fact that fathers from both treatment groups emphasised the importance of the father role may explain why we do not observe any significant differences between the groups on the social support variables pertaining to the partner.

Social support is a secondary outcome of the *PFL* intervention and as a result it was hypothesised that limited differences would be observed as a result of the high treatment intervention.

8.8 Childcare & Service Use

There were no significant differences among the 17 measures analysed between the high and low treatment groups in relation to childcare and service use. The research in relation to childcare and service use practices of families participating in home visiting programmes is particularly limited. Service use is commonly recorded in relation to referrals to health services, however childcare is rarely recorded. One exception is the Early Head Start home visiting programme in New Zealand, which reported higher rates of attendance at early childhood education facilities among intervention group participants (Fergusson et al., 2005). There are a number of academic and behavioural benefits of formal childcare for children who begin attending at age 2-3, particularly among those from a lower SES background (Barnett, 1995; Geoffroy et al., 2010, Sylva et al., 2011, NICHD, 2002). However, little is known about informal childcare, although it is reported that such arrangements do not produce the cognitive benefits associated with formal childcare (Andersson, 1989).

The figures observed in the *PFL* cohort in relation to childcare use differ somewhat from figures reported in the GUI study at three years (Williams et al., 2013). Overall, the rates of formal childcare use in Ireland appear lower and the cost of childcare appears higher than that experienced by the *PFL* sample. However, if only the lower socio-economic group of the GUI sample are considered, the figures are in keeping with those observed for the *PFL* cohort. This suggests that at thirty-six months, parents in lower socio-economic groups in Ireland are more likely to use formal childcare and are also likely to pay lower rates for this type of childcare than their higher SES counterparts.

A common support offered to both high and low treatment groups is facilitated access to enhanced preschool, thus it is noteworthy that over 60% of the sample attend a Síolta quality accredited crèche. In keeping with this, it is also noteworthy that the low treatment group were more likely to use formal childcare and less likely to use grandmother care than the comparison group. Findings from the five reports to date indicate very limited treatment effects relating to childcare. Limited significant differences

between the high and low treatment groups were observed at six months (6%) and no positive significant effects subsequently. We will continue to monitor childcare, however given the consistency of findings at previous time points in conjunction with the introduction of the free preschool year in Ireland, it is unlikely that we will observe differences between groups in this domain at forty-eight months.

In relation to service use, both the high and the low treatment groups reported similar levels of use across all types of services. Accounting for differential attrition using the IPW method produced one significant difference in the childcare and service use domain. Specifically, the use of community services was significantly higher in the high treatment group once the weights were applied. The findings in relation to service use are in conflict with the hypothesised theory of change, outlined at the outset of the intervention, that use of preventative health services would be an avenue for programme impact. Few other studies record service use at this time point. One study conducted by Schwarz et al. (2012) found that recipients of the US based MOM Program were more likely to have been referred to and to have received early intervention at thirty-three months. However, this indicator of service use is not directly comparable to the indicator adopted by the *PFL* evaluation which records the number of community services accessed by the participant in the last year, rather than access to specialised services where referral is necessary. Thus it is difficult to determine whether the *PFL* programme is having a comparable impact as other home visiting programmes in relation to service use.

8.9 Household Factors & SES

Of the 42 household factors and SES measures assessed, six (14%) were statistically significant in the hypothesised direction. There is substantial evidence documenting the impact of household factors and SES on children's development, their well-being and school readiness (Hirsch, 2007; Letourneau et al., 2011). The *PFL* evaluation uses multiple indicators of household factors and SES at each time point. At thirty-six months, families in the high treatment group did not differ significantly from families in the low treatment group in terms of household size or marital status. However, there were a number of significant differences in other areas. Similar to findings from previous reports, mothers in the high treatment group were less likely to be unemployed than mothers in the low treatment group. This finding was also reported at eighteen and twenty-four months. It is important to note that there was no significant difference in the percentage of mothers in the high and low treatment groups who reported that they were employed. However, more mothers in the high treatment group described themselves as looking after their homes and families on a full-time basis. Similar to the results at twenty-four months, it is likely that this difference accounts for the significant difference between the two groups on the measure of unemployment.

Several significant differences were identified at thirty-six months that were not evident at previous time points. For example, mothers in the high treatment group were more optimistic about their financial situation for the coming year. It is likely that this result is associated with the improvement in work status reported by mothers in the high treatment group. This change in work status is also reflected in the increase in the proportion of mothers in paid employment between twenty-four (36%) and thirty-six months (41%). These results may be related to support and encouragement provided by the mentors in relation to access to employment and education.

While the results in relation to maternal employment at thirty-six months are favourable, fathers in the high treatment group appear to be experiencing difficulties in relation to employment. Specifically, high treatment fathers were more likely to be unemployed and less likely to have had a positive change in their work status over the last year. These findings are most likely related to the negative treatment effect in relation to receipt of unemployment benefit within the household. These findings have not been observed at previous time points, thus it will be informative to determine if this pattern in relation to paternal employment is maintained at forty-eight months.

In relation to other household or SES factors, families in the high treatment group were less likely to be living in social housing. When this outcome was assessed at six months there was no significant difference between the two groups. In addition, participants in the high treatment group reported higher levels of satisfaction with their neighbourhood. Participants in the high and low treatment groups live in close proximity to one another, however it is possible that the high level of support provided by the programme leads to a greater sense of satisfaction with the area. This is a difference that was not observed at baseline, although a different indicator of neighbourhood satisfaction was used at both time points. Finally, mothers in the high treatment group reported that they were more likely to be experiencing difficulties with a past or present partner. This result was not observed at eighteen months and thus suggests that conflict with a past or present partner is not a consistent feature in the lives of high treatment group participants. In addition, given that there were no differences between the groups in relation to satisfaction with partner or support from baby's father, it is difficult to interpret the factors that may be influencing this result. Results from the father focus groups suggest that fathers are enthusiastic about engaging with their children, but sometimes experience tension when trying to balance the demands of fatherhood with other aspects of their life such as work and other commitments.

When we account for differential attrition, the proportion of significant differences in this domain decreases. The results in relation to unemployment status, improvement in work status, financial outlook, social housing, and neighbourhood satisfaction were no longer significant when IPW-weights were applied. Furthermore, two additional significant differences were observed when weights are applied; the high treatment group were less likely to receive one parent benefit, and more likely to be partnered with the father of their *PFL* child. Similar to previous time points, there were a number of significant differences between the low treatment group and the comparison group, with the comparison group generally outperforming the low treatment group in this domain.

The results in this domain, for the most part, are in keeping with the home visiting literature at thirty-six months. Similar to other home visiting evaluations, the *PFL* evaluation found no impact on multiple indicators of welfare receipt, indicators of education, and most indicators of family composition (Fergusson et al., 2005; Landsverk et al., 2002; Schwarz et al., 2012). However, in contrast to the results from the *PFL* evaluation, Love et al. (2002) found a significant negative programme effect in relation to continuous male presence. In addition, Landsverk et al. (2002) reported a favourable programme impact on maternal education attendance. The *PFL* evaluation did not identify any significant differences in relation to household composition or maternal education. In interpreting these comparisons it is important to note that as household and SES indicators used in home visiting evaluations vary across programmes, it is difficult to draw direct comparisons between the results.

8.10 Future Reports

A further report will assess primary and secondary outcomes of the programme when the *PFL* children are forty-eight months old. This report will also capture the pattern of treatment effects for families across the whole course of the programme. An additional report will examine children's school readiness as they enter primary school settings. Collectively these reports will provide the final overview of the programme's success in improving the development, health, and wellbeing of *PFL* 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