preparing for life

Early Childhood Intervention

Assessing the Impact of *Preparing For Life* at 12 Months

By the *PFL* Evaluation Team, UCD Geary Institute

UCD Geary Institute

preparing for life
planning together for our children

Preparing For Life:
Early Childhood Intervention

Assessing the Impact of
Preparing For Life at Twelve Months

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

By
PFL EVALUATION TEAM AT THE UCD GEARY INSTITUTE
December, 2012

UCD Geary Institute

preparing for life
planning together for our children

Table of Contents

List of Tables

List of Figures

Acknowledgements

Noel Kelly,

Manager, *Preparing For Life*

Executive Summary

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

Recruitment & Baseline Characteristics

In total, 233 pregnant women were recruited into *PFL* between January 2008 and August 2010. Randomisation resulted in 115 participants assigned to the high treatment group and 118 participants assigned to the low treatment group. In addition, 99 pregnant women were recruited into the comparison group. The population-based recruitment rate was 52%. Baseline data, collected before the programme began, was available for 104 high and 101 low *PFL* treatment group participants respectively, and 99 comparison group participants. Tests of baseline differences between the high and low *PFL* treatment groups found that the 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.

Findings from the Six Month Report

The six month evaluation of *PFL* indicated that the programme was progressing well. 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%), the high treatment and comparison groups (21%) and the low treatment and comparison groups (11%) 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 correspond directly to information provided by the *PFL* mentors. However, the programme had no significant impact on pregnancy behaviour, infant birth weight, breastfeeding and child development at six months. While attrition from the programme was low and participant satisfaction was high, the level of engagement was less than anticipated. 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.

Twelve Month Report

The aim of this report is to determine whether the *PFL* programme has had an impact on parent and child outcomes at and before twelve months of age, and to provide a detailed review of implementation practices in the programme regarding attrition, dosage, participant engagement and programme effectiveness.

Impact of *PFL* at Twelve Months: Main Results

In total, 247 twelve month interviews (n*Low* = 83; n*High* = 82; n*LFP* = 82) were completed. The main results 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 well-being, social support, childcare, and household factors and socioeconomic status (SES), incorporating 147 outcome measures.

Table ES.1 compares the significant findings from the home visiting literature with the significant findings from *PFL* at six and twelve months. Consistent with the literature, there were limited differences observed between the high and low treatment groups at six and twelve months. However, many of the outcomes were in the hypothesised direction, with the high treatment group reporting somewhat better outcomes than the low treatment group.

Table ES.1 - Summary Comparing *PFL* Findings with Findings from Home Visiting Literature

Domain	Six Months		Twelve Months	
	Studies w/ Favourable Outcomes	*PFL* Favourable Outcomes	Studies w/ Favourable Outcomes	*PFL* Favourable Outcomes
Child Development	✔			✔
Cognition, Communication, Problem Solving	✔			
Physical Development				✔
Socio-emotional Development	✔			✔
Child Health	✔	✔	✔	✔
Immunisations, Hospitalisations	✔	✔	✔	✔
Medical care	✔			
Physical Health	✔		✔	✔
Health Behaviours (Sleeping, Eating)				✔
Parenting	✔	✔	✔	
Risk Factors for Maltreatment		✔	✔	
Parental Knowledge			✔	
Parenting Behaviours	✔	✔	✔	
Home Environment	✔	✔	✔	
Maternal Health & Wellbeing	✔	✔	✔	✔
Physical Health	✔	✔		
Mental Health	✔		✔	
Health Behaviours			✔	✔
Social Support		✔		✔
Childcare	✔	✔	✔	✔
Household Factors & SES			✔	
Household Composition				
Parental Employment, Finances			✔	

Table ES.2 summarises the *PFL* results at six and twelve months. Based on the literature, we hypothesised that treatment effects at twelve months would be found in the domains of child health, parenting and maternal health. The results suggest partial support for our hypotheses. Although there were no significant effects found in the domain of parenting at twelve months, effects were found for maternal health and child health. Counter to our hypotheses, significant treatment effects were found in the domains of child development and social support. This is noteworthy as previous studies of home visiting programmes do not report effects in these domains at twelve months. Three of the 23 step-down categories remained significant in the multiple hypothesis analysis, including child development, child health, and maternal health. Specific findings include the following. Children in the high treatment group compared to those in the low treatment group displayed a higher level of fine motor skills and were less likely to be at risk for social and emotional difficulties. Additionally, they had more appropriate eating patterns and had a higher level of immunisation rates. Moreover, mothers in the high treatment group were more likely to regularly meet their friends.

Table ES.2 - Summary of Main Findings at Six & Twelve Months

PFL Low – *PFL* High	Proportion of Measures Significantly Different at Six Months		Proportion of Measures Significantly Different at Twelve Months	
	Individual Tests	Multiple Hypothesis Tests	Individual Tests	Multiple Hypothesis Tests
Child Development	0% (13)	0% (2)	7% (28)	20% (5)
Child Health	10% (30)	0% (3)	17% (23)	0% (4)
Parenting	23% (22)	20% (5)	0% (16)	0% (2)
Home Environment	36% (22)	50% (2)	0% (6)	0% (1)
Maternal Health & Wellbeing	5% (20)	25% (4)	4% (28)	25% (4)
Social Support	38% (13)	0% (2)	43% (7)	0% (2)
Childcare	7% (14)	0% (2)	~	~
Household Factors & SES	0% (26)	0% (5)	3% (32)	0% (5)
Total Statistically Different	14% (23/160)	12% (3/25)	8% (11/140)	9% (2/23)

Fewer significant findings were reported at twelve months compared to six months. However, this is mainly due to differences in the measures included at each time point. For instance, at six months the home environment domain focused on aspects of the physical environment and appropriateness of toys and activities, while at twelve months it focused on aspects of the family relationship.

Impact of *PFL* at Twelve Months: Dynamic Analysis Results

A number of standardised instruments used to evaluate the programme are collected at multiple time points. This allows for comparison of responses for the same participants over time in order to track changes in child and parent outcomes. Three common instruments were administered at six months and twelve months, three were administered at baseline and twelve months and one was administered at baseline, six and twelve months. These measures reflect parent and child outcomes, as well as participant satisfaction. Children in the high and low treatment groups did not differ significantly across many of the child developmental measures in this dynamic analyses. However, children in the high treatment group had significantly better fine motor skill development between six months and twelve months than those in the low treatment group. Additionally, children in the high treatment group were significantly less likely to experience parental oppression of their power and independence by twelve months than those in the low treatment group. There were no differences on the remaining 24 measures.

Impact of *PFL* at Twelve Months: Interaction Results

An interaction analysis was conducted to determine whether the programme had a varying impact on girls or boys, first time or non-first time mothers, lone or partnered parents, mothers with higher or lower cognitive resources, families with low or high familial risk and mothers with low or high emotional well-being. The results indicated that the programme had differential impacts with some groups benefiting more from the programme than others. For example, there was suggestive evidence that at twelve months the programme benefited partnered mothers, primiparous mothers, and boys. It is important to note that the number of significant findings within the interaction analysis were limited.

Impact of *PFL* at Twelve Months: Comparison Group Results

As hypothesised, the comparison of the twelve month outcomes of the two *PFL* treatment groups and the comparison group (*LFP*) found there were more significant differences in the outcomes of the high treatment group versus the comparison group than in the outcomes of the low treatment group versus the comparison group. Specifically, of the 140 individual outcomes analysed, there were positive significant differences between the high treatment group and the comparison group on 20 measures (14%), with most effects in the domains of child development, parenting and social support. However, only two of these effects remained significant in the multiple hypothesis analysis, and both of these were in the parenting domain. In addition, there were positive significant differences between the low treatment group and the comparison group on 12 measures (9%), with most effects in the domains of child development and parenting. Four of the 23 step-down categories remained significant in the multiple hypothesis analysis, including child development, child health, parenting and household factors.

Overall, the results of the high treatment group and comparison group analysis support the main findings, such that the additional supports provided to the high treatment group appeared to have some positive effects at twelve months. However, in some cases the domains in which effects were found differed. For example, the analyses suggest that parents in the high and low treatment groups read more to their children than those in the comparison group, while there was no significant difference between the high and low treatment groups. This indicates that programme components may have had some impact on parenting practices for *PFL* parents. Similarly, both the high and low treatment groups reported better child cognitive functioning than those in the comparison group, while no difference was detected between the high and low treatment groups. Furthermore, although no significant differences were found between the high and low treatment groups in the home environment domain, the high treatment vs. comparison group analyses revealed that the high treatment group reported less family conflict than the comparison group. These findings suggest that some of the common programme components, such as the developmental and reading packs, may have an impact on all of those participating in the *PFL* programme.

PFL Implementation Analysis at Twelve Months

ATTRITION

On average, 15% of the sample officially dropped out of the programme between the baseline assessment and twelve months (*High*=20%, *Low*=14%, *LFP*=9%) and 10% of the sample were classified as disengaged (*High*=9%, *Low*=15%, *LFP*=8%). In addition, the level of attrition between six and twelve months was extremely low (2 participants) and the proportion of participants in the high and low treatment groups completing a twelve month survey was almost identical. Very few individual participant characteristics were associated with programme attrition and disengagement. Overall, there is weak evidence to suggest that there are systematic differences based on relative disadvantage between those who completed the twelve month survey and those who did not.

ENGAGEMENT

Families in the high treatment group received an average of 21 home visits by the *PFL* mentors between programme intake and twelve months, with each visit lasting about one hour on average. The frequency and duration of visits did not differ significantly across each time period. The majority of participants reported meeting their mentor twice a month (61%). Few individual participant characteristics were associated with the frequency or duration of home visits. Two factors were associated with both the frequency and duration of home visits in the prenatal period – mothers who joined the programme earlier in pregnancy and mothers with higher self-esteem were more likely to engage in the programme prenatally. The only factor related to engagement in the post-birth period was cognitive resources, such that mothers with higher cognitive resources engaged in more home visits and these were of a longer duration.

SATISFACTION

Overall, participant satisfaction with the programme between six and twelve months was high. As expected, the high treatment group reported greater satisfaction with the programme than the low treatment group. However, the low treatment group still reported relatively high levels of satisfaction despite the minimal supports received. The areas where participants reported the highest and lowest levels of satisfaction were similar across the two groups. In line with the six month findings, both groups were generally satisfied with the whole programme, their children's progress and the type of help they receive from the programme. However, both groups reported being least satisfied with how the programme has improved relationships with their partner.

CONTAMINATION

A contamination analysis was conducted to determine whether the low treatment group received all or part of the additional services designed for the high treatment group. This analysis found that the potential for contamination was high as the participants reported knowing other participants in the programme and sharing their *PFL* materials. However, the direct measure of contamination suggests that these practices did not translate into improved knowledge about *Preparing For Life* for those in the low treatment group. These findings indicate that the level of contamination in the *PFL* programme up to twelve months was quite low and does not bias the twelve month results.

Conclusion

The twelve month evaluation of *Preparing For Life* suggests that the programme is progressing well regarding the retention of participants and programme satisfaction and the results are in line with evaluations of other home visiting programmes, which typically identify few significant effects at twelve months. Many of the relationships were in the hypothesised direction, with the high treatment group reporting somewhat better outcomes than the low treatment group. As hypothesised, there were some significant findings in the domains of child health and maternal health. Furthermore, contrary to much of the literature, there were significant findings in the domains of social support and child development across all groups which corresponded directly to information on the *PFL* Tip Sheets delivered to participants during this period. However, the programme had no significant impact on key factors such as parenting and the home environment, although it is important to note that different aspects of these two domains were measured at six and twelve months. An analysis of the eighteen month data, which includes many of the measures observed at six months, will prove informative regarding the consistency of programme effects and we will begin to form a more in-depth picture of the kinds of factors that are most affected by the programme. In addition, as more data is gathered, we will be able to expand the dynamic analysis to compare the developmental trajectories of children over time.

The detailed report of the six and twelve month *PFL* evaluations can be found at the following website under publications: **http://geary.ucd.ie/preparingforlife**

The Life of an Average PFL Child at 12 Months

Kirsty is a twelve month old girl from the *PFL* catchment area. She lives at home with her mother, father and her siblings. Although her parents are unmarried they are in a committed partnership. Her grandparents also play a significant role in her life. She sees them often and they look after her regularly. Kirsty's mother has been a participant in the high treatment group of *PFL* since she was 22 weeks pregnant. Kirsty's family has had 21 visits from their *PFL* mentor and Kirsty has become quite comfortable with her. Kirsty's family is at low risk for problems such as addiction, abuse or family violence, however there are some issues within her family such as family conflict. Both her mother and father completed their Junior Certs. However, her father is out of work and the family is receiving social welfare payments. Kirsty's family have received a lot of information from *PFL* about how to enhance Kirsty's development and deal with common parenting problems that often arise. Kirsty eats appropriate foods for her age and is up to date on her immunisations. At twelve months of age she sleeps in her own cot and sleeps throughout the night. Her mother is generally in good physical health compared to other women, although there is at least one adult in her home who smokes cigarettes. This puts the Kirsty at greater risk for bronchial issues, such as chest infections. Her mother drinks alcohol, but generally in moderation and she does not use drugs. However, her mother is at greater risk of experiencing mental health difficulties compared to other people in Ireland. Kirsty's mother has realistic expectations for her and is an empathetic and nurturing parent. Kirsty is read to at least a few times a week and her parents use appropriate punishment when disciplining her and have a good knowledge of child development. Although her mother is not worried about her development or her behaviour, Kirsty is at risk for cognitive delays and atypical development. Yet she is showing signs of normal development in the realms of socio-emotional functioning, communication and gross motor skills.

Chapter One

Background of the *PFL* Programme Twelve Month Evaluation

1.1 Introduction

This report is the second in a series of reports which presents the result of the *PFL* evaluation. The report **'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, an explanation of the theoretical underpinnings of home visiting interventions. Also, a discussion of the outcomes at six months for other home visiting interventions, in addition to the results of the impact and implementation evaluation at six months was included. The present report focuses on information specific to the twelve month evaluation, including new measures collected as part of the twelve month interview, the results of the evaluation at twelve months, and new implementation data collected between six and twelve months. In addition, as there are now multiple waves of *PFL* data, the results of longitudinal analyses (dynamic analyses), which examine the impact of the programme on changes in child and parent outcomes over time, are included.

Chapter 1 of this report provides a brief summary of the recruitment process, analysis of baseline data, and the results of the evaluation at six months. It then presents a review of relevant findings from the literature on the impact of home visiting programmes at twelve months of age. Updated hypotheses on the likely impact of *PFL* on twelve month outcomes is then presented, as well as information regarding the collection of twelve 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, was available for 104 and 101 high and low *PFL* treatment group participants respectively, and 99 comparison group participants. Tests of baseline differences between the high and low *PFL* treatment groups found that the 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 '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 *Preparing For Life* programme was progressing well. In total, 257 six month interviews (n*Low* = 90; n*High* = 83; n*LFP* = 84) were completed. Analysis of the six month data across eight domains revealed there were limited significant differences reported between the high and low treatment groups (14%). This was 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 directly to information on the *PFL* Tip Sheets delivered to participants during this period.

Specifically, children in the high treatment group compared to those in the low treatment group had more appropriate eating patterns, had a higher level of immunisation rates, had more parental interactions, and parent-child interactions were of a higher quality. Additionally, children in the high treatment group were exposed to less parental hostility, a safer home environment, and more appropriate learning materials

and childcare. Moreover, mothers in the high treatment group were more likely to be socially connected in their community and less likely to be hospitalised after birth. The results of the multiple hypotheses tests strengthen these findings by showing that the high treatment group reported higher scores on the quality of the home environment and in the domain of maternal physical health, and lower scores on parental stress compared to the low treatment group. However, it is important to note that the programme had no significant impact on key factors such as pregnancy behaviour, child birth weight, breastfeeding, and child development. The interaction and subgroup analysis revealed that the programme had differential impacts with some groups benefiting more from the programme than others. For example, there was suggestive evidence that the programme benefited mothers with relatively higher cognitive resources, mothers with multiple children, and families who have experienced familial risk.

These lack of effects may be attributed to dosage and timing. 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. These results were also supported by the findings from the qualitative interviews which highlighted the small changes in behaviour and attitudes in the participants witnessed by the mentors. They acknowledged that these changes, while small, may be indicative of cumulative effects for the parents, children and community in the future. Despite these relatively modest effects, the low level of attrition (10% dropped-out and 8% disengaged) and high participant satisfaction were indications that programme engagement was high which may result in positive future outcomes.

The results comparing the high and low treatment groups to the comparison community largely confirmed the main treatment effect results, as well the integrity of the RCT design. The high treatment group differed from the comparison group in a positive direction on 21% of the measured analysed. While the comparison of the low treatment and comparison groups suggest that, as expected, the *PFL* programme is not having a significant impact on the outcomes of the low treatment group (only 11% of the differences between the low treatment group and comparison group were significant in a positive direction). This finding echoes the results of the contamination analysis which suggest that despite the high risk of contamination within the community between the high and low treatment groups, contamination was not a significant issue at this stage of the study.

1.4   Evidence on Short-term Effectiveness of Home Visiting Programmes

The six month report reviewed the evidence on the effectiveness of home visiting programmes on outcomes observed up to six months of age. This section reviews the evidence on outcomes reported between six and twelve months. Overall, the evidence on the early effectiveness of home visiting programmes between six and twelve months is mixed. Many studies find no significant differences between intervention and control groups during this period (Anisfeld, Sandy, & Guterman, 2004; Barlow et al., 2007; Barnes, Senior, & MacPherson, 2009; Duggan et al., 2007; Duggan, Fuddy et al., 2004; Duggan, McFarlane et al., 2004; El-Kamary et al., 2004; LeCroy & Crysik, 2011; Olds et al., 2002; Wagner, Cameto, & Gerlach-Downie, 1996; Wagner & Spiker, 2001). While some studies identify significant improvements in parent and child outcomes at twelve months of age (Barnes-Boyd, Norr, & Nacion, 1996; Culp et al., 2004; DuMont et al., 2008; Jungman et al., 2012; Koniak-Griffin et al., 2002; Mitchell-Herzfeld, Izzo, Greene, Lee, & Lowenfels, 2005), others find significant negative outcomes for parents and families (Duggan et al., 1999).

Table 1.1 reflects the impact of home visiting programmes on twelve month outcomes 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.

In addition, the *PFL* evaluation team conducted an extensive literature search according to the criteria outlined by HomVee and added any additional relevant studies. The table consists of findings on outcomes measured at twelve months postpartum from the sources after the year 1989.

Several studies identified in a review by Avellar and Paulsell (2011) found significant outcomes for home visiting programmes at twelve months postpartum in a variety of domains. For instance, in the domain of child health evaluations of the Resources, Education and Care in the Home (REACH) and Early Intervention Program for Adolescent Mothers (EIP) home visiting programmes found that children in the programmes had fewer hospitalisations at twelve months, were more likely to be adequately immunised (Koniak-Griffin et al., 2002) and experienced fewer illnesses in the first year of life than children in the control group. Multiple evaluations of the Healthy Families America programme found that participation in a home visiting intervention was associated with lower frequencies of physical aggression and abuse and a reduction in the prevalence of neglect (Duggan et al., 1999; DuMont et al., 2008; Mitchell-Herzfeld et al., 2005). However, in the domain of child development and school readiness none of the programmes identified by Avellar and Paulsell (2011) found significant outcomes for home visiting programmes at twelve months postpartum. In the domains relating to the parent and family including maternal health, positive parenting practices, family economic self-sufficiency, and linkages and referrals, several significant outcomes were identified (Avellar and Paulsell, 2011). Investigations of the Parents as Teachers programme found that parents in the intervention group demonstrated higher levels of parental responsivity and were more likely to provide their child with appropriate play materials (Wagner et al., 1996). In addition results from the Healthy Families America programme suggest that parents in the intervention group were less likely to engage in negative disciplinary practices than those in control group at twelve months (LeCroy & Crysik, 2011). However, evaluations of this programme also found that at twelve months the intervention group demonstrated lower levels of maternal mental health and maternal employment than the control group (Duggan et al., 1999; Mitchell-Herzfeld et al., 2005).

Other programmes, not identified by Avellar and Paulsell (2011), also identified significant outcomes in these domains. Results from a German pilot home visiting project, Pro Kind, found that at twelve months there was a significant difference in parental self-efficacy, attachment and perceived stress, with parents in the intervention group reporting more favourable levels of self-efficacy and attachment and lower levels of stress (Jungmann et al., 2011). Jungmann et al. (2011) also found that parents in the intervention group were more likely to use contraception at twelve months than those in the control group. Culp et al. (2004) investigated the Community-Based Family Resource Service (CBFRS) in the US and found significant differences at twelve months regarding parents' understanding of developmental expectations and the use of non-corporal punishment. This study also found that mothers were more accepting of their children, the home environment was safer, and parents were more likely to utilise community services at twelve months (Culp et al., 2004). Furthermore, an evaluation of a home visiting intervention in the UK based on the Family Partnership Model found that parents in the intervention group were more likely to demonstrate sensitivity towards their children and their children were more likely to be cooperative at twelve months (Barlow et al., 2007).

High-quality published studies of other home visiting programmes, such as the Nurse Family Partnership (NFP) and Home-Start UKHFA either did not report any significant early outcomes or did not assess early outcomes (Barnes et al., 2009; Olds et al., 2002).

Table 1.1 Evaluations of Early Outcomes for Home Visiting Programmes at 12 Months.

Outcome	Author	Sample Size	Programme
Child Health	Duggan et al. (1999)	564	Healthy Families America
	Mitchell-Herzfeld et al. (2005)	1061	Healthy Families America
	Barnes-Boyd et al. (1996)	372	Resources, Education and Care in the Home (REACH)
	Koniak-Griffin et al. (2002)	102	Early Intervention Program for Adolescent Mothers (EIP)
	Wagner et al. (1996)	236	Parents as Teachers (PAT)
Child Development & School Readiness	Anisfeld et al. (2004)	348	Healthy Families America
	Duggan et al. (1999)	564	Healthy Families America
	Jungman et al. (2011)	367	ProKind
	Wagner et al. (1996)	236	Parents as Teachers (PAT)
	Barlow et al. (2007)	131	Family Partnership Model
Family Economic Self-Sufficiency	Anisfeld et al. (2004)	348	Healthy Families America
	LeCroy & Crysik, (2011)	171	Healthy Families America
	Mitchell-Herzfeld et al. (2005)	1061	Healthy Families America
	Olds et al. (2002)	425	Nurse Family Partnership (NFP)
	Wagner et al. (1996)	236	Parents as Teachers (PAT)
Linkages & Referrals	LeCroy & Crysik, (2011)	171	Healthy Families America
	Culp et al. (2004)	263	Community Based Family Resource Service Programmes (CBFRS)
Maternal Health	Anisfeld et al., (2004)	348	Healthy Families America
	LeCroy & Crysik, (2011)	171	Healthy Families America
	Duggan et al. (1999)	564	Healthy Families America
	El-Kamary et al. (2004)	564	Healthy Families America
	Mitchell-Herzfeld et al. (2005)	1061	Healthy Families America
	Jungman et al. (2012)	397	ProKind
	Barnes et al. (2009)	250	Home- Start UK
	Koniak-Griffin et al. (2002)	102	Early Intervention Program for Adolescent Mothers (EIP)
	Wagner et al. (1996)	236	Parents as Teachers (PAT)
	Barlow et al. (2007)	131	Family Partnership Model
Positive Parenting Practices	Anisfeld et al., (2004)	348	Healthy Families America
	LeCroy & Crysik, (2011)	171	Healthy Families America
	Duggan et al. (1999)	564	Healthy Families America
	Duggan, McFarlane et al. (2004)	558	Healthy Families America
	Mitchell-Herzfeld et al. (2005)	1061	Healthy Families America
	Culp et al., (2004)	263	Community Based Family Resource Service Programmes (CBFRS)
	Koniak-Griffin et al. (2002)	102	Early Intervention Program for Adolescent Mothers (EIP)
	Wagner & Spiker (2001)	344	Parents as Teachers (PAT)
	Wagner et al. (1996)	236	Parents as Teachers (PAT)
	Barlow et al. (2007)	131	Family Partnership Model
Reductions in Child Maltreatment	Duggan et al. (1999)	564	Healthy Families America
	Duggan et al. (2007)	268	Healthy Families America
	Duggan, McFarlane et al. (2004)	558	Healthy Families America
	DuMont et al. (2008)	1060	Healthy Families America
	Mitchell-Herzfeld et al. (2005)	1061	Healthy Families America
Reductions in Juvenil Delinquency, Family Violence, and Crime	Duggan, Fuddy et al. (2004)	564	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 at 12 Months	Effect
	Immunisations, Hospitalisations	None	None
	Health Insurance, Primary Caregiver, Medical Care, Breastfeeding, Well Child Visits	None	None
	Postneonatal Mortality Rate, Physical Problems, No. of Illnesses in 1st year of life	No. of Illnesses in 1st year of life	Favourable
	Hospitalisations, Immunisations	No. of days nonbirth related child hospitalisations	Favourable
	Medical Care	None	None
	ASQ (Ages & Stages Questionnaire), Bayley Scales of Child Development	None	None
	Bayley Scales of Child Development, Mental Development Index	None	None
	Cognitive development (MDI), Psychomotor development (PDI)	None	None
	Cognitive Development (DPII)	None	None
	Social and Emotional Adjustment (BITSEA), Development (BSID), Temperament (ITS)	None	None
	Maternal Employment/Education	None	None
	Maternal Education, School/Training	School/Training	None
	Family Welfare Receipt, Mother Employed, Education	Mother Employed	Unfavourable
	Months Employed, Welfare Receipt	None	None
	Education, Marital Status, Father Living in Household, Employment, Teen Mother Only Adult, Welfare Receipt	None	None
	Use of Resources	Use of Resources	Favourable
	Active Utilisation of Community Services	Active utilisation of community services	Favourable
	MSSI (Perceived Social Support), CES-D (Depressive Symptoms), PSI (Parenting Stress)	None	None
	Parenting Support, The Emotional/Social Loneliness Inventory, The Adult Hope Scale	None	None
	Community Life Skills Scale (CLSS); Maternal Social Support Index (MSSI), Substance Abuse, Depressive Symptoms (CES-D), General Mental Health (CTS), Parenting Stress (PSI)	General Mental Health (CTS)	Unfavourable
	Family Planning & Birth Control Practices	None	None
	Health Insurance, Depressive Symptoms (CES-D), Mastery (PSM), Smoking, Alcohol & Drug Use, Social Support (ISSB)	None	None
	Use of Contraception, Further Pregnancies, Perceived Stress, Partnership Satisfaction, Attachment, Self-Efficacy	Use of Contraception, Further Pregnancies, Perceived Stress, Attachment, Parental Self-Efficacy	Favourable
	Edinburgh Postnatal Depression Scale (EPDS), Structured Clinical Interview for Diagnostic and Statistical Manual (SCID)	None	None
	Depression (CES-D), Self Esteem (RSEI), Perceived Stress (PSS), Repeat Pregnancy Rate, Substance Use	None	None
	Family Planning, Contraception	None	None
	Self-esteem, Self-efficacy, Parenting Stress (PSI), Social Support	None	None
	NCAST (Nursing Child Assessment Satellite Training)	None	None
	Revised Parent-Child Conflict Tactics Scale, AAPI-2	Revised Parent-Child Conflict Tactics Scale (Verbal Aggression, Minor Corporal Aggression)	Favourable
	HOME Inventory (learning environment), NCAST (Nursing Child Assessment Satellite Training)	None	None
	Discipline strategies (CTS-PC), HOME Inventory	None	None
	Parenting Attitudes (AAPI), Child Safety Checklist, Knowledge of Child Development (KIDI), Parenting Practices (CTS2)	None	None
	Massachusetts Safety Checklist, Adult-Adolescent Parenting Inventory (AAPI), HOME inventory	Massachusetts Safety Checklist, Adult-Adolescent Parenting Inventory (AAPI: developmental expectations, noncorporal punishment); Parenting knowledge HOME Inventory (acceptance)	Favourable
	Nursing Child Assessment Teaching Scale (NCATS: mother's score, total score)	None	None
	Parental Knowledge (AAPI, PSOC), HOME Inventory	None	None
	HOME Inventory, Parenting Knowledge (KIDI), Sense of Competence (PSOC)	HOME Inventory (total score, parental responsivity, appropriate play materials)	Favourable
	Maternal Sensitivity and Child Cooperativeness (CARE index), Maternal Psychopathology (GHQ), Parenting Attitudes and Competence (AAPI, WBPB, Parenting Sense of Competence Scale), HOME Inventory	Maternal Sensitivity and Child Cooperativeness (CARE index)	Favourable
	Conflict Tactics Scale (CTS)	Conflict Tactics Scale (CTS; Neglectful Behaviour)	Favourable
	Child Protective Services Reports	None	None
	Aggressive Maternal Parenting Behaviours	None	None
	Abuse, Neglect, Aggression (CTS-PC)	Frequency of very serious physical abuse, minor physical aggression, psychological aggression and harsh parenting in the past week (CTS-PC)	Favourable
	Emergency Room Visits, CPS Reports	Prevalence of Neglect, Chronicity of Psychological Aggression (CPS)	Favourable
	Partner's psychological abuse (CTS)	None	None
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.			

1.5 Hypotheses

The primary aim of the *PFL* Programme is to change parental knowledge, attitudes, and feelings leading to improved parenting behaviour, which will then positively impact on child development, ultimately increasing a child's school readiness. *PFL* also hypothesises that the programme will have an effect on other child and family outcomes (e.g. social support, 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 on the primary and secondary outcomes at twelve months of age are informed by the evidence described above on the early impact of home visiting programmes. Results from previous studies indicate that at twelve months home visiting programmes have limited positive effects on child health and development. Consistent with these findings we expect that the impact of *PFL* at twelve months also will be limited. Based on findings from similar programmes, we hypothesise that negative indicators of child health such as hospitalisation and reported number of illnesses may be decreased by the *PFL* programme. Other studies further suggest that home visiting can positively impact maternal health and parenting. Therefore, we expect to find a programme effect in these realms. Further, consistent with previous findings, risky parenting behaviours, such as parent-child conflict and use of physical punishment, may be reduced by the *PFL* programme. Consistent with the evaluation design, for the comparison group results, we expect that that there will more significant differences in the outcomes of the high treatment group versus the comparison group than in the outcomes of the low treatment group versus the comparison group. Given previous findings our hypotheses for results at 12 months are as follows:

- Children in the high treatment group will have fewer heath concerns than children in the low treatment group, such as fewer hospitalisations and fewer reported illnesses,
- Mothers in the high treatment group will report better overall health than those in the low treatment group,
- Mothers in the high treatment group will be more likely to report healthy and appropriate parenting practices than those in the low treatment group, such that they will report less incidence of risky parenting behaviours, like parent-child conflict and use of physical punishment,
- For comparison group analyses, there will more significant differences in the outcomes of the high treatment group versus the comparison group than in the outcomes of the low treatment group versus the comparison group.

1.6 Description of Twelve Month Survey & Data Collection Process

Between July 2009 and March 2012, a third research interview was conducted by the *PFL* Evaluation Team, within two weeks before or two months after each *PFL* child reached twelve months of age. In total, 248 twelve month interviews (n*Low* 82; n*High* = 83; n*LFP* = 83) were completed. The average age of the target child at time of completion was 12.3 months old (*SD* = 2.7 weeks). Three *PFL* participants (n*Low* = 2; n*High* = 1) dropped out of the evaluation after completing the baseline and six month interviews, but prior to completing a twelve month interview. There were no comparison group drop outs at this time point. Fifty participants in total dropped out prior to the twelve month interview and an additional thirteen participants were disengaged. A comprehensive analysis of attrition rates may be found in Chapter 4 of this report.

The twelve month interviews lasted approximately one to one and a half hours and were conducted using a Computer Assisted Personal Interviewing (CAPI) technique on tablet laptops. The interviews were conducted by trained interviewers who were blinded to the participant's treatment assignment. Immediately prior to the interview, participants were asked to complete the MacArthur-Bates Communicative Development Inventory (CDI) on paper. Although home interviews were encouraged, participants had the option of conducting the interview either in her home or in a local community centre. The majority of participants completed the interview in their homes (79.5% in the high treatment group, 84.1% in the low treatment

group, and 92.7% in the comparison community). Each participant was given a €20 shopping voucher after the twelve month interview was completed as a thank you for taking the time to complete the interview. During the interview the interviewer asked some of the questions that were asked during the baseline interview and the six month interview as well as several new questions, particularly in relation to the *PFL* child. The repeated questions included family demographics and socio-economic profile, maternal physical and psychological health, substance use by the mother, family risk factors, parenting knowledge, use of childcare, child motor skills, cognitive development, behavioural and emotional functioning, temperament, and social emotional development, child health and routines. Questions new to the twelve month questionnaire included items related to child development including child communication and language skills, cognitive development and more information about the child's temperament and problem behaviours, child's diet, family environment, maternal future outlook, parenting resources and job stability.

The twelve month survey was divided into ten modules, each containing questions with a common theme.

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

Similar to the six month report, this report focuses on eight domains incorporating 24 categories and 147 outcome measures. The domains and categories within each domain are – child development (Ages Stages Questionnaire, Brief Child Toddler Social and Emotional Assessment, MacArthur-Bates Communicative Development Inventory, Temperament and Atypical Behaviour Scale, Developmental Profile-3, difficult temperament and mother's concern about child' language and behaviour development), child health (child physical health; mother's health decisions for her child; sleep routines and diet), parenting (Adult Adolescent Parenting Inventory 2, Knowledge of Infant Development), home environment (Family Environment Scale; social worker involvement), maternal health and wellbeing (maternal physical health; maternal mental health; current substance use; Future Outlook Inventory), social support (satisfaction with father involvement; social support measures), childcare (childcare measures), and household factors and socioeconomic status (household factor measures; parental education; maternal employment; paternal employment; household finances and expectations of future finances). Note that while the same domains as the six month report are reported, the measures included in each domain may differ as different questions are asked in sequential interviews.

1.7 | Aims & Overview of Report

The aims of this report are threefold. First, to determine whether the *PFL* programme had an impact on parent and child outcomes at and before twelve months, second, to examine the impact of the programme on changes in mother and child behaviour over time through a dynamic analysis comparing outcomes at baseline, six months and twelve months, and three, to provide a detailed review of implementation practices in the *PFL* programme regarding attrition, dosage, and participant engagement.

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 socio-economic status). It also presents the results from the interaction analyses examining whether the *PFL* programme has differential effects by gender, primiparous status, partner status, risk status, maternal cognitive resources and maternal emotional wellbeing. Finally it presents a summary the treatment groups and comparison group analysis. Chapter 3 presents the results from the dynamic analysis examining changes in child and parent outcomes over time. Chapter 4 presents an implementation analysis of the *PFL* Programme between programme intake and twelve months, including an analysis of contamination in the *PFL* Programme. Chapter 5 summarises and concludes the results from the impact and implementation analyses.

Chapter Two

Main Results High and Low Treatment Groups

2.1 Introduction

This chapter presents the main results comparing the twelve 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 twelve months are indicative of a programme effect. The analysis focused on eight main domains – child development, child health, parenting, home environment, maternal health and wellbeing, social support, childcare and household factors and socioeconomic status. Although both the six month and twelve month reports contain the same overarching eight domains, measures which focus on different aspects of these domains were utilised at each time point. Therefore, it is not possible to make a direct comparison between the findings from the two reports on some domains, most notably the parenting and home environment domains. This chapter contains relevant literature for the new measures which were not included in the six month report, and considers the relevance and effectiveness of previous home visiting programmes on these measures at twelve 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' as these will be referenced where relevant. This report can be found at the following website under publications: http://geary.ucd.ie/preparingforlife.

This chapter also presents a summary of the interaction analysis which was conducted to determine whether the impact of the *PFL* programme differed for certain groups of participants. Specifically, it examined whether the programme had a greater impact on girls or boys, primiparous or multi-primiparous mothers, lone or partnered parents, higher or lower cognitive resource mothers, families with higher or lower risk, and mother's with higher or lower emotional wellbeing.

The chapter proceeds as follows: Section 2.2 describes the methods used to conduct the analyses and information on how to interpret the outcomes tables presented in the report. Sections 2.3 to 2.11 present the results for each of the eight main domains under analysis. Section 2.12 presents the results of the interaction analysis. Section 2.13 presents a summary of the treatment and comparison group analyses.

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, including conditional permutation testing, the step-down procedure which is used for multiple hypothesis testing, and the procedure for dealing with missing data .

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 difference between each observed response and the average response. This sum is then divided by the total number of observations to derive the average difference between responses and the mean. It serves as a useful indication 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 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. Given that this is a twelve month comparison, low *p-values* (i.e., significant results) would be a positive outcome indicating that the high treatment group is outperforming the low treatment group, and the *PFL* groups are outperforming the comparison group. *p-values* are presented for significant differences only. Non-significant differences are denoted by *ns*. A significant result in the non-hypothesised direction is denoted by *s~*. 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 twelve month outcomes of the high treatment, low treatment and comparison 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).
Step-down Test *p*[2]	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 in that category 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 in that category. 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 *p-value*s, such that the measure with the smallest *p-value* is listed first and the outcome with the highest *p-value* 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. As 147 outcome measures are considered in this report, it is possible that we may reject some of these null hypothesis by chance (i.e. we may identify a significant difference between the high and low treatment groups on certain outcomes when there is, in fact, no significant difference). Multiple hypothesis testing allows us to test for the joint significance of multiple outcomes at the same time, thus minimising the likelihood of finding treatment effects that are false. The multiple hypothesis method we use is called the Step-down procedure. 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.
Effect Size *d*	*Effect size (d)* illustrates the magnitude of the difference between the groups. While the *p-value* allows the reader to determine whether or not there is a statistically significant difference between groups, it does not indicate the strength of the difference. As the strength of a relationship can provide valuable information, the effect size was calculated using Cohen's d. A Cohen's d ranging from 0.0 to 0.2 is deemed a small effect; values ranging from 0.2 to 0.8 represent a medium effect; and values greater than 0.8 illustrate a large effect (Gravetter & Wallnau, 2004).

2.3 | Child Development

During the second six months of life, children develop across a broad range of domains including cognition, communication and language, physical health and motor skills, and social and emotional understanding (Charlesworth, 2010). Children become more mobile and thus more capable of exploration (Bredekamp & Copple, 1997; Meggitt, 2007). They also begin to have a more active participation in games and activities and they demonstrate intention in their actions (Bremner et al., 1997). While all children follow the same general trajectory in their development there is huge variance in the timing of the acquisition of each new skill. For example, while the average child walks alone by twelve months, a healthy child can be expected to walk anytime between 8 and 18 months (WHO Multicentre Growth Reference Study Group, 2006).

COGNITIVE DEVELOPMENT, COMMUNICATION, PROBLEM SOLVING & LANGUAGE

As children develop they adapt their behaviour through interaction with the environment and then assimilate or accommodate new information into their way of thinking. According to Piaget, between eight and twelve months of age children begin to demonstrate intentional or goal directed behaviour (Berk, 2012). Cognitive development progresses rapidly in the second six months of life, by nine months most children have a sense of object permanence and will recognise familiar pictures. By twelve months most children will use trial and error to explore objects and will begin to treat objects differentially, for example hugging a teddy bear or attempting to use a hair brush correctly (Meggitt, 2007).

Language skills also begin to emerge during this time period and children's communicative intent becomes more established (Eadie et al., 2010; Woodward & Guajardo, 2002). Between six and twelve months children progress from producing unintentional communicative gestures, such as burping or vocalisations with no intended message, to purposefully using gestures and non-linguistic vocalisations (Capone, 2010). By nine months most children will understand simple directions, such as 'wave to teddy'. They may also imitate sounds that an adult makes and respond to the 'no' command (Meggitt, 2007). By approximately twelve months of age most children will recognise simple direction, like 'no', 'show me' or 'look'. At this age, many children will have produced their first words, and many will be able to speak two or more recognisable words (Capone, 2010; Eadie et al., 2010; Meggitt, 2007). However, age of initial language expression may not be an indicator of future language ability as there is huge variance in the timing and the number of first words. In general, vocabulary size is not used as an indicator for clinical risk of language delay until after 24 months of age (Eadie et al., 2010). However, other communication skills such as the use of gestures and the development of communicative intent may be more important for future language development than initial word production (Eadie et al., 2010; Capone, 2010). During this time children typically use pre-linguistic skills such as making eye contact and turn taking. These skills, learned from imitating caregiver behaviour, may serve as models for future linguistic development (Capone, 2010).

PHYSICAL DEVELOPMENT (GROSS MOTOR & FINE MOTOR)

Motor development is an important aspect of child development (Wijnhoven, et al., 2004) as an increase in mobility and precision facilitates further exploration of the environment (Bredekamp & Copple, 1997). During the second six months of life children progress through several stages of motor development that build to independent walking (WHO Multicentre Growth Reference Study Group, 2006). By nine months children are capable of maintaining a sitting position and can usually sit unsupported for a substantial amount of time. They tend to be more mobile and will move around the floor by crawling, shuffling or wriggling. They also show progression towards walking. They may be able to pull themselves up to a standing position and can generally stand while holding onto something. They may even take several steps while their hands are held (Meggitt, 2007). By the time children reach twelve months of age, most can rise to a standing position without help and many can stand alone for a brief time. Some children at this age may have taken their first steps. If not, older children can usually move around using the support of furniture or by crawling on their hands and knees (Meggitt, 2007).

As children progress from six months to one year of age, they develop more co-ordinated fine motor skills. By the time they reach twelve months most children can grasp small objects between their thumb and the tip of the index finger. They also become more deliberate in their actions and are capable of releasing a toy into an adult's hand, turning pages of a book and building bricks or arranging toys on the floor (Meggitt, 2007).

PERSONAL, SOCIAL, EMOTIONAL DEVELOPMENT & TEMPERAMENT

Personal, social and emotional skills are reflected in a child's ability to pay attention, transition from one activity to another and cooperate with others (Nisha, 2006). According to Erikson (1968) during the first year children learn to develop trust and security through their attachment with their primary caregiver. The acquirement of these skills is strongly influenced by child temperament, as temperament influences the child's reactivity and self-regulation in their interactions with individuals and their environment (Rothbart & Bates, 2006).

While it is often difficult to pinpoint milestones of social and emotional development, in general children begin to develop self-awareness and learn that they are separate from others and the environment in their first year of life (Nisha, 2006). As this occurs they discover that they can influence and are influenced by the people and the environment around them (Broadhead et al., 2010). Between six and nine months of age children may have the ability to clearly express different emotions. They may respond to language and gestures and will show displeasure when a favourite toy is taken away. By twelve months most children will be affectionate towards familiar people, but are generally shy around strangers. In line with their increased communication abilities they may begin to join in conversation and will play more interactive games (Meggitt, 2007). They will also mimic simple actions and may begin to engage in skills, such as self-feeding (Nisha, 2006).

IMPACT OF HOME VISITING PROGRAMMES ON CHILD DEVELOPMENT AT TWELVE MONTHS

Several studies assess the impact of programme participation on a variety of child developmental domains including physical development, gross and fine motor skills, cognitive development, language, communication, problem solving, and personal, social and emotional development (Barlow et al., 2007; Duggan et al., 1999; Jungman et al., 2011; Wagner et al., 1996). However, there are very few home visiting studies that identify an impact on child development outcomes at twelve months postpartum. The measures used to assess child development in these interventions include some of the standardised measures used in the *Preparing For Life* evaluation. However, as shown in Table 1.1 in Chapter One none of these studies found a significant programme effect for child development at twelve months.

2.3.1 | Child Development Instruments

AGES AND STAGES QUESTIONNAIRE

Child development in the *PFL* Evaluation was assessed using the twelve month version of the Ages and Stages Questionnaire (ASQ; Squires et al., 1999). The ASQ was designed as an effective screening measure for young children who were considered to be at risk for developmental delay. The ASQ child monitoring system consists of 19 screening questionnaires at specific age intervals ranging from four to 60 months of age and provides scores across five domains of child development, with each domain comprising six items. Communication (α=..48) measures the child's babbling, vocalisation, listening and understanding. The gross motor domain (α=..80) measures the child's arm, body and leg movements. The fine motor domain (α=..51) assesses the child's finger and hand movements. Problem solving (α=..53) measures the child's learning and playing with toys. Finally, the personal-social domain (α = .42) provides a rating of solitary social play with toys and other children. During the interview, the interviewer asked the participant questions related to different activities the child is capable of. The participant responded by indicating if her child exhibits the behaviour regularly, sometimes, or not yet. If the participant did not know whether her child was capable of the behaviour, where appropriate, 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 =15.8; gross motor=18.0; fine motor=28.4; problem solving=25.2; and personal-social=20.1). 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 on a domain are considered to be at risk of developmental delay in that domain. Furthermore, an ASQ standardised total score was calculated by creating a standardised score, 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, 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 the Ages and Stages Questionnaire: Social-Emotional (ASQ:SE; Squires, Bricker, & Twombly, 2003). The ASQ:SE (α=0.67) 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 participant questions related to different behaviours the child displays. The participant responded by indicating if her child exhibited the behaviour most of the time, sometimes, or never. Additionally, the participant 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 participant. Scores were summed to provide a total ASQ:SE score, with a possible range of 0 to 285. 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 48 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.

MACARTHUR-BATES COMMUNICATIVE DEVELOPMENT INVENTORIES: WORDS AND GESTURES (CDI-WG)

The MacArthur-Bates Communicative Development Inventories: Words and Gestures (CDI-WG: Fenson et al., 2000) short form is a parent report instrument for assessing language and communication skills in children. It provides norms for children aged eight to sixteen months of age. The CDI inventories measure a range of early communicative and representational skills that are related to language development in typically developing and language-delayed children. The CDI-WG consists of three sections; first signs of understanding, first communicative gestures, and a vocabulary checklist.

Participants were asked to complete the CDI-WG with pen and paper before beginning the main part of the interview. The first section, 'first signs of understanding', contained 3 questions with the response options yes or no. In the second section, 'first communicative gestures', there were 12 questions with the response options not yet, sometimes and often. The final section contained an 89-word 'vocabulary checklist' with 3 separate columns; understands, understands and says and does not understand or say. First signs of understanding (α = not available due to lack of variation/variables), and first communicative gestures (α=.94) each produced a summed raw score. The vocabulary checklist columns understands and understands and says provided 2 scores: words understood (α=.96) and words produced (α=.96). These were then normed by age and gender, according to Fenson et al. (2000). In total, the CDI produces four scores.

BRIEF CHILD-TODDLER SOCIAL AND EMOTIONAL ASSESSMENT (BITSEA)

The Brief Child-Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2006) is a 42-item screening tool for social-emotional/behavioural problems and delays in competence in children aged twelve months to thirty-six months. This version is a shortened version of the Child-Toddler Social and Emotional Assessment (ITSEA). The BITSEA yields a Problem score (α=.79) and a Competence score (α=.65). Problem behaviour (α=.82) items include externalising (6 items), internalising (8 items) and dysregulation problems (8 items). Competencies (α=.59) include areas of attention, compliancy, mastery,

motivation, pro-social peer relations, empathy, play skills and social relatedness (11 items). The interviewer asked participants 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 total score, with higher Problem scores indicating greater levels of social-emotional or behavioural problems and lower Competence scores indicating possible delays/deficits in competence. These scores were normed by child gender. In addition, the BITSEA provides cut-off scores which indicate a 'Possible Problem' (25th percentile) or 'Possible Deficit/Delay' (15th percentile) respectively. There are different cut-off scores for different ages: twelve to seventeen months, eighteen to twenty three months, twenty four to twenty nine months and thirty to thirty six months. At twelve months, the cut-off scores are a score greater than 12 for the Problem score and less than 12 for the Competence score, which suggest that children with scores above/below these cut-offs respectively may be at risk. In line with these cut-off scores, binary variables were calculated to illustrate if the child was displaying potential problem or competence difficulties.

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 included the DP-3 cognitive section which measures cognitive abilities in an indirect manner (α=.88). This is a 38 item scale, starting at number 1 and continuing until the stop rule is satisfied (i.e. when five consecutive no responses are recorded). Each of the items refer to tasks which require cognitive skill and are arranged in order of difficulty, for example, 'When an adult points to something, does the child usually look where the adult has pointed?'. For each item, participants 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. In addition, a binary variable was created to indicate those above the average score, that is, a score of above 115.

TEMPERAMENT AND ATYPICAL BEHAVIOUR SCALE

Temperament and Atypical Behaviour Scale (TABS; Neisworth, Bagnato, Salvia & Hunt, 1999) is a 15 item norm referenced screening tool for temperament and regulation problems, which can identify a child's risk for an overall developmental delay (α=.82). These overall developmental delays include learning disability, Autism and Foetal Alcohol Syndrome. It is suitable for children age 11 months to 5 years, 11 months. In the *PFL* interview, the interviewer listed 15 target behaviours and asked participants to verbally reply 'yes' or 'no' indicating whether these behaviours were a problem for their child. 'Yes' scores were summed and any score, other than zero, was indicative of atypical development. In clinical practice, the TABS Assessment tool would then be recommended to evaluate the nature of the concern. In line with these scoring guidelines, a binary was calculated to illustrate if the child was at risk of atypical development.

DIFFICULT TEMPERAMENT

Seven items were used to assess the temperament of the *PFL* child. Specifically, the participant was asked questions related to how well her child behaves and was instructed to answer each item on a zero to six point scale. These seven items (α=.68) were taken from the Quebec Longitudinal Study of Child Development (QLSCD) and are originally based on the Child Characteristics Questionnaire (Bates, Freeland, & Lounsbury, 1979). Scores to each item were summed providing a total possible range of 0 to 42, with higher scores indicative of more difficult child temperament.

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, four of the six child development measures were in the hypothesised direction and one of these, ASQ Fine Motor Score, was statistically significant. The high treatment group scored an average of 54.33 on this subscale while the low treatment group scored an average of 51.87 ($p<.05$, $d=.26$) indicating that children in the high treatment group were more likely to engage in developmentally appropriate finger and hand movements than children in the low treatment group. The step-down test showed that the joint effect of all six measures in the ASQ Scores category was not statistically significant.

ASQ CUT-OFF SCORES

Within the ASQ Cut-off Scores category, which measures the proportion of children at risk of development delay in each group, five of the six measures were in the hypothesised direction. These differences were not statistically significant for any of the six measures. In addition, the step-down test showed that the joint effect of all six measures in the ASQ Cut-off Scores category was not statistically significant.

MACARTHUR-BATES COMMUNICATIVE DEVELOPMENT INVENTORY (CDI)

Within the CDI category, one of the four variables, CDI: First signs of understanding was in the hypothesised direction. There was one variable which was statistically significant in the non-hypothesised direction. Children in the low treatment group scored an average of 9.70 in the CDI: First Communicative Gestures variable, compared with an average score of 8.84 for children in the high treatment group ($p<.01$, $d=.37$). This indicates that children in the low treatment group use more communicative gestures than children in the high treatment group.

BITSEA

Within the BITSEA category both the Competence Score and the Problem Scores were in the hypothesised direction. However, these differences were not statistically significant. Furthermore, the step-down test showed that the joint effect of the BITSEA scores was not statistically significant.

BITSEA CUT-OFF SCORES

Within the BITSEA cut-off score category both the Competence Cut-off Score and Problem Cut-off Score were in the hypothesised direction, and the BITSEA Competence Cut-off Score was statistically significant. 16% of children in the high treatment group were below the Competence Score, compared with 28% of children in the low treatment group ($p<.05$, $d=.29$). In addition, the step-down test showed that the joint effect of the two measures in the BITSEA Cut-off Score category was statistically significant ($p<.10$) indicating that there is a significant difference between the expression of social and emotional difficulties for low and high treatment group children. The joint effect finding is driven by the significant results found for the BITSEA Competence Cut-off Score.

NON STEP-DOWN MEASURES

All six of the non step-down measures were in the hypothesised direction. There were no statistically significant differences between the high and low treatment groups.

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

	Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Individual Test p^1	Step-down Test p^2	Effect Size d
ASQ Scores										
	ASQ Fine Motor Score	165	(82/83)	54.33	(8.63)	51.87	(10.29)	$p<.05$	ns	0.26
	ASQ Personal Social Score	165	(82/83)	49.88	(8.82)	48.55	(10.46)	ns	ns	0.07
	ASQ Gross Motor Score	165	(82/83)	42.07	(18.34)	40.72	(18.27)	ns	ns	0.07
	ASQ Problem Solving Score	165	(82/83)	46.40	(11.71)	46.33	(13.07)	ns	ns	0.01
	ASQ Communication	165	(82/83)	49.88	(10.74)	50.18	(10.55)	ns	ns	0.14
*	ASQ Social-Emotional Score	165	(82/83)	23.48	(21.51)	21.14	(16.05)	ns	ns	0.12
ASQ Cut-off scores										
*	ASQ Problem Solving Cut-off	165	(82/83)	0.06	(0.24)	0.11	(0.31)	ns	ns	0.17
*	ASQ Communication Cut-off	165	(82/83)	0.00	(0.00)	0.01	(0.11)	ns	ns	0.16
*	ASQ Personal Social Cut-off	165	(82/83)	0.01	(0.11)	0.04	(0.19)	ns	ns	0.16
*	ASQ Gross Motor Cut-off	165	(82/83)	0.11	(0.31)	0.14	(0.35)	ns	ns	0.10
*	ASQ Fine Motor Cut-off	165	(82/83)	0.02	(0.16)	0.04	(0.19)	ns	ns	0.07
*	ASQ Social-Emotional Cut-off	165	(82/83)	0.10	(0.30)	0.07	(0.26)	ns	ns	0.09
MacArthur-Bates Communicative Development Inventory (CDI)										
	First Signs of Understanding	151	(76/75)	2.97	(0.16)	2.96	(0.20)	ns	ns	0.08
	Vocabulary Words Produced NORM	165	(82/83)	55.27	(35.57)	55.66	(32.44)	ns	ns	0.12
	Vocabulary Words Understood NORM	165	(82/83)	74.43	(23.79)	77.71	(20.24)	ns	ns	0.15
	First Communicative Gestures	162	(81/81)	8.84	(2.30)	9.70	(2.02)	s~	ns	0.37
Brief Child-Toddler Social and Emotional Assessment (BITSEA)										
	BITSEA Competence Score	165	(82/83)	15.39	(3.38)	14.84	(3.56)	ns	ns	0.16
*	BITSEA Problem Score	165	(82/83)	8.82	(5.74)	8.90	(6.49)	ns	ns	0.01
BITSEA Cut-off scores										
*	BITSEA Competence Score Cut-off	165	(82/83)	0.16	(0.37)	0.28	(0.45)	$p<.05$	$p<.10$	0.29
*	BITSEA Problem Score Cut-off	165	(82/83)	0.22	(0.42)	0.24	(0.43)	ns	ns	0.05
Non Step-down Measures										
	ASQ Standardised Total Score	165	(82/83)	100.08	(14.34)	98.09	(16.60)	ns	-	0.13
	DP3: Cognitive Development Standardised Score	165	(82/83)	116.20	(13.66)	115.13	(16.03)	ns	-	0.07
	DP3: Cognitive Development Above Average Cut-off	165	(82/83)	0.54	(0.50)	0.52	(0.50)	ns	-	0.04
*	TABS Score (a score of > 0 => risk of atypical development)	165	(82/83)	0.95	(1.74)	1.22	(2.01)	ns	-	0.14
*	TABS Score Binary (risk of atypical development)	165	(82/83)	0.33	(0.47)	0.40	(0.49)	ns	-	0.14
*	Difficult Temperament	165	(82/83)	12.66	(5.52)	13.30	(5.76)	ns	-	0.11

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

2.4 | Child Health

A number of evaluations of home visiting interventions have identified positive effects on child health such as a decrease in the number of child illnesses (Barnes-Boyd et al., 1996) and a reduced rate of hospitalisation (Koniak-Griffin et al., 2002). Information specific to child health at twelve months of age and the impact of home visiting interventions on such measures is discussed below.

BREATHING PROBLEMS

On average, children typically have seven respiratory illnesses in their first year (Bee, 1995) and the severity of these illnesses can negatively impact a child's development. One study found that 35% of twelve month old children were reported to have mild asthma-like symptoms, 11% moderate and 4% severe symptoms (Mohangoo et al., 2012). Furthermore, child quality of life was associated with the severity of symptoms, with more severe symptoms resulting in a lower quality of life. There is little research on the effects of home visiting interventions on breathing problems or rates of asthma very early in life.

CHILD WEIGHT

According to the WHO growth charts, the average weight for twelve month old children is 8.94kg for girls and 9.64kg for boys (http://www.cdc.gov/growthcharts/who_charts.htm). There are a number of difficulties associated with being overweight and underweight during infancy. Being overweight during infancy may negatively impact an child's gross motor skill development (Shibli, 2008), while being underweight during this period is associated with longer term effects, such as a higher risk of developing metabolic and cardiovascular disease in later life (Fabricius-Bjerre et al., 2011). There are no known evaluations that measure the impact of home visiting interventions on child weight at twelve months.

IMMUNISATIONS

In Ireland, children are scheduled to receive vaccinations for MMR (measles, mumps and rubella) plus a vaccine against Pneumococcal Disease at twelve months of age. These vaccinations are provided free of charge in local GP surgeries and health clinics (www.citizensinformation.ie). There is mixed evidence regarding the impact of home visiting programmes on immunisation rates at twelve months old, with some studies reporting positive effects on immunisation uptake (Koniak-Griffin et al., 2002) and others not (Duggan et al., 1999).

CHILD SLEEPING

The literature regarding child sleeping at twelve months is mixed, with some studies reporting that almost all children sleep through the night by six months (Eaton-Evans & Dugdale, 1988; Henderson et al., 2010), while others report that few children sleep undisturbed for 12 hours (Anders, Halpern & Hua, 1992). It is suggested that while most children may wake during the night, most instances of waking do not require parental attention to return to sleep.

In relation to sleeping location, there is much debate about the appropriateness of co-sleeping and the cultural context of this (Ball, 2003). One study found that the consistency of sleep location was related to positive mother-child interactions, rather than the actual sleep location (Taylor, Donovan & Leavitt, 2008). Child sleeping patterns (including location) have not been shown to be affected by home visiting interventions at twelve months.

CHILD FEEDING

Finally, in relation to diet, it is recommended that a varied diet is introduced to children from 6 months of age, and breastfeeding is recommended alongside complementary food up to 24 months of age (Lande et al., 2003; WHO, 2009(a)). Few studies report the effects of home visiting interventions on feeding habits at twelve months; however a study by Mitchell-Herzfeld et al. (2005) reported no impact on breastfeeding rates at twelve months.

| 2.4.1 | Child Health Instruments |

GENERAL HEALTH

A number of variables were used to assess child health. A variable representing the overall general health of the child was asked with response options given on a five point scale ranging from excellent to poor. This measure was dichotomised to create a binary variable denoting whether the child had ill health (poor, fair) or not (good, very good, excellent). The number of health problems the child had in the last six months was assessed by asking the participant whether her child had ever been taken to the GP, Health Centre, or Casualty for any problems on a list of 13 possible options. A variable denoting the total number of health problems the child had was created by summing the number of child health problems endorsed by the mother. Two binary variables were created to determine whether the child had a chest infection or asthma in the last six months. A binary variable was also created based on whether or not the child had ever stayed overnight in hospital in the last six months for any illness.

MOTHER'S HEALTH DECISIONS FOR HER CHILD

Participants were asked about the vaccinations their child had received. Two binary variables denoting whether the child had received all recommended vaccinations up until six and twelve months were created. In addition, a binary variable indicating whether the child had ever been vaccinated was created. Participants were asked if they knew their child's current weight, and if so, what it was. A binary variable was created to indicate whether the mother knew her child's weight or not, and a variable of child weight (in kilograms) was created.

SLEEPING ROUTINE

Participants were asked whether their child had any difficulty falling asleep at night. The response options to this question were never, sometimes, often, or always. A binary variable denoting whether or not the child ever had any difficulty falling asleep was constructed (never versus sometimes/often/always). Participants were asked whether their child slept undisturbed through the night most of the time and a binary variable was created indicating whether the child slept through the night. Participants were also asked to rate on a three point scale how much of a problem their child awakening has been. A binary variable was created indicating whether it had been not much of a problem or somewhat/quite a bit of a problem. Finally, participants were asked where their child generally sleeps. The four possible response options were divided into a binary variable with the 'appropriate' category including, Alone in his/her bedroom, Shares his/her bedroom with another family member (brother, sister, grand-parents etc.) and In your bedroom BUT not in your bed and the 'inappropriate' category including In your bedroom AND in your bed.

APPROPRIATE FOOD

Participants 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 daily'. The sugars and fats category was reverse scored to indicate that more of these foods was not beneficial. Participants were asked to name both their child's favourite things to drink, and drinks that their child would not drink. The appropriateness of these drinks was coded and a binary was created with water and milk being rated as appropriate, while sugary drinks were not.

BREASTFEEDING

Participants were asked several questions relating to breastfeeding including whether they ever tried to breastfeed their child, if they were still breastfeeding, and at what age they stopped both exclusively and non-exclusively breastfeeding (that is they continued to breastfeed their child but also gave him/her other food or drink). They were also asked whether their child drank anything other than breast milk, where relevant. The first two questions were simple yes/no questions while the latter two questions were continuous variables with response options given in weeks, months, or days. In order to make responses similar across all participants, the answers to both questions were converted into weeks.

2.4.2 Child Health Results

CHILD HEALTH

Four of the five measures in the Child Health category were in the hypothesised direction. There was one statistically significant difference between the high and low treatment groups regarding whether the child ever had a chest infection. 24% of children in the high treatment group were reported to ever have had a chest infection, compared with 34% in the low treatment group ($p<.10$, $d=.21$). Overall, the step-down test showed that the joint effect of all five measures in the Child Health category was not statistically significant.

MOTHER'S HEALTH DECISIONS FOR HER CHILD

Within the Mother's Health Decisions for her Child category, two of the three measures showed differences in the hypothesised direction and one was statistically significant, immunisations at twelve months. 27% of children in the high treatment group had received all vaccinations up to twelve months, compared with 14% of children in the low treatment group ($p<.05$,$d=.31$). Overall, the step-down test showed that the joint effect of all three measures in the Mother's Health Decisions category was not statistically significant.

SLEEP ROUTINE

One of the four measures in the Sleep Routine category was in the hypothesised direction. There was, however, one statistically significant difference between the high and low treatment groups in a non-hypothesised direction regarding the appropriateness of the child's sleeping location. 95% of low treatment group parents reported an appropriate sleeping location for their children, compared with 87% of children in the high treatment group ($p<.05$, $d=.30$). This indicated that children in the high treatment group were more likely to have an inappropriate sleeping location such as sleeping in their parent's bed. The step-down test showed that the joint effect of the four measures in the Sleep Routines category was not statistically significant.

APPROPRIATE FOOD

Five of the seven measures in the Appropriate Food category were in the hypothesised direction, two of which were statistically significant. Children in the high treatment group were reported to eat more grains ($p<.05$, $d=.29$) and more dairy ($p<.05$, $d=.29$) than children in the low treatment group. The step-down test showed that the joint effect of the seven measures in the Appropriate Food category was not statistically significant.

NON STEP-DOWN MEASURES

Three of the four measures which were not included in the above Step-down categories were in the hypothesised direction, however none were statistically significant.

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										
*	Ever had chest infections	165	(82/83)	0.24	(0.43)	0.34	(0.48)	$p<.10$	ns	0.21
*	No. of health problems taken to GP/ health centre/casualty	164	(81/83)	1.31	(1.41)	1.46	(1.25)	ns	ns	0.11
	Child's health in the last 6 months	165	(82/83)	0.94	(0.24)	0.92	(0.28)	ns	ns	0.09
*	Ever had asthma	165	(82/83)	0.11	(0.31)	0.13	(0.34)	ns	ns	0.07
*	Ever stayed in hospital for at least one day	165	(82/83)	0.06	(0.24)	0.06	(0.24)	ns	ns	0.00
Mother's Health Decisions for her Child										
	Necessary immunisations at 12 months	165	(82/83)	0.27	(0.45)	0.14	(0.35)	$p<.05$	ns	0.31
	Necessary immunisations at 6 months	165	(82/83)	0.99	(0.11)	0.96	(0.19)	ns	ns	0.16
	Mother knows child's current weight	165	(82/83)	0.20	(0.40)	0.22	(0.41)	ns	ns	0.03
Sleep Routine										
*	Child awakening a problem	151	(76/75)	0.27	(0.45)	0.29	(0.46)	ns	ns	0.05
*	Difficulty falling asleep	165	(82/83)	0.63	(0.48)	0.63	(0.49)	ns	ns	0.02
	Sleeps undisturbed through the night	165	(82/83)	0.66	(0.48)	0.71	(0.46)	ns	ns	0.11
	Appropriateness of sleeping location	162	(81/81)	0.87	(0.34)	0.95	(0.22)	s~	ns	0.30
Appropriate Food										
	How often the child eats grains	165	(82/83)	6.36	(1.01)	6.01	(1.39)	$p<.05$	ns	0.29
	How often the child eats dairy	165	(82/83)	6.57	(1.31)	6.20	(1.22)	$p<.05$	ns	0.29
*	How often the child eats other food (e.g. sugars and fats, etc., sweets, crisp, chips.)	164	(81/83)	4.37	(1.88)	4.72	(1.82)	ns	ns	0.19
	How often the child eats protein	165	(82/83)	5.63	(1.09)	5.47	(1.32)	ns	ns	0.14
	How often the child eats vegetables	165	(82/83)	6.04	(0.84)	6.02	(0.92)	ns	ns	0.01
	Appropriateness of drinks	160	(80/80)	0.63	(0.49)	0.63	(0.49)	ns	ns	0.00
	How often the child eats fruits	165	(82/83)	6.20	(1.35)	6.20	(1.42)	ns	ns	0.01
Non Step-down Measures										
	Current Weight (kilograms)	33	(17/16)	10.08	(1.57)	10.28	(2.10)	ns	-	0.11
	Ever vaccinated	165	(82/83)	0.99	(0.11)	0.96	(0.19)	ns	-	0.16
	Age (in weeks) exclusive breastfeeding ends	40	(20/20)	5.95	(9.47)	3.84	(5.12)	ns	-	0.28
	Age non-exclusive breastfeeding ends	40	(20/20)	8.07	(12.31)	4.66	(6.34)	ns	-	0.36

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

2.5 | Parenting

PARENTING RISK FACTORS FOR MALTREATMENT AND ABUSE

Traditionally, research in the field of parenting has focused on the conceptualisation of parenting patterns and has identified parenting styles based on parents' relative use of each of these dimensions 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 high control are commonly associated with negative developmental outcomes (Petito & Cummins, 2000). Promoting sensitive and responsive parenting to high risk families may support positive development for children who are at increased risk for poor developmental outcomes, as well as prevent parental abuse and neglect. To this effect, research has demonstrated that at risk mothers who participated in home visiting programmes during pregnancy displayed lower risk of potential child abuse when the children were 3 months old compared to comparison mothers not receiving a home visiting intervention (Guthrie, Gaziano, & Gaziano, 2009).

PARENTAL KNOWLEDGE OF CHILD DEVELOPMENTAL NORMS AND PROCESSES

Parental understanding of child behaviour may affect the way child behaviour is interpreted by the parent (Mills & Rubin, 1990) and it has been argued that knowledge of typical child behaviour has the ability to influence parent-child interactions (see Goodnow, 1988 for review). Furthermore, parental knowledge of child development is consistently found to be lower among parents living in low socioeconomic environments (McLoyd, 1998), parents experiencing depressive symptoms (Cunningham & Boyle, 2002), and primiparous parents (Pleck, 1997). Therefore, increasing maternal knowledge of child development has the potential to positively influence child development.

Parenting knowledge and attitudes are a common outcome measure used to assess the impact of home visiting interventions. A meta-analysis by Kendrick et al. (2000) reported that although some evaluations of home visiting programmes had a significant impact on parenting, in relation to parental knowledge of developmental milestones, many did not. One study by Culp et al. (2004) found that programme effects on parental knowledge of infant development can be seen as early as six and twelve months. However, many others suggest that improvements in parenting knowledge are not evident at twelve months (Barlow et al., 2007; Mitchell-Herzfeld et al., 2005; Wagner & Spiker, 2001).

READING WITH CHILD

A meta-analysis found that joint book reading between parents and children was related to children's language growth, emergent literacy and reading achievement, and that the frequency of book reading was related to effect size (Bus, van IJzendoorn & Pellegrini, 1995). One study which investigated the relationship between shared parent-child book reading at four and eight months and subsequent language abilities at twelve and sixteen months found that shared reading at eight months was related to language abilities at twelve month and also language abilities sixteen month which were over and above the twelve month scores, however shared reading at four months was not significantly related to later language development (Karrass & Braungart-Rieker, 2005).

2.5.1 | Parenting Instruments

ADULT ADOLESCENT PARENTING INVENTORY 2 (AAPI-2)

Parenting risk of abuse and neglect was assessed using the Adult Adolescent Parenting Inventory 2 (AAPI-2; Bavolek & Keene, 1999). This 40-item measure assesses the parenting and child-rearing attitudes of adult and adolescent parent and non-parent populations. Based on the known behaviours of abusive parents, responses to the AAPI-2 provide an index of risk for practicing parenting behaviours known to contribute to the maltreatment of children. The AAPI-2 yields scores on five subdomains, the names of which have been adapted to indicate positive parenting, these include realistic parental expectations of children (7 items; α=.76), parental empathy (10 items; α=.78), belief in the use of appropriate punishment (11 items; α=.75), appropriate parent-child roles (7 items; α=.82), and promoting children's power and

independence (5 items; α=.33). Participants were asked to rate how much they agreed or disagreed with a series of questions regarding parenting on a five point Likert scale ranging from 1 meaning strongly agree to 5 representing strongly disagree. Raw scores for the five subdomains were converted to standard scores, ranging from one to ten. The cut-off scores for each 'at risk' domain is a score of less than or equal to 3.

In addition to these five subdomains, the AAPI-2 provides an overall score of parenting risk (40 items; α=.86) that is presented as an average of the standard scores for each subdomain. Higher scores, i.e. from 8 to 10, on the AAPI-2 are indicative of lower risk for abusive parenting, such that higher scores are representative of positive, nurturing, parenting attitudes and a low risk of abuse. Standard scores ranging from 1 to 3 are considered to be low and are indicative of high risk for abusive parenting and neglect. Standard scores ranging from 4 to 7 represent the normal range of parenting behaviours and illustrate moderate risk for parenting abuse and neglect. A binary variable was also created using a cut-off score which is indicative of high risk parenting. An additional continuous variable was calculated to represent the total number of scales on which participants scored in the at risk range (i.e., 1 to 3).

KNOWLEDGE OF INFANT DEVELOPMENT (KIDI)

Maternal knowledge of child development was assessed using the 14-item Knowledge of Child Development – Short Form (KIDI-SF; MacPhee, 1981), a measure designed to assess knowledge of developmental processes and child developmental norms. Participants were presented with 14 items (α=.34) related to child developmental milestones and norms and were asked how much they agree or disagree with each statement. Response options ranged from one representing strongly agree to five signifying strongly disagree. Responses were summed, providing a range of scores from 14 to 70. An indicator of knowledge of child development was obtained from these raw scores and is represented as the proportion of accurate responses about child development or the raw score divided by the total possible number of points (i.e., 70). This figure ranges from 0 to 100 and can be interpreted as an indicator of maternal knowledge of child development with higher scores representing greater knowledge. In addition to this continuous score, a binary variable was created to represent the proportion of participants who scored in the lowest 10% of the entire *PFL* cohort on the KIDI-SF.

READING WITH CHILD

Participants were asked whether they read to their child. If they said yes, they were asked how often, with responses of everyday, 2 or 3 times a week, once a week, once a month or less than once a month. A binary variable was created indicating whether the participant read to their child everyday or not.

CONCERN ABOUT CHILD'S DEVELOPMENT

Participants were asked how worried they were about their child's language development and their child's behaviour, emotions and relationships. Responses were recorded as; not at all worried, a little worried, worried or very worried. A binary variable was created where responses of 'a little worried', 'worried' or 'very worried' were scored as '1' while parents who were 'not worried at all' were scored '0'.

2.5.2 Parenting Results

ADULT ADOLESCENT PARENTING INVENTORY 2- AAPI

While all five of the AAPI subscales were in the hypothesised direction, none were statistically significant. The step-down test showed that the joint effect of the five AAPI subscales was not statistically significant.

ADULT ADOLESCENT PARENTING INVENTORY 2- AAPI CUT-OFF SCORES

Two of the five AAPI cut-off subscales were in the hypothesised direction, yet neither were statistically significant. There was one significant result in the non-hypothesised direction. 24% of parents in the low treatment scored above the risk cut-off on the Parental lack of empathy subscale, while 34% of high treatment parents did so (p<.10, d=.22). The step-down test showed that the joint effect of the five AAPI cut-off subscales was not statistically significant.

NON STEP-DOWN MEASURES

Two of the eight measures in the non step-down category were in the hypothesised direction. There were no significant differences between the high and low treatment groups on any of these outcomes.

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
AAPI Standardised Scores										
	Realistic parental expectations of children	165	(82/83)	6.62	(1.91)	6.46	(1.89)	ns	ns	0.09
	Belief in the use of appropriate punishment	165	(82/83)	6.40	(1.14)	6.29	(1.47)	ns	ns	0.09
	Promoting children's power and independence	165	(82/83)	5.35	(2.23)	5.27	(2.13)	ns	ns	0.04
	Appropriate parent-child roles	165	(82/83)	6.11	(2.15)	6.07	(2.22)	ns	ns	0.02
	Parental empathy	165	(82/83)	4.93	(2.40)	4.92	(2.00)	ns	ns	0.01
AAPI Cut-offs										
*	Belief in the use of inappropriate punishment - At risk cut-off	165	(82/83)	0.00	(0.00)	0.02	(0.15)	ns	ns	0.22
*	Oppressing children's power and independence - At risk cut-off	165	(82/83)	0.23	(0.42)	0.28	(0.45)	ns	ns	0.10
*	Inappropriate parent-child roles - At risk cut-off	165	(82/83)	0.12	(0.33)	0.12	(0.33)	ns	ns	0.00
*	Unrealistic parental expectations of children - At risk cut-off	165	(82/83)	0.06	(0.24)	0.05	(0.22)	ns	ns	0.06
*	Parental lack of empathy - At risk cut-off	165	(82/83)	0.34	(0.48)	0.24	(0.43)	s~	ns	0.22
Non Step-down Measures										
	Total AAPI score	165	(82/83)	124.18	(14.25)	125.71	(12.68)	ns	-	0.11
*	AAPI - total number of scales at risk	165	(82/83)	0.76	(0.85)	0.76	(1.09)	ns	-	0.00
	KIDI %	165	(82/83)	70.19	(7.82)	69.72	(6.78)	ns	-	0.06
*	KIDI cut-off (lowest 10%)	162	(81/81)	0.10	(0.30)	0.10	(0.30)	ns	-	0.00
	Mother reads to child	165	(82/83)	0.90	(0.30)	0.90	(0.30)	ns	-	0.00
	Mother reads to child daily	149	(74/75)	0.46	(0.50)	0.53	(0.50)	ns	-	0.15
	Worried about child's behaviour	165	(82/83)	0.07	(0.26)	.07	(0.26)	ns	-	0.00
	Worried about child's language development	164	(81/83)	0.06	(0.24)	0.02	(0.15)	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. * 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 sample size reported are those used in the individual tests and may differ from the sample size used in the Step-down analyses which are based on the number of observations present in all variables included in the Step-down category. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

2.6 | Home Environment

The family environment is a child's primary environment during infancy and it plays a significant role in child outcomes (Carr, 2006). Aspects of the family environment include the sense of cohesion within the family, conflict and communication patterns as well as how organised the family is (Moos & Moos, 2009). Bullock and Pennington (1988) found that children's cognitive competence and social acceptance was related to greater family cohesion and expressiveness, as assessed by the Family Environment Scale (FES). Other studies found that children of families with low cohesion and organisation were more likely to have a difficult temperament (Lopez & Thurman, 1993; Matheny et al., 1987; Stadelmann et al., 2007). In assessing academic motivation and success in children, some studies found that children from more cohesive and expressive families, that emphasise intellectual-cultural development, study more, adjust better in school, have higher academic expectations, and have higher grade point averages (Feiring & Lewis, 1998; Kesetzis et al., 1998; King, 1998).

In relation to family-centred interventions, Armbruster and Fallon (1994) found that families lacking in cohesion and achievement orientation were associated with higher attrition from children's mental health clinic interventions. Furthermore, the Parent Effectiveness Training (PET) programme found that better family communication and a more supportive family environment was achieved as a result of the programme (Pinsker & Geoffroy, 1981; Wood & Davidson, 1987). There is little research on the effects of home visiting interventions on such family environment characteristics.

SOCIAL WORK INVOLVEMENT

Social Work involvement with a family indicates that the family may be 'at risk', whether due to child abuse or neglect, or indeed the presence of an individual in the household experiencing social or emotional difficulties (www.citizensinformation.ie, as accessed on 7th August, 2012). By reducing the involvement of social work, home visiting interventions can reduce the burden on crisis services as well as reduce the instance of abuse and neglect in the home. Few studies investigate the impact of home visiting intervention on social work involvement, and the few that do find mixed results. In the U.S., Mitchell-Herzfeld et al. (2005) reported a reduction in Child Protective Service Reports for abuse and neglect at twelve months as a result of a home visiting intervention whereas Duggan et al. (2007) did not find any difference.

2.6.1 | Home Environment Instruments

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, the 'real' family environment is measured, across two of the dimensions; the *Relationship* dimension includes cohesion, expressiveness and conflict, and the *System Maintenance* dimension includes organisation and control. This reduces the total number of items to 45. Mothers were asked to answer each of the 45 statements by choosing whether it was true or false for their family. The responses are matched to the score card (each item has a preferred direction, either true or false), and the scores are converted into a standard score using the FES conversion table.

SOCIAL WORK INVOLVEMENT

The participant was also asked if there was a social worker working with the family, and a binary was created for yes/no.

2.6.2 Home Environment Results

FAMILY ENVIRONMENT SCALE (FES) - STANDARDISED

Two of the five measures in the FES-standardised category were in the hypothesised direction, however neither of these effects were statistically significant. There was a statistically significant difference in the non-hypothesised direction on the cohesion standard score variable. The high treatment group reported a standardised score of 55.51, compared with the low treatment group average standardised score of 58.07 ($p<.05$, $d=.27$). This indicates that the high treatment group report less commitment and support within their families than the low treatment group. The step-down test showed that the joint effect of the five measures in the FES category was not statistically significant.

NON STEP-DOWN MEASURES

The measure in the non step-down category was in the non-hypothesised direction, and was not statistically significant.

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
FES Standardised									
Organisation Standard Score	165	(82/83)	48.15	(9.46)	46.88	(8.56)	ns	ns	0.14
Conflict Standard Score	165	(82/83)	70.43	(7.88)	69.63	(10.16)	ns	ns	0.09
Expressiveness Standard Score	165	(82/83)	51.76	(11.48)	52.08	(9.31)	ns	ns	0.03
Control Standard Score	165	(82/83)	52.26	(7.89)	52.65	(7.17)	ns	ns	0.05
Cohesion Standard Score	165	(82/83)	55.51	(10.19)	58.07	(8.62)	s~	ns	0.27
Non Step-down Measures									
Social Worker Working With Family	165	(82/83)	0.06	(0.24)	0.05	(0.22)	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. * 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 sample size reported are those used in the individual tests and may differ from the sample size used in the Step-down analyses which are based on the number of observations present in all variables included in the Step-down category. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

2.7 | Maternal Health & Wellbeing

The importance of maternal health in the postnatal period is often overlooked (Cheng et al., 2006), however maternal health difficulties often continue beyond the initial birthing period (Brown & Lumley, 1998). These difficulties can present challenges to child development. This section presents information relating to maternal health after pregnancy including physical health, mental health, and maternal substance use. The effects of each factor will be considered in light of the available research, with regards to the developmental outcomes for children and the effectiveness of interventions on these areas.

PHYSICAL HEALTH

Many women experience physical health difficulties related to pregnancy well into the postnatal period. One study found that rates of maternal fatigue remain as high as 50% after twelve months (Cheng et al., 2006). Another study found that 94% of women reported having one or more health problems in the postnatal period. The most commonly experienced problems in the postnatal period were tiredness, backache, sexual problems and haemorrhoids (Brown & Lumley, 1998). These difficulties can have knock on effects for child and maternal health difficulties and are often associated with negative outcomes for children. The American National Maternal and Child Health survey in 1988 and the longitudinal follow up in 1991 found that poor physical health was associated with poor maternal physical health of the child, tantrums and difficult peer interactions (Cheng et al., 2006).

Home visiting programmes that investigate maternal health tend to focus on mental health, stress and substance use. However at twelve months postpartum, some programmes investigated several other factors that may influence maternal physical health such as family planning, further pregnancies and contraceptive use. In general these studies do not find that home visiting interventions have a significant impact, although one study of a German intervention programme, ProKind, found that contraceptive use was higher and further pregnancies were lower in the intervention group than in the control group (Jungman et al., 2012).

MENTAL HEALTH

The uncertainty of pregnancy and motherhood can trigger mental health problems and post-partum depressive symptoms are common. The prevalence rates in Ireland were reported as 13% at six weeks postpartum and 10% at twelve weeks post-partum (Leahy-Warren, McCarthy & Corcoran, 2011) with rates reducing to 6% at four months postpartum (Matthey, Barnett, Judy & Waters, 2000). This is supported by a Finnish study, which identifies that symptoms of psychological distress continue beyond the immediate postnatal period (Vänskä et al., 2011). Postnatal depression is associated with a number of negative child outcomes including poor cognitive and emotional development (Beck, 1998), insecure attachment (Murray, 1991; Teti, Gelfand, Messinger, & Isabella, 1995), and behavioural problems (Murray, 1991). Mothers who suffer from postnatal depression may engage in less responsive parenting, which is commonly associated with negative developmental outcomes in children (Coolahan, 1997; Cunningham & Boyle, 2002; Snyder, Reid, & Patterson, 2003; Steinberg, Lamborn, Darling, Mounts & Dornbusch, 1994).

Another important aspect of mental health that has been shown to have both direct and indirect effects on child development is self-efficacy (Junttila, Vauras, & Laakkonen, 2007; Weaver, Shaw, Dishion, & Wilson, 2008). Self-efficacy refers to a person's belief in their ability to influence events which affect their lives (Bandura, 2010), while parental self-efficacy refers to a parents beliefs in his or her ability to influence the development of the child (Ardelt & Eccles, 2001). High self-efficacy is characterised by competence in the face of demands, less negative emotional arousal when stressed, and conceptualisation of difficult situations as challenges. While low self-efficacy is characterised by self-doubt, high levels of anxiety in the face of adversity, and avoiding difficult tasks (Jerusalem & Mittag, 1995). Parents with high parental self-efficacy are more likely to be engaged in positive parenting strategies which, in turn, increase the likelihood of their children's success in both academic and social-psychological domains (Ardelt & Eccles, 2001). Parenting self-efficacy can also have a direct impact on child outcomes through the modelling of attitudes and beliefs. Furthermore, Weaver and colleagues (2008) reported that the relationship between parenting self-efficacy and child behaviour problems is at least partially mediated by maternal depression which also has negative consequences for child development.

Parent's future outlook, or their ability to foresee short and long term consequences, may also affect child outcomes. Although this area of research is less developed, it suggests that parents' future orientation has an impact on their children's economic behaviour (Webley & Nyhus, 2006), as children learn through observing the behaviour of adults in their environment (Bandura, 1977). The ability to foresee short and long term consequences is also related to the ability to delay gratification and research suggests that children who are able to delay gratification at age four years have been later described as more successful in school and better able to cope with stress and frustration than those who were not able to delay gratification (Mischel, Shoda, & Rodriguez, 1989), illustrating the importance of such behaviours in young children.

Home visiting interventions commonly use outcome measures related to maternal mental health and general psychological wellbeing, however measures of future outlook are not commonly reported. Few evaluations of home visiting programmes at twelve months find positive outcomes in relation to maternal mental health. A German based study, ProKind, found that at twelve months participants in the intervention group reported lower levels of perceived stress and greater feelings of attachment and parental self-efficacy than those in the control group (Jungman et al., 2012). However, one study of the Healthy Families America programme found that at twelve months the programme seemed to have a negative impact on the general mental health of mothers (Duggan et al., 1999).

MATERNAL SUBSTANCE USE

While there is a multitude of research which reports the negative impact of maternal smoking, drinking and drug taking during pregnancy, there is a growing body of research highlighting the impact of exposure to these factors in the postnatal period (Pattenden et al., 2006). The lifestyle associated with drug addiction may also be damaging. For example, Das Eiden, Peterson and Coleman (1999) linked cocaine use in mothers with more violent environments.

Several evaluations of home visiting programmes have investigated cigarette, alcohol and drug use at twelve months, however these studies do not find any significant effects on these behaviours at this time point (Duggan et al., 1999; Koniak-Griffin et al., 2002; Mitchell-Herzfeld et al., 2005).

2.7.1 Maternal Health & Wellbeing Instruments

GENERAL HEALTH

The mother's current health status was assessed using a self-rated report of general health measured on a five point scale ranging from excellent to poor. This measure was dichotomised to create a binary indicator of ill health if the participant reported fair or poor health. The mother was considered to not have ill health if she indicated her current health was good, very good, or excellent. Participants were also asked how many times they visited the GP in the last six months (not including visits for the child). Participants were asked how many hours of sleep they got per night over the past six months. Sleep hours are recorded as less than 6 hours, more than 6 hours but no than 8 or more than 8 hours of sleep. A binary variable was created with less than six hours of sleep scored as inappropriate. Worried about getting a good night's sleep was recorded as '0' if the response was no, not at all, and as '1' if the response was sometimes or always.

Participants were asked if they were pregnant at the time of interview, and whether the pregnancy was planned. If they were not pregnant, they were asked whether they used birth control, and asked 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'. There were also some common responses to the 'other' choice; 'the coil', the 'patch', the 'bar', 'tubes tied', 'injection', 'husband sterilised', all of which were scored as valid. Some responses were scored as not valid, for example 'I have my partner withdraw'.

CURRENT SUBSTANCE USE

Three yes/no questions were used to assess whether participants smoked, drank alcohol or took drugs in the past six 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 variable 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 reported consuming more than 6 units of alcohol in one sitting, and also a regular binge drinking variable was constructed for those consuming more than 6 units in a sitting, more than twice a week. Whether the participant changed her smoking and/ or drinking habits was also calculated, based on the reported smoking/drinking habits at six months.

WEIGHT, DIET AND EXERCISE

Participants were asked how they viewed their eating habits, i.e. as very healthy, healthy, average, unhealthy or very unhealthy. A binary variable was created with healthy and very healthy being rated as 1, and average, unhealthy and very unhealthy rated as 0. Participants were then asked how often they ate different foods (never, rarely, sometimes, often or always). These included low fat foods, wholemeal breads, fish, fruit and vegetables and fatty foods. The responses were scored 1-5 with higher scores indicating healthier choices on the healthy eating scale. They were then asked how often they exercise, with exercise defined as 20 minutes of activity that increases heart rate. Responses for this question ranged from never/ once a month to 5-6 times a week/everyday. A regular exercise variable was coded as exercising more than three times a week. The mother was asked to report her current weight and mother's BMI was calculated using the mother's height as reported at baseline. A binary variable was calculated indicating whether the respondent was obese or not.

MENTAL HEALTH

WHO-5

Maternal wellbeing was assessed using the five item (α=.88) WHO-5 (World Health Organisation, 1998) instrument completed by the participant. The WHO-5 is a measure of emotional wellbeing. Participants were presented with five statements related to how they have been feeling over the past two weeks and asked to rate how often they have felt that way on a six point scale ranging from 0 meaning at no time to 5 meaning all of the time. A raw score was obtained by summing all of the responses, giving a range of 0 to 25 with lower scores, particularly those below 13, indicative of poor wellbeing. A binary variable was calculated for participants who obtained a score below 13. It is recommended that anybody who falls in this range is tested for depression.

MATERNAL SELF EFFICACY

Maternal self-efficacy was measured using 7 items of the mastery subscale from the Pearlin Self Efficacy Scale (Pearlin & Schooler, 1978) and 6 items of parental self-efficacy from the Abecedarian study (Borkowski, et al., 2001). Participants were presented with these 13 items related to how they feel about themselves, their life so far, and becoming a parent, and asked to rate how much they agree or disagree with each item on a scale ranging from 0 meaning strongly disagree to 4 signifying strongly agree. These measures provides scores on two subdomains including mastery (7 items, α=.76) or the degree to which the mother feels she has control over things that happen to her and parental self-efficacy (6 items, α=.72) or the mothers' belief that she is able to effectively parent her child/children, as well as an overall 'maternal efficacy score' (13 items, α=.81) score. The 'maternal efficacy score' was generated by summing the responses to each of the 13 individual items and dividing by 13 to get the average score. All scores represent the average response to all items within that scale or subscale and range from zero to four with higher scores indicating higher self-efficacy. The 'Pearlin cut-off score' was calculated using the mastery items from the Pearlin Self Efficacy Scale only. The cut-off represents the lowest 10th percentile of the whole sample (high treatment, low treatment and comparison group).

FUTURE OUTLOOK INVENTORY

The Future Outlook Inventory (FOI; Cauffman & Woolard, 1999) is an 8-item measure of an individual's consideration of future events and consequences. Participants were asked to verbally rate how true each item was about themselves on a Likert scale ranging from 1 never true to 4 always true. The scores of each item were summed and divided by 8 to give a mean total score. Higher scores indicate a higher degree of future consideration and planning.

2.7.2 | Maternal Health & Wellbeing Results

MATERNAL PHYSICAL HEALTH AND HEALTH BEHAVIOURS

Of the four measures included in the Maternal Physical Health category, none were in the hypothesised direction. The high and low treatment groups differed significantly in the non-hypothesised direction on one measure; health compared with other woman of the same age. 89% of the low treatment group reported good health in comparison to other women, compared with 78% of the high treatment group ($p<.05$, $d=.31$). The step-down test showed that the joint effect of the four measures in this category was not statistically significant.

CURRENT SUBSTANCE USE

Two of the four measures in the Current Substance Use category were in the hypothesised direction, with one significant difference between the high and low treatment groups. 83% of the high treatment group reported drinking alcohol in the past six months, compared with 94% of the low treatment group ($p<.05$, $d=.36$). In addition, the step-down test showed that the joint effect of the four measures in this category was statistically significant ($p<.10$). The joint effect was driven by the significant results found for alcohol use in the past six months.

EATING HABITS

Of the three measures in the Eating habits category, none were in the hypothesised direction. Further, none were statistically significant. The step-down test showed that the joint effect of the three measures in this category was not statistically significant.

MATERNAL SELF-EFFICACY

Neither of the self-efficacy sub-domains were in the hypothesised direction, with one significant difference between the high and low treatment groups in the non-hypothesised direction. The high treatment group reported an average score of 3.16 for the Pearlin Self-Efficacy Score, compared with the low treatment group average of 3.29 ($p<.10$, $d=.24$), indicating greater efficacy among the low treatment group. Overall, the step-down test showed that the joint effect of the two measures in this category was not statistically significant.

NON STEP-DOWN MEASURES

Ten of the fifteen measures in the non step-down categories were in the hypothesised direction. However, there were no significant differences between the high and low treatment groups on any of these outcomes.

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										
*	No. of GP visits in past 6 months	164	(81/83)	2.51	(3.43)	2.17	(4.23)	ns	ns	0.09
*	Less than 6 hours sleep per night	162	(81/81)	0.21	(0.41)	0.17	(0.38)	ns	ns	0.09
*	Worried about getting a good night's sleep	165	(82/83)	0.41	(0.50)	0.33	(0.47)	ns	ns	0.19
	Good health compared with other women same age	164	(81/83)	0.78	(0.42)	0.89	(0.31)	s~	ns	0.31
Current Substance Abuse										
*	Drank alcohol in past 6 months	164	(81/83)	0.83	(0.38)	0.94	(0.24)	$p<.05$	$p<.10$	0.36
*	Drug use in past 6 months	164	(81/83)	0.02	(0.16)	0.04	(0.19)	ns	ns	0.07
*	Currently a smoker	164	(81/83)	0.53	(0.50)	0.52	(0.50)	ns	ns	0.03
*	Consumed alcohol above recommended level	159	(80/79)	0.11	(0.32)	0.10	(0.30)	ns	ns	0.04
Eating Habits & Exercise										
	Regular exercise	164	(81/83)	0.41	(0.49)	0.42	(0.50)	ns	ns	0.03
	Healthy eating habits	164	(81/83)	0.38	(0.49)	0.40	(0.49)	ns	ns	0.03
	Healthy eating scale	163	(81/82)	19.14	(5.96)	19.52	(5.71)	ns	ns	0.07
Maternal Self-Efficacy										
	Pearlin Mastery Mean Score	165	(82/83)	2.87	(0.60)	2.88	(0.57)	ns	ns	0.02
	Parental Self-Efficacy Mean Score	165	(82/83)	3.16	(0.54)	3.29	(0.54)	s~	ns	0.24
Non Step-down Measures										
	WHO-5 Percentage Score	165	(82/83)	57.61	(24.03)	56.39	(24.84)	ns	-	0.05
*	Below WHO-5 Score of 13	165	(82/83)	0.35	(0.48)	0.42	(0.50)	ns	-	0.14
	Maternal Efficacy Mean Score	165	(82/83)	3.00	(0.51)	3.07	(0.49)	ns	-	0.13
*	Pearlin Mastery cut-off	165	(82/83)	0.12	(0.33)	0.14	(0.35)	ns	-	0.07
	Future Outlook Inventory	165	(82/83)	2.66	(0.56)	2.63	(0.57)	ns	-	0.06
	Changed smoking between 6 and 12 months (Number of cigarettes)	79	(39/40)	1.00	(6.29)	0.63	(4.68)	ns	-	0.07
*	Number of cigarettes per day	86	(43/43)	12.81	(7.45)	13.04	(5.58)	ns	-	0.04
	Changed drinking between 6 and 12 months (Number of Drinks)	127	(59/68)	0.22	(8.62)	0.20	(5.19)	ns	-	0.00
*	Binge drinking (> 6 units in any sitting)	140	(66/74)	0.52	(0.50)	0.49	(0.50)	ns	-	0.06
*	Binge drinking (> 6 units and drink >= 1-2 times per week)	91	(43/48)	0.26	(0.44)	0.21	(0.41)	ns	-	0.11
*	Mother's BMI	101	(51/50)	24.76	(5.94)	24.99	(4.45)	ns	-	0.04
*	Obese	101	(51/50)	0.16	(0.37)	0.14	(0.35)	ns	-	0.05
*	Currently pregnant	164	(81/83)	0.06	(0.24)	0.04	(0.19)	ns	-	0.12
	Currently using a valid form of birth control	155	(75/80)	0.77	(0.42)	0.74	(0.44)	ns	-	0.08
	New pregnancy planned	8	(5/3)	0.40	(0.55)	0.00	(0.00)	ns	-	1.03

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

2.8 | Maternal Social Support

Maternal social support is an important predictor of positive outcomes for both mothers and children during pregnancy and early infancy, especially for those families living in disadvantaged circumstances. Social support continues to play a protective role in the first year of a child's life, particularly in relation to maternal wellbeing, parenting and mother-child interaction. Life with a young child can be stressful and maternal stress in the first year has been linked to negative outcomes for both mothers and their children. Crnic, Greenberg and Slough (1986) observed that maternal stress was associated with less positive maternal attitudes and a less secure child-mother attachment. Social support has been shown to particularly influence the quality of parenting and attachment in the presence of stress (Crnic, et al., 1986). Indeed, Crockenberg (1981) noted that the positive effect of support was strongest for families under stress. A further study of maternal stress, involving an at-risk population living in poverty, reported that mothers with more social support showed greater increases in the frequency in positive parent-child activities (Green, Furrer & McAllister, 2007). The authors noted that social support influenced parenting by helping parents feel less anxious about their relationships and, thus, engage more with their children.

Social support comes in many different forms. For example, Levitt, Webber and Cherie (1986) found that maternal support is provided primarily by a spouse or partner, followed by the maternal grandmother. Other family members and close friends were also found to provide support, but to a lesser degree. These different sources of support vary in their impact on mother and child outcomes. The spouse or partner is cited most frequently as being the strongest predictor of maternal wellbeing (Crnic et al., 1983), and child functioning at twelve months (Crnic, Greenberg & Slough, 1986), particularly in first-time mothers (Levitt, Webber & Cherie 1986). In contrast, Levitt and colleagues (1986) observed that support from one's own mother was related to the wellbeing of second and third-time mothers, while community support was related to the quality of mother-child interaction. Of interest to the *PFL* programme, professional support, although found to be related to greater satisfaction with parenting and positive affect at 8 months, had no effect when the child was twelve months old. An experiment conducted by Jacobson and Frye (1991), in which first time mothers received social support for the first year, observed that those who had received support scored higher on an attachment rating than the control group. While a number of evaluations of home visiting interventions measured the impact on social support, none reported any positive effects at twelve months (Anisfeld et al., 2004; Barlow et al., 2007; Duggan et al., 1999; LeCroy & Crysik, 2011; Mitchell-Herzfeld et al., 2005).

To our knowledge, there are no previous studies investigating the impact of home visiting programmes on voting behaviour

2.8.1 | Maternal Social Support Instruments

SATISFACTION WITH FATHER'S INVOLVEMENT

Participants were asked questions relating to how involved the child's father is in her child's life and how satisfied she is with his level of involvement. A binary variable was used to assess whether or not the father was involved in the child's life. Participants were asked to rate their level of satisfaction with the support the child's father provides based on 14 questions (α=.91) assessing satisfaction in relation to helping with household chores, playing with the child, helping with transportation, helping with childcare, etc. Participants answered very dissatisfied, somewhat dissatisfied, neither/neutral, somewhat satisfied, very satisfied, or he does not help in this way. The responses to these questions were summed to create a scale representing mother's satisfaction with the father's involvement. This scale ranged from a minimum of 14 to a maximum of 90, with higher scores illustrating greater satisfaction.

SOCIAL SUPPORT

Participants were asked how often they meet with friends or relatives not living in their household and how often their child sees his/her grandparent. Response options to these questions were dichotomised into regularly (daily), or irregularly (less than once a day).

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

SATISFACTION WITH FATHER INVOLVEMENT

All three of the measures within the Satisfaction with Father Involvement category were in the hypothesised direction, however none were statistically significant. The step-down test showed that the joint effect of the three measures in this category was not statistically significant.

SOCIAL SUPPORT

Both measures in the social category were in the hypothesised direction, with one being statistically significant. 65% of the high treatment group reported meeting their friends regularly, compared with 54% of the low treatment group ($p<.10$, $d=.23$). The step-down test showed that the joint effect of the two measures in this category was not statistically significant.

NON STEP-DOWN MEASURES

Both measures in this category were statistically significant in the hypothesised direction. 65% of the high treatment group reported voting in the last general election, compared with 46% of the low treatment group ($p<.01$, $d=.39$). Similarly, 55% of the high treatment group reported voting in the last local/European elections, compared with 38% of the low treatment group ($p<.05$, $d=.35$).

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
Satisfaction with Father Involvement									
Father part of child's life	164	(81/83)	0.94	(0.24)	0.89	(0.31)	ns	ns	0.17
Satisfied with father's involvement	165	(82/83)	62.91	(7.60)	62.54	(8.97)	ns	ns	0.04
Father has daily contact with child	164	(81/83)	0.68	(0.47)	0.67	(0.47)	ns	ns	0.01
Social Support									
Meet friends frequently	164	(81/83)	0.65	(0.48)	0.54	(0.50)	$p<.10$	ns	0.23
Frequently sees grandparent	164	(81/83)	0.85	(0.36)	0.83	(0.38)	ns	ns	0.06
Non Step-down Measures									
Voted in last General Election	158	(78/80)	0.65	(0.48)	0.46	(0.50)	$p<.01$	-	0.39
Vote in last Local Elections and European Elections	157	(76/81)	0.55	(0.50)	0.38	(0.49)	$p<.05$	-	0.35

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

2.9 Childcare

There are many elements which mediate the effect of childcare on child outcomes. These include the socio-economic status and home environment of the child (Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007; Scaramella, Neppl, Ontai, & Conger, 2008), the quality and type of childcare, and the duration of childcare (both in how long the child has been attending the centre and the number of hours) (NICHD, 2004; Sylva, Stein, Leach, Barnes & Malmberg, 2011).

Formal centre-based childcare has been associated with increases in cognitive development (Sylva et al., 2011, NICHD, 2002), yet it has also been associated with higher incidences of externalising behaviours in children (Loeb et al., 2007; NICHD, 2004). These behavioural effects appear to be more pronounced for children entering childcare at less than one year of age (Loeb et al., 2007; NICHD, 2004). Furthermore, children receiving more centre-based care between the ages of zero to seventeen months displayed a reduction in academic skills indicating that the cognitive benefits may not apply to younger children (NICHD, 2004). A study using a UK sample found that centre-based childcare during the first year was particularly beneficial to children of low educated mothers, and had no effect on the later cognitive development of children from higher educated families (Cote, Doyle, Petitclerc & Timmins, forthcoming). The suggested optimal age to begin childcare is 2-3 years old, with children who begin younger displaying more negative behavioural effects (Barnett, 1995; Loeb et al., 2007). In contrast, it was found that children who attended centre-based childcare in Scandinavia scored highest in cognitive development at eighteen months (Berglund, 2005).

Further investigation indicates that higher quality childcare can enhance academic performance, regardless of the duration, whereas behaviour problems are related to the duration of care, regardless of the quality (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 internalised behaviours (Crockenberg, 2005).

There are few studies which report benefits of informal childcare at any age. Groeneveld (2010) report that children cared for in a home other than their own, received higher caregiver sensitivity and consistency, along with a less noisy environment, yet their cortisol levels (a by-product of stress) were the same when compared with children in centre-based care. Overall, centre-based childcare is reported to have greater cognitive benefits, yet more behaviour difficulties, whereas informal childcare, by its nature is less consistent.

As discussed, there are cognitive benefits and behavioural disadvantages associated with centre-based childcare during infancy. The quality of the centre, as well as the duration of attendance, play a role. There is limited research which assesses the impact of childcare in general for children aged twelve months or younger and it is unclear whether the benefits of centre-based care outweigh the disadvantages. In addition, there is limited research on the impact of home visiting programmes on childcare use for children aged twelve months or younger.

2.9.1 | Childcare 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. Those who indicated that they used childcare in the last six months were then asked to choose what type of childcare they mainly used from of 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 the child's grandparnet provide childcare to them. Additionally, participants were asked how many hours per week their child was in childcare, whether they paid for this childcare and how much, as well as what age their child was when he/she first started 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.

The analysis of the childcare measures is based on two-tailed tests as there is some debate regarding the optimal childcare situation for twelve month olds, therefore we cannot hypothesise the expected direction of the results.

2.9.2 | Childcare Results

CHILDCARE: TWO-SIDED INDIVIDUAL TESTS

Two measures in the Childcare category indicated a statistically significant difference between the high and low treatment groups. 30% of the high treatment group used 'any type of childcare' compared to 45% of the low treatment group ($p<.10$, $d=.30$). This indicates that there is a significant difference in childcare use across the two treatment groups. Also, 100% of the high treatment group who used childcare, reported that they were satisfied with the childcare their child received, compared with 84% of the low treatment group ($p<.05$, $d=.56$).

Table 2.7 - Results for High and Low Treatment Groups: Childcare and Service Use: Two-Sided

Variable	N	(n_{HIGH}/n_{LOW})	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	Two Sided Individual Test	Effect Size d
Childcare								
Any type of childcare	163	(80/83)	0.30	(0.46)	0.45	(0.50)	$p<.10$	0.30
Formal childcare	163	(80/83)	0.11	(0.32)	0.20	(0.41)	ns	0.25
Grandparent care	63	(25/38)	0.40	(0.50)	0.37	(0.49)	ns	0.07
Hours per week in childcare	26	(9/17)	18.11	(6.19)	18.53	(3.26)	ns	0.10
Age started childcare	61	(24/37)	6.88	(2.64)	6.68	(3.21)	ns	0.07
Satisfaction with childcare	61	(24/37)	1.00	(0.00)	0.84	(0.37)	$p<.05$	0.56
Childcare cost per hour	26	(9/17)	1.64	(0.74)	1.87	(1.55)	ns	0.18

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

2.10 | Household Factors & SES

The mediational role of household and socio-economic factors such as household composition, lone parent status, parental education, parental employment, parental relationship status and family finances on child development have been well established in the literature and are discussed extensively in the six month report. The impact of these factors on child outcomes, including academic attainment, begins in early infancy as they can determine the extent and quality of parent-child interaction (Bradley & Corwyn, 2002). For example, Harrison and Ungerer (2002) found that infants at twelve months were more likely to demonstrate secure attachment when their mothers expressed more commitment to work and less anxiety about using non-family child care, and returned to work earlier.

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 and whether or not the child's grandparent lives in the household. Additionally, the participant reported her current relationship status from a list of seven options. This information was used to generate two separate binary indicators denoting 1) whether the participant was currently in a relationship (married, cohabitating, or boyfriend) or 2) 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 six months old.

MATERNAL AND PARENTAL EMPLOYMENT

Several questions assessed the current work status of both the mother and the father. 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 FÁS training, or unpaid FÁS training. Responses to this question were dichotomised to represent the proportion of mothers and fathers in paid work versus not in paid work, and the proportion of mothers and fathers currently unemployed. Unemployed individuals were asked for how many months they have been without paid work. A binary variable denoting long term unemployment (greater than twelve months) was created. Participants also reported on whether they worked in full or part time employment and the approximate annual income of both parents.

FAMILY FINANCES

Participants' perception of financial difficulty was assessed by asking them to consider the total income of their household, and to rate on a seven point scale, ranging from with great difficulty to very easily, how difficult it was for the household to make ends meet. Responses to this variable were used to generate a binary variable indicating whether the participants make ends meet with difficulty or not. Participants were also asked to compare their financial situation to twelve months ago and a binary variable was created indicating whether their financial situation had got better or worse. Finally, they were asked 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 also if anyone in the household was currently in receipt of any social welfare payments. 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, it would be misleading to compare household income across participants without accounting for the number of people living in the household. To overcome this issue a variable representing the household equivalised weekly income was created. This was calculated by assigning a weight to each household member. 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.

DOMESTIC RISK

Participants were presented with a list of 8 potential domestic risks and asked to indicate if any of these factors had been an issue for anyone in the family. These included separation, parenting problems, domestic violence, abuse, suicidal thoughts, mental health issues, addiction and other. A total number of domestic risks score was also calculated by summing the number of risks each participant indicated.

2.10.2 Household Factors & SES Results

HOUSEHOLD FACTORS

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

MATERNAL EMPLOYMENT

One of the two measures in the Maternal Employment category was in the hypothesised direction this was significant. 15% of mothers in the high treatment group were unemployed long-term, where as 23% of the low treatment mothers were ($p<.10$, $d=.21$). The step-down test showed that the joint effect of the two measures in the Maternal Employment category was not statistically significant.

PATERNAL EMPLOYMENT

Neither of the measures in the Paternal Employment category was in the hypothesised direction. In addition, the step-down test showed that the joint effect of the two measures in the Paternal Employment category was not statistically significant.

FAMILY FINANCES

Two of the six measures in the Family Finances category were in the hypothesised direction, however, neither of these indicated a statistically significant difference between the high and low treatment groups. One measure was significantly different in the non-hypothesised direction. 73% of the low treatment group stated that their finances had improved over the past year, compared with 59% of the high treatment group ($p<.05$, $d=.29$). The step-down test showed that the joint effect of the six measures in the Family Finances category was not statistically significant

DOMESTIC ISSUES

Four of the eight measures in the domestic issues category were in the hypothesised direction, yet none of these were statistically significant. The step-down test showed that the joint effect of the eight measures in the domestic risk category was not statistically significant.

NON STEP-DOWN MEASURES

Four of the eleven measures which were not included in the above Step-down categories were in the hypothesised direction. However, none were significantly significant. One, father unemployed, was significant in the non-hypothesised direction. 40% of fathers in the low treatment group were unemployed compared to 50% of those in the high treatment group ($p<.10$, $d=.21$).

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										
	Married	164	(81/83)	0.16	(0.37)	0.13	(0.34)	ns	ns	0.08
	Household size	162	(81/81)	4.69	(1.82)	4.83	(1.93)	ns	ns	0.07
	Has a partner	164	(81/83)	0.70	(0.46)	0.77	(0.42)	ns	ns	0.15
Maternal Employment										
*	Mother long-term unemployed	162	(80/82)	0.15	(0.36)	0.23	(0.42)	$p<.10$	ns	0.21
	Mother in paid employment	163	(80/83)	0.33	(0.47)	0.40	(0.49)	ns	ns	0.15
Paternal Employment										
*	Father long-term unemployed	147	(71/76)	0.25	(0.44)	0.21	(0.41)	ns	ns	0.10
	Father in paid employment	150	(76/74)	0.45	(0.50)	0.53	(0.50)	ns	ns	0.16
Finances										
	Household financial situation over the next 12 months	153	(76/77)	0.79	(0.41)	0.75	(0.43)	ns	ns	0.09
*	Difficulty making ends meet	163	(81/82)	0.26	(0.44)	0.28	(0.45)	ns	ns	0.05
	Equivalised weekly household income	142	(70/72)	230.64	(104.56)	231.57	(96.20)	ns	ns	0.01
*	Receives social welfare payments	163	(81/82)	0.84	(0.37)	0.80	(0.40)	ns	ns	0.09
	Saves regularly	163	(81/82)	0.48	(0.50)	0.51	(0.50)	ns	ns	0.06
	Household current financial situation compared to 12 months ago	162	(81/81)	0.59	(0.49)	0.73	(0.44)	s~	ns	0.29
Domestic Issues										
*	Other risks	165	(82/83)	0.02	(0.16)	0.06	(0.24)	ns	ns	0.18
*	Suicidal thoughts	165	(82/83)	0.04	(0.19)	0.06	(0.24)	ns	ns	0.11
*	Separation	165	(82/83)	0.07	(0.26)	0.10	(0.30)	ns	ns	0.08
*	Parenting problems	165	(82/83)	0.05	(0.22)	0.06	(0.24)	ns	ns	0.05
*	Abuse	165	(82/83)	0.02	(0.16)	0.02	(0.15)	ns	ns	0.00
*	Domestic violence	165	(82/83)	0.04	(0.19)	0.02	(0.15)	ns	ns	0.07
*	Mental health issues	165	(82/83)	0.11	(0.31)	0.08	(0.28)	ns	ns	0.09
*	Addiction	165	(82/83)	0.06	(0.24)	0.01	(0.11)	ns	ns	0.26
Non Step-down Measures										
*	Total number of domestic risks	159	(79/80)	0.63	(1.17)	0.63	(1.14)	ns	-	0.01
*	Resides with grandparent	153	(77/76)	0.31	(0.47)	0.30	(0.46)	ns	-	0.02
*	Mother unemployed	164	(81/83)	0.27	(0.45)	0.34	(0.48)	ns	-	0.14
*	Father unemployed	154	(76/78)	0.50	(0.50)	0.40	(0.49)	s~	-	0.21
	Mother took or is currently on maternity leave	164	(81/83)	0.01	(0.11)	0.02	(0.15)	ns	-	0.09
	Same partner as 6 months ago	121	(57/64)	0.98	(0.13)	0.98	(0.13)	ns	-	0.02
	Partner is the child's father	121	(57/64)	0.98	(0.13)	0.94	(0.24)	ns	-	0.23
	Mother in part-time employment	58	(25/33)	0.64	(0.49)	0.61	(0.50)	ns	-	0.07
	Mother improvement in work status (between 6 and 12 months)	157	(79/78)	0.05	(0.22)	0.05	(0.22)	ns	-	0.01
	Mother's annual wage	22	(9/13)	24,061	(7,564)	23,986	(11,957)	ns	-	0.01
	Father's annual wage	51	(24/27)	25,494	(10,265)	24,797	(10,817)	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. * 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 sample size reported are those used in the individual tests and may differ from the sample size used in the Step-down analyses which are based on the number of observations present in all variables included in the Step-down category. The variables are reported in order of the largest to the smallest T statistic within each Step-down category.

2.11	**Main Results Summary**

INTRODUCTION

The twelve month results are consistent with the home visiting literature, such that there were limited significant findings between the high and low treatment groups. At this early stage of the programme, there has been an average of 21 home visits by the programme staff to the high treatment group families. While the majority of outcomes were in the hypothesised direction, the results are mixed with few statistically significant differences identified, as well as some significant differences in the non-hypothesised direction.

CHILD DEVELOPMENT

Children in the high treatment group and children in the low treatment group did not differ significantly across the majority of the child development domains, including developmental at-risk cut-off scores, cognitive development, atypical development, temperament and parental worry about their child's development. We did not hypothesise that we would find significant findings in the child development domain at twelve months, and none were found at six months.

Overall, three statistically significant differences were found between children in the high treatment and low treatment group in the Child Development domain, two in the hypothesised direction and one in the non-hypothesised direction.

- Children in the high treatment group displayed more developed fine motor skills than children in the low treatment group.
- Children in the high treatment group were less likely to be at risk of social and emotional difficulties than children in the low treatment group.
- Children in the low treatment group were reported to display more communicative gestures than children in the high treatment group.

CHILD HEALTH

Children in the high treatment group and children in the low treatment group did not differ significantly across the majority of the child health measures, including number of health problems, incidence of asthma, sleeping difficulties and breastfeeding. However, we did find significant results as hypothesised. These findings mirror the findings at six months in relation to food and sleeping location.

Overall, there were five statistically significant differences between the high and low treatment groups in the Child Health domain at twelve months, four of which were in the hypothesised direction.

- Children in the high treatment group were less likely to have had a chest infection compared to children in the low treatment group.
- Children in the high treatment group were more likely than children in the low treatment group to have received the necessary immunisations at twelve months.
- Children in the high treatment group were reported to eat more dairy and grains than children in the low treatment group.
- Children in the low treatment group were more likely have an appropriate sleep location than children in the high treatment group.

PARENTING

Although we hypothesised that we would find results in this domain, mothers in the high treatment group and mothers in the low treatment group did not differ significantly across the majority of the parenting measures, including parental attitudes toward child rearing, knowledge of child development and reading practices. There were differences found in the parenting domain at six months, specifically in the areas of parent-child interaction and parental attitude towards their child. However, the measures used at the two time points reflect different aspects of parenting. Given these findings, it may be that some aspects

of parenting may be more affected by the programme than others. There was one significant finding in the parenting domain in the non-hypothesised direction.

- Parents in the low treatment group were less likely to be at risk for lack of empathy than those in the high treatment group.

HOME ENVIRONMENT

As hypothesised, family home environment in the high and low treatment groups did not differ significantly on the majority of the environment domains, including household organisation and social work involvement. There were differences between the high and low treatment groups found at six months in home environment and safety. However, the home environment and safety measures used at six months focused on aspects of the physical home environment and appropriateness of child toys and activities, whereas measures at twelve months focused on the family relationship. Therefore, findings from the home environment domain at six and twelve months are not directly comparable. There was one significant difference between the high and low treatment groups at twelve months in the non-hypothesised direction.

- Families in the low treatment group were more likely to report a more cohesive family relationship than families in the high treatment group.

MATERNAL HEALTH AND WELLBEING

Counter to our hypothesis, mothers in the high treatment group and mothers in the low treatment group did not differ significantly across the majority of the maternal health characteristics including maternal weight, emotional wellbeing, contraception and eating habits. There was a significant programme effect for alcohol consumption. In comparison, at six months, there was no effect for alcohol consumption was found. The only effect at six months in the maternal health domain between the high and low treatment was related to hospitalisation post-birth.

At twelve months there were three significant differences between the high and low treatment group in the Maternal Health and Wellbeing domain, two were in the non-hypothesised direction.

- Mothers in the high treatment group were more likely to report that they had not drank alcohol in the previous six months, compared with mothers in the low treatment group.
- Mothers in the low treatment group rated their health as better when compared with other women their age, than did mothers in the high treatment group.
- Mothers in the low treatment group reported higher self-efficacy scores than mothers in the high treatment group.

MATERNAL SOCIAL SUPPORT

As expected, mothers in the high treatment group and mothers in the low treatment group did not report differences on some of the social support measures including father involvement and satisfaction with this involvement. This is in contrast with the findings at six months which indicated that mothers in the high treatment group were less satisfied with the father's involvement than mothers in the low treatment group. There were similar findings with regards to maternal socialising, that is the frequency of seeing friends and grandparents.

Overall, there were three significant differences in the Maternal Social Support domain at twelve months:

- Mothers in the high treatment group were more likely than mothers in the low treatment group to regularly meet with friends.
- Mothers in the high treatment group were more likely than mothers in the low treatment group to report voting in the last General Election and Local/European Elections.

CHILDCARE

Families in the high treatment group did not differ significantly from families in the low treatment group in terms of the number of hours per week the child was enrolled in childcare, the age at which the child started childcare, the use of grandparent care and the cost of childcare. At twelve months those in the low treatment were more likely to utilise childcare than those in the high treatment group. This is consistent with childcare findings at baseline which suggested that parents in the low treatment group intended to use more childcare. At six months there were no differences between the high and low treatment groups in relation to childcare, yet at twelve months there were two significant differences between the treatment groups in the childcare domain. Please note that the direction of these effects was not hypothesised in this analysis.

- Children in the low treatment group were more likely to attend childcare than those in the high treatment group.
- Mothers in the high treatment group were more likely to be satisfied with their childcare provider than mothers in the low treatment group.

HOUSEHOLD FACTORS AND SES

As hypothesised, there were very few differences between families in the high treatment group and families in the low treatment group with respect to household and SES factors. These include relationship status, parental employment, social welfare dependency, mental health and abuse risk factors, living arrangements and wages. There was only one difference between the high and low treatment groups at six months; more mothers in the high treatment group resided with the child's grandparent. At twelve months there were three significant differences between the groups in the non-hypothesised direction.

- More mothers in the high treatment group were less likely to be unemployed long term than those in the low treatment group.
- More mothers in the low treatment group reported that their financial situation had improved since the previous year, than in the high treatment group.
- More fathers in the high treatment group were more likely to be unemployed compared to those in the low treatment group.

SUMMARY

Overall 147 outcome measures were assessed at twelve months. One-tailed tests were estimated with 140 of the measures and two-tailed were estimated with 7 of the measures. Among the one-tailed tests, 77 were in the hypothesised direction such that the high treatment group had better outcomes than the low treatment group, and 11 (8%) of these differences were statistically significant. These differences were found in the domains of child development, child health, maternal health and wellbeing and social support. 48 of the measures were in the non-hypothesised direction, such that the low treatment group had better outcomes than the high treatment group, and 8 (6%) of these were statistically significant. These differences were found across all domains except social support. Of the 23 step-down categories, 2 were significant, BITSEA cut-offs in the child development domain and Current substance abuse in the maternal health and well-being domain. Among the two-tailed tests, there were 2 significant differences between the high and low treatment groups in the childcare domain.

Consistent with previous studies of home visiting programmes at 12 months, we hypothesised that

treatment effects would be found in the domains of child health, parenting and maternal health. The results suggest partial support for our hypotheses. In the child health domain, children in the high treatment group had fewer chest infections, ate more appropriate foods and were more likely to have the suggested immunisations at 12 months; but no significant programme effects were found for child hospitalisations. No significant programme effects were found in the parenting domain. In the mother's health domain, mothers in the high treatment group were less likely to drink alcohol than those in the low treatment group, however there were no other significant effects found. Moreover, mothers in the low treatment group were more likely than mothers in the high treatment group to rate their health more favourably than others. Counter to our hypotheses, significant treatment effects were found in the domains of child development and social support. This is note-worthy as previous studies of home visiting programmes do not report effects in these domains at 12 months.

2.12 Differential Interaction Results

INTRODUCTION

There is some evidence to suggest that certain groups of participants may benefit more from home visiting programmes than others (e.g. Heckman, Malofeeva, Pinto, & Savelyev, 2010). It is possible that the main results reported above, which include all participants, may mask treatment effects for particular kinds of individuals. In order to investigate differential effects of the *PFL* programme, interaction analyses were conducted based on child and family charactersitics that were found to be of relevance to early childhood interventions - child gender, primiparous status, marital status, maternal cognitive resources, maternal emotional wellbeing and domestic risk. All of the subgroup categorisations are based on baseline characteristics.

METHODS

An interaction analysis was conducted using a regression framework whereby the characteristic of interest (for example, gender) was interacted with the treatment indicator, while controlling for the main treatment effect and the main gender effect. The analysis shows a) whether the programme is having an overall impact on everyone in the sample, b) whether boys and girls have different oucomes regardless of what treatment group they are in, and c) whether the programme is having a different effect on girls or boys. While this analysis shows whether the programme has a different effect on girls and boys, it does not show the direction of this effect. Therefore, for all cases in which a significant interaction effect was found, a further sub-group analysis was conducted whereby the sample was divided into boys and girls to determine the direction of the effect

2.12.1 Testing for Differential Effects

Based on the literature, we identified a number of child and parent characteristics where differential programme effects may be found.

The programme may have different effects on:

- Girls (56% of the sample) and boys (44% of the sample)
- First-time parents (49% of the sample) and non first-time parents (51% of the sample).
- Non-partnered mothers (18% of the sample) and partnered mothers (82% of the sample)
- High cognitive resource mothers (48% of the sample) and low cognitive resource mothers (52% of the sample). As measured by the Wechsler Abbreviated Scale of Intelligence (WASI).
- High family risk (30% of the sample) and low family risk (70% of the sample). As measured by a composite score relating to financial difficulties, presence of biological father, social support, stressful life events, maternal education, maternal rigidity/flexibility, planned pregnancy, mental health and wellbeing, and experience of addiction.
- High emotional well-being (62%) and low emotional well-being (38%). As measured by World Health Organisation Wellbeing Index.

2.12.2 Differential Effects Results Summary

Table 2.9 summarises the results by indicating whether a significant interaction effect was identified in each domain (based on the interaction analysis) and the direction of that effect (based on the sub-group analysis). The results are presented by each domain – Child Development, Child Health, Parenting, Home Environment, Maternal Health & Wellbeing, Maternal Social Support, and Household Factors & SES.

The table reflects aggregate results for each subdomain category. Positive programme effects (i.e. those instances where the high treatment group had more positve effects than the low treatment group) are indicated by a + symbol. Negative programme effects or those instances where the low treatment group performed better are indicated by a – symbol. There were significant results in both the hypothesised and non-hypothesised directions.

In regards to gender, there were more effects, both positive and negative, found for boys than girls. Significant treatment effects for boys were found in the domains of child development, child health and maternal health. However, there were both negative effects found for boys and significant positive effects found for girls in the parenting domain. For primiparious status, there were more positive effects for first time mothers than for participants who were not first time mothers. There were positive treatment effects found for first time mothers in the domains of child development, parenting and maternal health. For relationship status, more positive programme effects were found for partnered parents, but more negative programme effects were found for non-partnered parents. Siginificant positive effects for first partnered mothers were found in the domains of child development and maternal health. Negative effects for non-partnered mothers were found in the domain of child development, however positive effects were found in the home environment domain. There were mixed effects found for high and low risk families with both having both positive and negative effects in the domains of child development and child health. Results were also mixed for families with mothers who had high and low cognitive resources. Although there were modest positive effects for mothers with low cognitive resources in the domain of child development and mothers with both high and low coginitive resources had positive effects in child health, both negative and positive effects were found in the maternal health domain for those with high cognitive resources. For maternal emotional well-being, there were positive treatment effects found for those with low emotional well-being in the domains of child development and parenting. No significant positive treatment effects were found for those with high emotional well-being.

The results of the interaction analysis are fairly mixed at twelve months and it is difficult to establish a coherent pattern at present. Based on the significant interaction effects, the results suggest that at 12 months the programme may have more effects for boys, partnered mothers, first time mothers and those with low emotional well-being. For instance, for the gender subgroup, boys had both more positive and more negative treatment effects. Moreover, although positive effects were found for boys in the domain of child development and child health, positive treatment effects were found for girls in the parenting domain. It is plausible that the programme may have different effects for different types of families and that these may change over time, such that families who benefit more at early stages of the programme may not necessarily be those who benefit most at later time points.

Table 2.9 - Differential interaction results

	Gender		Primiparous Status		Relationship Status		Risk		Cognitive Resources		Emotional Wellbeing	
	Boys	Girls	Primiparous	Multiparous	Partnered	Not Partnered	High	Low	Low	High	Low	High
Child Development												
ASQ Scores	+		+	-	+	-	-					
ASQ Cut-off Scores			+		+	+	-					
BITSEA							-	+				
BITSEA Cut-off scores					+	-					+	
CDI								-				
Non-Stepdown Measures			+	-	+					+	+	
Child Health												
Child Health	+			(+)				+				
Mother's Health Decisions for her Child	+									+		
Sleep Routines						-	-	+				
Appropriate Food							+		+			
Non Step Down Measures										+		
Parenting												
AAPI Standardised	-	+	+			-	(+)					
AAPI Cut Offs	-	+			(+)						+	
Non Step Down Measures						(-)						
Environment												
FES Standardised					+							
Non Step Down Measures					+					(-)		
Maternal Health												
Maternal Physical Health & Health Behaviours							-		-			
Current Substance Abuse	+			+	+							
Eating Habits	-					-					-	
Maternal Self-Efficacy												
Non Step Down Measures	+	-	+	-		-	(-)	-		+		(-)
Social Support												
Satisfaction with Father Involvement							+		(+)			
Social Support												
Non Step Down Measures												
Household Factors & SES												
Household Factors												
Maternal Employment												
Paternal Employment				-								
Finances			(+)					(+)	(-)			
Domestic Issues	-	+			+			+				
Non Step Down Measures		+		(-)	+	+		(+)		+		

Notes: This table reflects significant interaction effects by domain for each sub group of parents. + indicates a favourable treatment effect. - indicates an unfavourable treatment effect. Where an interaction did not refer to a statistically significant treatment effect (favourable or unfavourable) the 'strongest' p value was identified in each of the 2 subgroups (e.g. Low-v-High Boys and Low-v-High Girls). This means that the value closest to 0 or 100 was chosen, and is identified on the table by a (brackets), with (+) being a more favourable effect, and (-) being more negative.

2.13   Treatment & Comparison Group Summary

This is a summary of results comparing the twelve month outcomes of the two treatment groups to the twelve month outcomes of the comparison community group. For a detailed report of these findings, please see the following website: http://geary.ucd.ie/preparingforlife

2.13.1   Hypotheses

As the high treatment group and the low treatment group have received some level of treatment compared to the comparison group, we hypothesise that there may be statistically significant differences between the *PFL* treatment groups and the *LFP* community comparison group. Since the high treatment group supports were more intensive and substantial than the supports provided to the low treatment group, we hypothesise that there may be more differences between the high treatment group and the comparison group than the low treatment group and the comparison group.

Controlling for the baseline differences between the groups, a finding that the high treatment group have outperformed the comparison group at twelve months suggests that the high treatment supports were effective. Similarly, a finding that the low treatment group have outperformed the comparison group suggests that the low treatment supports were effective. If there are no statistical differences between the high/low treatment groups and the comparison group this may suggest that the treatment was not effective.

2.13.2   Key Findings: HighTreatment Group and Comparison Group

This section summarises the findings comparing the outcomes of the high treatment group to the outcomes of the comparison group. While 147 items were analysed, 7 were two tailed tests and had no hypothesised direction. Of the remaining 140 items, there were findings in the hypothesised direction for 57 of the 140 measures. There were positive significant differences between the high treatment group and the comparison group on 24 of these measures (17%), with most effects in the domains of child development, parenting and social support. Six of these effects remained significant in the multiple hypothesis analysis. These were the parenting, home environment, maternal health & wellbeing and social support domains. In addition, there were 74 differences in the non-hypothesised direction, and 9 of these were statistically significant (6%). These findings were in the areas of child health, maternal health & wellbeing and household factors.

2.13.3   Key Findings: Low Treatment Group and Comparison Group

This section summarises the findings comparing the outcomes of the low treatment group to the outcomes of the comparison group. While 147 individual outcomes were analysed, 7 of these were two tailed tests, i.e. there was no hypothesised direction. Of the remaining 140 measures, 49 were in the hypothesised direction. There were positive significant differences between the low treatment group and the comparison group on 17 measures (12%), with most effects in the domains of child development and parenting. 4 of the 23 step-down categories remained significant in the multiple hypothesis analysis, including child development, child health, parenting and household factors. There were 87 measures in the non-hypothesised direction, in that the comparison group outperformed the low treatment group. 13 of these were statistically significant (9%). These results were found in all domains except parenting and childcare.

2.13.4   Summary: Treatment & Comparison Group

For both comparison analyses, the treatment groups displayed some positive effects. As expected, there were more differences in the hypothesised direction between the high treatment and comparison groups (17%) than between the low treatment and comparison groups (12%). These differences were found across all domains, particularly child development and parenting. The high treatment group displayed positive effects in the social support domain, which were not found for the low treatment group when compared with the comparison group. In contrast, effects were seen for child health when the low treatment and comparison groups were compared, but not when the high treatment and comparison groups were

46

compared. There were 6% and 9% significant differences in the non-hypothesised direction in the high and low treatment analyses respectively. These differences were found across all domains between the low treatment and comparison groups, but were concentrated in household factors, child development, parenting and child health when the high treatment and comparison groups were compared. In the two-tailed childcare domain, there were two differences in each analysis.

Overall, the results of the high treatment group and comparison group analysis support the main findings, such that the additional supports provided to the high treatment group appeared to have some positive effects at twelve months. However, in some cases different effects were found for the low treatment group vs. the comparison group and the high treatment group vs. the comparison group than were found for low vs. high treatment groups. These results suggest that the low treatment is having a lesser impact on participant outcomes at twelve months, with a similar amount of hypothesised and non-hypothesised effects.

Chapter Three

Dynamic Analysis - Analysing Change Over Time

Introduction

A significant component of the *Preparing For Life* evaluation involves comparing the outcomes of participants over time. Such analysis is necessary in order to fully understand the dynamics of child and parent development and to identify the causal impact of the programme on these outcomes. A number of standardised instruments used to evaluate the *Preparing For Life* 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 treatment, low treatment and comparison groups. If the programme is effective, we would hypothesise that the rate of improvement in the high treatment group over time is greater than that of the low treatment and comparison groups. In addition, we would expect that the rate of improvement in the low treatment group over time is greater than that of the comparison group.

Table 3.1 lists the instruments collected at multiple time points between baseline and the twelve month data collection point. Two child development instruments (ASQ & ASQ-SE) were used at the six and twelve month data collection points, three parenting instruments were used at both baseline and the twelve month data collection point (AAPI, KIDI, Pearlin), and a measure of participant satisfaction was measured at six and twelve months (CSQ). One instrument measuring mental wellbeing (WHO-5) was used at all three time points. For the purposes of these analyses, the WHO-5 was analysed at both baseline to six months and six months to twelve months.

Table 3.1 - Instruments included in the dynamic analysis

	Measure	Baseline	6 Month Interview	12 Month Interview
Child outcomes	Ages & Stages Questionnaire (ASQ)		X	X
	Ages & Stages Questionnaire (ASQ) Social/Emotional		X	X
Parent outcomes	Adult-Adolescence Parenting Inventory 2 (AAPI)	X		X
	Knowledge of Infant Development (KIDI)	X		X
	Pearlin Self Efficacy Scale	X		X
	WHO-5 Mental Well-being	X	X	X
Participant satisfaction	Client Satisfaction Questionnaire (CSQ)		X	X

This chapter briefly describes some literature on this topic and outlines the statistical method which is used to compare changes in these outcomes across groups over time. It then presents the results of the analysis and concludes.

Dynamic Analysis Literature Review

Longitudinal studies which track changes in development over time generally focus on cognitive development and suggest that there are multiple pathways to academic competence in the early years (Burchinal et al., 2002). Results from these studies indicate that there is a general upward trend in developmental outcomes over time, with older children generally performing better than younger children (Hale, 1990). However this upward trend is filled with peaks and troughs from one time point to another (Wilson, 1983) and is also mediated by several factors such as parenting and childcare (Burchinal et al., 2002). Many home visiting interventions identify long-term benefits from home visitation (Howard & Brooks-Gunn, 2009), but relatively few have tracked changes in outcomes over time. The absence of such analyses suggests that many evaluations of early intervention programmes are not gaining a full understanding of the process of change brought about through home visitation (Lyons-Ruth & Easterbrooks, 2006).

The few evaluations of home visiting programmes which have analysed data over multiple time points include Armstrong et al., (1999), Olds et al., (2007), Culp et al., (2004), and Jungman et al., (2009). These studies use repeated measures analysis of variance (ANOVA) in order to assess group and time effects. However there is no consistent pattern of change in outcome variables across studies. In terms of child development, health and behavioural outcomes, a nine year follow up of the Nurse Family Partnership found that children in the treatment group consistently demonstrated better school performance than did their control group counterparts over time (Olds et al., 2007). An evaluation of a German based home visiting programme, ProKind, found that in terms of infant behaviour and neurodevelopment, children in the intervention group showed a significant improvement between six and twelve months that was not evident in the control group (Jungman et al., 2009). In relation to maternal outcomes, Olds et al. (2007) found that mothers in the intervention group reported a greater sense of mastery over difficulties in their lives. They also found that a sense of mastery was particularly evident for the two years that the programme was operating; however in the nine year follow up this difference was no longer evident (Olds et al., 2007). An evaluation of an Australian based home visiting intervention identified a significant reduction in post-natal depression scores for intervention group mothers between the immediate post-natal period and 6 week follow up, a decrease which was not identified in control group mothers (Armstrong et al., 1999). Moreover, a study by Culp et al. (2004) found that mothers in the group receiving home visits scored consistently better in terms of parenting knowledge, appropriate home environment and use of community services between six and twelve months.

The use of a repeated measures design allows for the identification of changes in outcomes as well as the stability of change over time for each group (Talpin, 2005). Results from these studies, using repeated measures ANOVAs, identify varied patterns of change for both intervention and control groups. For some measures significant differences between groups emerged due to a marked improvement in outcomes for one group, but for other measures significant differences were identified due to the deterioration in outcomes for one group (Culp et al., 2004; Armstrong et al., 2000; Armstrong et al., 1999). For example, a study by Culp et al. (2004) identified an increase in service use by intervention group parents over time that was not evident in the control group. However Lyons-Ruth and Easterbrooks (2006) found that a significant difference between groups on teacher rated behaviour was due to deterioration in behaviour scores for the control group rather than an improvement in the intervention group. Repeated measures analyses of changes over time by treatment group gives researchers the ability to identify these trends, which might otherwise go undetected in static time point analyses.

Although in developmental research the most commonly used method for analysing longitudinal data are based on ANOVA techniques, ANOVA methods may not be an ideal type of analysis for all types of longitudinal data. For many data sets there is a risk of misinterpreting results as the assumptions of ANOVA are often violated (Gueorguieva & Krystal, 2004; Hertzog & Rovine, 1985). One advantage of using ANOVA techniques is that they take into account variability within the sample, however in a randomised control trial this is not an issue (OECD, 2010). Methods such as Differences-in-Differences, the method adopted by the *Preparing For Life* evaluation, are most frequently used as a means of estimating the effects of certain policy interventions or programmes in a non-experimental context (Lechner, 2011). Differences-in-Differences is a less complex analysis to conduct than ANOVA (OECD, 2010) and several quasi-experimental investigations have utilised this method in the evaluation of early childhood and home visiting interventions (Edwards et al., 2009; Baqui et al., 2008; Liu et al., 2010; Armecin et al., 2006; Behrman et al., 2005; Havnes & Mogstad, 2010). In general Differences-in-Differences analysis is used in evaluations where samples are not randomly assigned into intervention and control groups. In these evaluations the use of this method controls for differences between groups for observed and unobserved characteristics (Armecin et al., 2006; Behrman et al., 2005). However, the use of Differences-in-Differences analysis can also be used to analyse the process of change in experimental interventions, as this type of analysis allows for both group-specific and time-specific effects (Edwards et al., 2009). By conducting this type of analysis in the *Preparing For Life* evaluation we can gain greater insight into the mechanisms of change.

3.3 Methodology

The methodology adopted to evaluate change over time in the *Preparing For Life* programme is the Differences-in-Differences method. This method, adapted from Ashenfelter and Card (1985), is necessary to account for any underlying trends within the data and reduce the probability of producing a biased estimate of programme impact. A simple comparison of outcomes in the treatment group pre- and post- intervention may potentially lead to a biased estimate of the impact of the programme due to an underlying trend over time (e.g. observations on child weight or height over time). Observing the before-and-after change in outcomes for the treatment group will not produce an estimate of the programme's causal impact as other external factors may also have influenced outcomes over time (e.g. natural development of a child). Similarly, comparing treatment and control groups at each time point may be problematic if there are pre-intervention differences across the two groups which are non-random. Even following randomisation into the high and low treatment groups, there is still the possibility that observed and unobserved differences may exist prior to the intervention occurring.

By combining these two methods and comparing the before-and-after changes in outcomes for the treatment group and the before-and-after changes in outcomes of the control group we can produce reliable estimates of the effect of the treatment over time. This method, "Differences-in-Differences" (Ashenfelter & Card, 1985), or Diff-in-Diffs for short, is a widely used quasi-experimental method that measures the effect of a treatment in a given time period. Diff-in-Diffs is considered a good non-experimental method for impact evaluation if the control and treatment groups are similar, and if the assumption that trends in the outcomes of the treatment and control groups, in absence of the treatment, would have been the same is true.

Simple two-group, two-period comparisons can be estimated by subtracting the pre-post difference in the treatment group from the pre-post difference in the comparison group. Consider the two-period, two-group example in Table 3.2. Within each group, before and after the treatment, an outcome is measured and the mean of each outcome for each group is calculated. The Diff-in-Diffs estimate of this example is calculated by subtracted the pre-post difference in the treatment group (D - B) from the pre-post difference in the control group (C – A).

Table 3.2 - Differences-in-Differences method for a two-period analysis

	Control	Treatment	Diff-in-Diffs
Period 1	Mean A	Mean B	
Period 2	Mean C	Mean D	
	(C - A)	(D - B)	DiD=((D-B)-(C-A))

The identifying assumption underlying Diff-in-Diffs is that trends in an outcome over time would have been the same in the absence of the treatment. The treatment prompts a deviation from this trend which can then be measured using the hypothetical counterfactual. Graphically this can be presented in Figure 3.1. The line BE is parallel to AC and controls for the pre-treatment differences between the treatment and control groups. The point E is a hypothetical point, representing where the treatment group would have been in the absence on the treatment (the counterfactual). The Diff-in-Diffs estimate of the treatment is therefore (D – E). It accounts for pre-treatment differences between the treatment and control groups (A – B).

Figure 3.1 - Graphical representation of the Differences-in-Differences method

In *Preparing For Life*, the Diff-in-Diffs method is used in a regression framework using OLS (Ordinary Least Squares) and permutation testing to evaluate the effectiveness of the intervention. Consider the following equation (EQ1).

$$\text{EQ1:} \quad Y = a + \beta_1(\text{Time2}) + \beta_2(\text{Treatment}) + \beta_3(\text{Time*Treatment}) + \epsilon$$

Here Y is the outcome obtained by each individual at each time point (i.e. the instrument score obtained pre- and post-intervention); Time2 is a binary variable identifying which observations are post intervention; Treatment is a binary variable denoting whether the observations relate to an individual in the control or treatment group; Time2*Treatment is an interaction of the previous two binary variables, β_3 is the coefficient of interest; finally, ϵ is the error term. Following each regression, permutation tests with 100,000 replications are conducted to determine the significance level associated with each treatment effect. Unlike the main analysis included in this report, the Diff-in-Diffs method does not use conditional or unconditional Step-down permutation testing. The method accounts for pre-existing differences (at baseline or six months) between the treatment and control groups and therefore does not require conditional testing. More complex analysis, involving multiple time points is also possible using econometric methods.

3.4 | Dynamic Analysis Results: High & Low Treatment Groups

Tables 3.3 to 3.5 and figures 3.2 to 3.5 present the results of the dynamic analysis which tests for significant differences between changes in the outcomes of the high treatment group and the low treatment group. The tables report the mean and standard deviation of the instruments at each time point for each group and the p-value indicating whether there was a significant difference in the outcomes of the groups over time. The figures provide a graphical representation of the average scores of the measures at each time point for each group. Only figures for the significant findings are presented.

3.4.1 Dynamic Child Development Results

Using ASQ Scores collected at the six and twelve month interviews, changes in child development can be assessed over time. The ASQ Scores were standardised to a mean of 100 and standard deviation of 15 to ensure comparability between the time points.
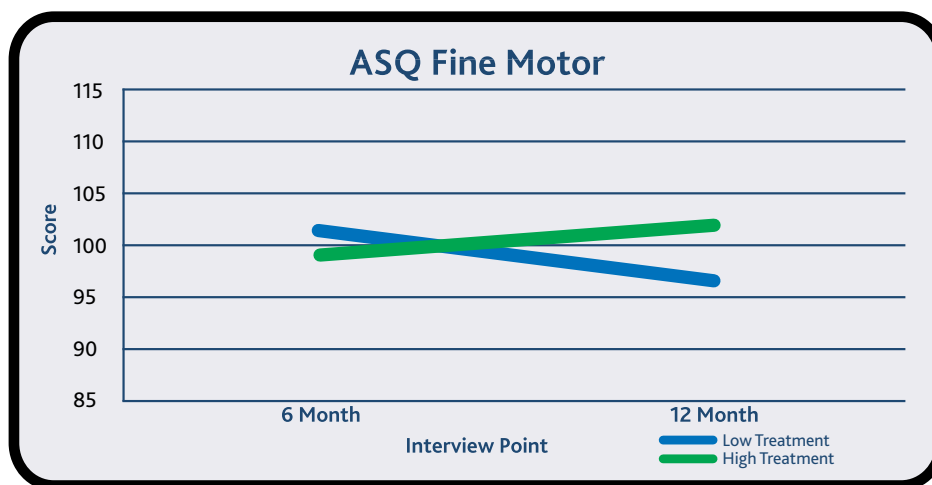
The comparison of the high and low treatment groups found that two of the six measures of child development were in the hypothesised direction (Table 3.3). However, the differences between the high and low treatment groups were only statistically significant for Fine Motor skills. Figure 3.2 presents the Differences-in-Differences result for ASQ Fine Motor Skills in graphical form. The fine motor skills of the high treatment group increased between six and twelve months, while the fine motor skills of the low treatment group decreased over time. Therefore, overall, the Differences-in-Differences estimate suggests a significant increase in fine motor skills for the high treatment group between six and twelve months.

Table 3.3 - Dynamic Child Development Results

	Variable	N	(n_{HIGH}/n_{LOW})	$M_{HIGH\ PRE}$	(SD_{HIGH})	$M_{LOW\ PRE}$	(SD_{LOW})	$M_{HIGH\ POST}$	(SD_{HIGH})	$M_{LOW\ POST}$	(SD_{LOW})	*Diff-in-Diffs* p^1
ASQ Scores												
	ASQ Communication Score	165	(82/83)	103.09	(18.78)	101.15	(20.45)	98.51	(20.80)	99.10	(20.42)	ns
	ASQ Gross Motor Score	165	(82/83)	102.31	(18.81)	99.25	(20.46)	99.92	(20.90)	98.38	(20.82)	ns
	ASQ Fine Motor Score	165	(82/83)	99.20	(19.43)	101.45	(20.41)	102.04	(18.99)	96.63	(22.63)	*p<.05*
	ASQ Problem Score	165	(82/83)	99.70	(19.85)	100.25	(22.01)	99.03	(19.60)	98.90	(21.87)	ns
	ASQ Personal Social Score	165	(82/83)	98.98	(19.56)	99.13	(21.12)	100.89	(18.48)	98.12	(21.93)	ns
*	ASQ Social-Emotional Score	165	(82/83)	98.27	(16.51)	98.66	(21.37)	102.13	(24.07)	99.41	(17.87)	ns

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. PRE refers to pre-intervention, POST refers to post-intervention. [1] one-tailed (right-sided) p value from a permutation test with 100,000 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. 's~'indicates that the variable was significant in a left-sided test.

Figure 3.2 - Graphical representation of Diff-in-Diffs results for ASQ Fine Motor

3.4.2 | Dynamic Parenting Results

The Adult-Adolescent Parenting Inventory (AAPI) and Knowledge of Infant Development Inventory (KIDI) questionnaires were measured at baseline and at twelve months.

The comparison of the high and low treatment groups found nine of the twelve measures of parenting were in the hypothesised direction (Table 3.4). However, only one of the differences between the high and low treatment groups was statistically significant. Figure 3.3 presents the result for being at risk of independence and power oppression (AAPI) graphically. The risk of oppression in the high treatment group decreased between baseline and twelve months, while the risk of oppression in the low treatment group increased over time. Therefore, overall, the Differences-in-Differences estimate suggests a significant decrease in the risk of oppression for the high treatment group between baseline and twelve months. Additionally, two Differences-in-Differences measures were statistically significant in the non-hypothesised direction. Figure 3.4 and 3.5 present the results for parental expectations and parental empathy. The risk of unrealistic parental expectations in the high treatment group increased between baseline and twelve months, while in the low treatment group it decreased over time. Parental lack of empathy in the high and low treatment groups decreased between baseline and twelve months. Overall, the Differences-in-Differences estimates suggest significant increases in the risk of unrealistic parental expectations and parental lack of empathy for the high treatment group between baseline and twelve months.

Table 3.4 - Dynamic Parenting Results

	Variable	N	(n_{HIGH}/ n_{LOW})	$M_{HIGH\,PRE}$	(SD_{HIGH})	$M_{LOW\,PRE}$	(SD_{LOW})	$M_{HIGH\,POST}$	(SD_{HIGH})	$M_{LOW\,POST}$	(SD_{LOW})	*Diff-in-Diffs p^1*
AAPI Standardised												
	Belief in the use of appropriate punishment	165	(82/83)	43.54	(5.62)	44.39	(6.09)	45.84	(4.48)	45.14	(5.63)	ns
	Promoting children's power and independence	165	(82/83)	19.45	(2.49)	19.91	(2.33)	20.07	(2.51)	19.90	(2.52)	ns
	Parental empathy	165	(82/83)	37.76	(5.33)	37.26	(5.73)	39.65	(6.21)	39.52	(5.50)	ns
	Realistic parental expectations of children	165	(82/83)	23.28	(3.99)	22.24	(4.26)	25.17	(5.21)	24.64	(4.90)	ns
	Appropriate parent-child roles	165	(82/83)	25.62	(4.41)	24.94	(4.92)	28.04	(5.05)	27.78	(4.90)	ns
AAPI Cut Offs												
*	Unrealistic parental expectations of children - At risk cut off	165	(82/83)	0.02	(0.16)	0.09	(0.28)	0.06	(0.24)	0.05	(0.22)	s~
*	Parental lack of empathy - At risk cut off	165	(82/83)	0.41	(0.50)	0.46	(0.50)	0.34	(0.48)	0.24	(0.43)	s~
*	Inappropriate parent-child roles - At risk cut off	165	(82/83)	0.18	(0.39)	0.23	(0.42)	0.12	(0.33)	0.12	(0.33)	ns
*	Belief in the use of inappropriate punishment - At risk cut off	165	(82/83)	0.02	(0.16)	0.06	(0.24)	0.00	(0.00)	0.02	(0.15)	ns
*	Oppressing children's power and independence - At risk cut off	165	(82/83)	0.28	(0.45)	0.17	(0.38)	0.23	(0.42)	0.28	(0.45)	*p*<.10
KIDI												
*	KIDI cutoff (lowest 10%)	165	(82/83)	0.10	(0.30)	0.16	(0.37)	0.10	(0.30)	0.10	(0.30)	ns
	KIDI %, higher score implies higher knowledge of child development	165	(82/83)	72.33	(7.04)	70.70	(8.30)	70.19	(7.82)	69.72	(6.78)	ns

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. PRE refers to pre-intervention, POST refers to post-intervention. [1] one-tailed (right-sided) p value from a permutation test with 100,000 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. 's~'indicates that the variable was significant in a left-sided test.

Figure 3.3 - Graphical representation of Diff-in-Diffs results for AAPI Oppression Risk Score
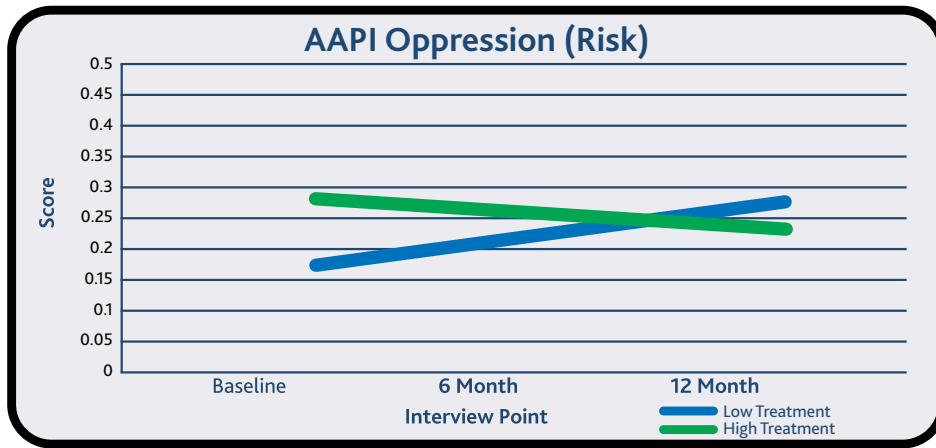

Figure 3.4 - Graphical representation of Diff-in-Diffs results for AAPI Unrealistic Expectations Risk Score
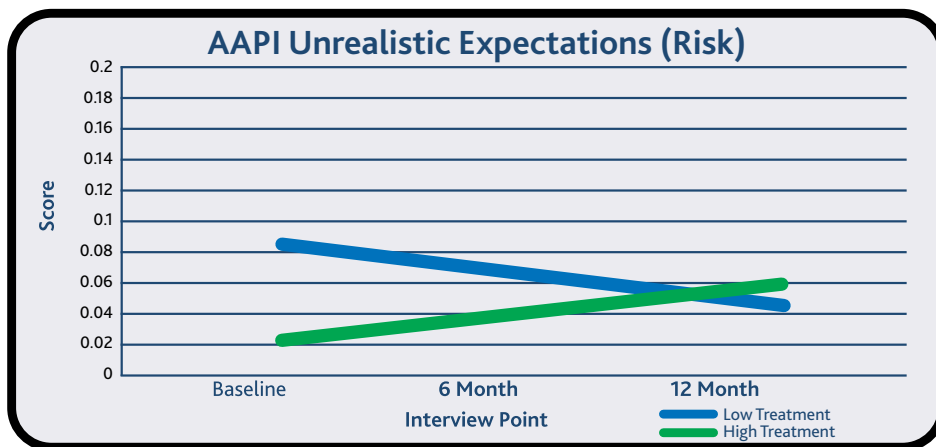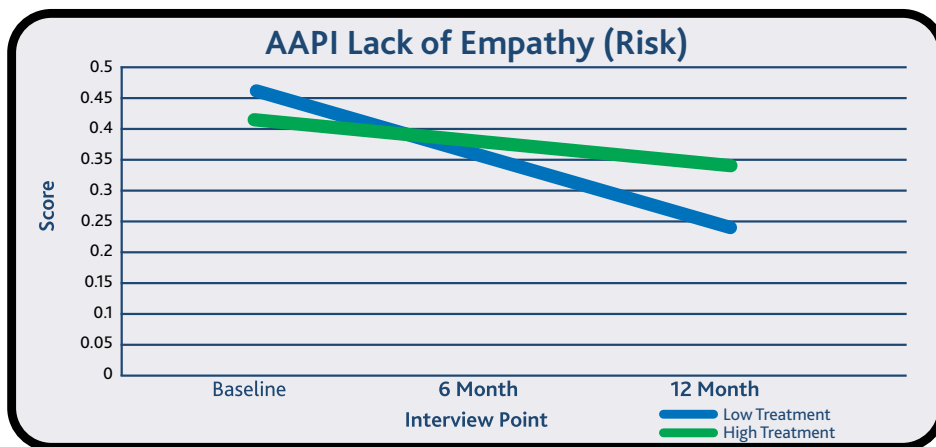

Figure 3.5 - Graphical representation of Diff-in-Diffs results for AAPI lack of Empathy Risk Score

3.4.3 | Dynamic Mental Health & Participant Satisfaction Results

The Pearlin measure of self-efficacy, measured at baseline and twelve month, and the WHO-5 measure of mental wellbeing, measured at baseline, six month and twelve month, are used to measure maternal mental health and wellbeing. Additionally, the Client Satisfaction Questionnaire was used to measure the participant's satisfaction with the programme at six and twelve months.

The comparison of the high and low treatment groups found three of the eight items measured in the hypothesised direction. However, none of the differences between the high and low treatment groups were statistically significant.

Table 3.5 - Dynamic Maternal Mental Health and Participant Satisfaction Results

Variable	N	(n_{HIGH}/n_{LOW})	$M_{HIGH\ PRE}$	(SD_{HIGH})	$M_{LOW\ PRE}$	(SD_{LOW})	$M_{HIGH\ POST}$	(SD_{HIGH})	$M_{LOW\ POST}$	(SD_{LOW})	$Diff\text{-}in\text{-}Diffs$ p^1
Pearlin											
Parental Self-Efficacy Mean Score	165	(82/83)	3.04	(0.54)	3.18	(0.53)	3.16	(0.54)	3.29	(0.54)	ns
Pearlin Mastery Mean Score	165	(82/83)	2.79	(0.60)	2.91	(0.62)	2.87	(0.60)	2.88	(0.57)	ns
Pearlin Mean Score	165	(82/83)	2.91	(0.49)	3.05	(0.52)	3.00	(0.51)	3.07	(0.49)	ns
WHO-5											
WHO-5 Percentage Score (0-6 months)	165	(82/83)	54.94	(19.50)	57.6	(22.94)	63.69	(21.37)	64.98	(20.66)	ns
WHO-5 Percentage Score (6-12 months)	165	(82/83)	63.69	(21.37)	64.98	(20.66)	57.42	(24.30)	56.34	(25.05)	ns
WHO-5 Below (0-6 months)	165	(82/83)	0.38	(0.49)	0.36	(0.48)	0.26	(0.44)	0.24	(0.43)	ns
WHO-5 Below (6-12 months)	165	(82/83)	0.26	(0.44)	0.24	(0.43)	0.35	(0.48)	0.42	(0.50)	ns
Client Satisfaction Questionnaire											
Total Satisfaction	165	(82/83)	70.80	(5.94)	64.20	(9.91)	69.07	(6.65)	61.81	(10.56)	ns

Notes: *'N'* indicates the sample size. *'M'* indicates the mean. *'SD'* indicates the standard deviation. PRE refers to pre-intervention, POST refers to post-intervention. [1] one-tailed (right-sided) p value from a permutation test with 100,000 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. *'s~'* indicates that the variable was significant in a left-sided test.

3.5 | Dynamic Key Findings

Overall, the dynamic analysis using the Differences-in-Differencess method echo the findings of the main (high vs. low treatment group) results. Fourteen of the twenty-six items were in the hypothesised direction, however only two of these were statistically significant (8%). Furthermore, two results were also statistically significant in the non-hypothesised direction (8%). Analysis of the dynamic results is on-going, further analyses which explore the within group differences in each group over time are planned for future papers and reports.

- Children in the high and low treatment groups did not differ significantly across many child developmental domains over time. However, children in the high treatment group have significantly better fine motor skill development between six and twelve months than the low treatment group.
- Children in the high treatment group were significantly less likely to experience parental oppression of their power and independence by twelve months than those in the low treatment group.
- Mothers in the high and low treatment groups did not differ significantly from each other over time across various mental health domains.

3.6 | **Dynamic Findings Summary: Treatment Groups & Comparison Group**

This is a summary of dynamic results comparing the two treatment groups to the comparison community group. For a detailed report of these findings, please see the following website:
http://geary.ucd.ie/preparingforlife

For the high vs. comparison group, sixteen of the twenty-five items measured were in the hypothesised direction, and five (20%) were statistically significant. Two were significant in the non-hypothesised direction (8%). For the low treatment vs. comparison group, twelve of the twenty-five items measured were in the hypothesised direction, and two (8%) were statistically significant. Three were significant in the non-hypothesised direction (12%). The results are detailed below.

HIGH TREATMENT GROUP AND COMPARISON GROUP

- Children in the high treatment group and comparison group did not differ significantly across many child developmental domains over time. However, children in the high treatment group have significantly lower communication development between six months and twelve months than the comparison group.
- Children in the high treatment group were significantly more likely to have parents who believe in appropriate punishments and who promote their child's power and independence by twelve months than those in the comparison group.
- Mothers in the high treatment group and comparison group differed significantly from each other across various mental health domains. Mothers in the high treatment group had significantly higher efficacy scores by twelve months and better mental health scores by six months than the comparison group.

LOW TREATMENT GROUP AND COMPARISON GROUP

- Children in the low treatment group and comparison group did not differ significantly across many child developmental domains over time. However, children in the low treatment group had significantly lower communication development and fine motor development between six months and twelve months than the comparison group.
- Children in the low treatment group were at a significantly lower risk of a parental lack of empathy by twelve months than those in the comparison group.
- Mothers in the low treatment group were at a significantly lower risk of mental health problems than mothers in the comparison group by six months. No statistical differences were found between six months and twelve months.

Chapter Four

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 and engagement from 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'. This chapter describes and analyses *PFL* implementation practices regarding participant attrition, engagement, satisfaction, and potential contamination between programme intake and when the *PFL* child was twelve months of age.

4.1 | *PFL* Attrition up to Twelve 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. This section investigates the level and determinants of attrition in the *PFL* sample between baseline and the twelve month survey.
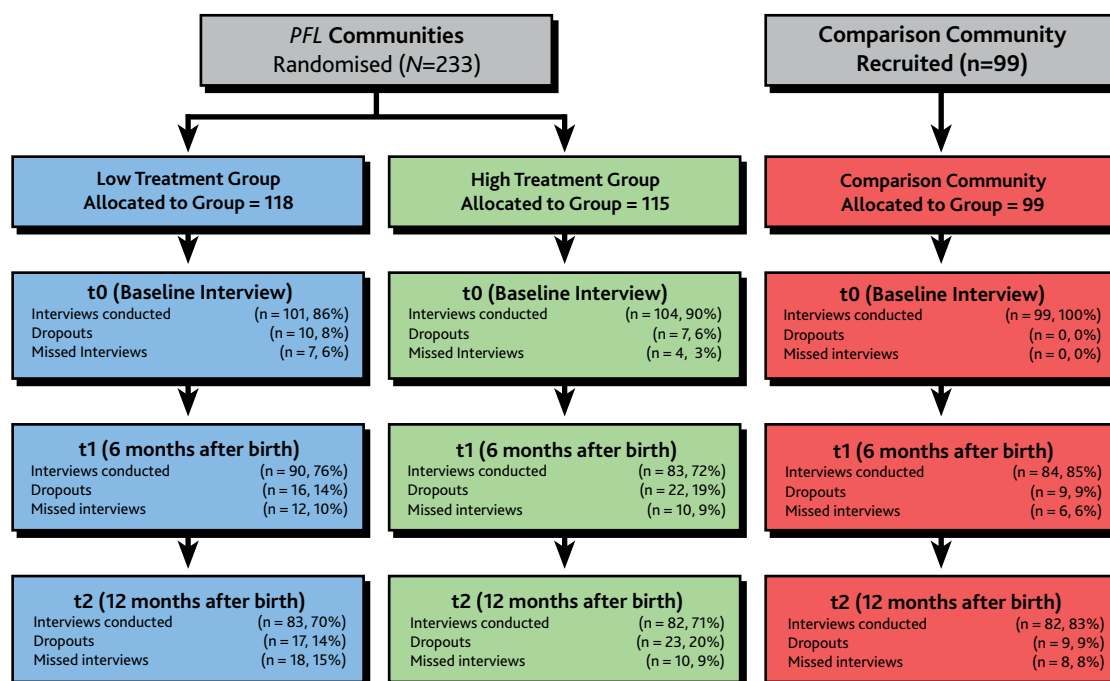
4.1.1 | Attrition/Disengagement in *PFL*

The Consort Diagram (Figure 4.1) describes the progression of the participants between programme entry and twelve months. In total, 247 twelve month interviews (n*High* = 82; n*Low* = 83; n*LFP* = 82) were completed. These 247 participants represent 70% of the original sample recruited into the study (n*High* = 115; n*Low* = 118; n*LFP* = 99). The twelve-month completion rate was similar for both the high (70%) and low treatment groups (71%). The comparison group had a slightly higher completion rate (83%).

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, 15% of the sample were classified as official 'dropouts' between baseline and twelve months, with the highest dropout rate experienced among the high treatment group at 20%, while the low treatment group experienced a dropout rate of 14%. Nine percent of the comparison group dropped out of the evaluation after completing the baseline interview, but prior to completing a twelve month interview. The dropout rate between six and twelve months was minimal however (n*High* = 1; n*Low* = 1; n*LFP* = 0). The most frequently cited reason given by those who left the programme prior to twelve months was time constraints. Many former participants felt that their busy schedules restricted them from participating. A number of former participants also suggested that their children did not need the programme, that the interview questions were too personal and that a five year commitment was too long.

In addition to those who dropped out, 10% of the sample did not complete a twelve 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 (those who missed interviews) are those who did not respond to repeated attempts by the evaluation team to be contacted or declined to be interviewed. The rates across the high and low treatment groups were 9% and 15% respectively, the rate for the comparison group was 8%. It is possible that some of these participants will re-engage with the programme at later data collection waves. Overall, the level of attrition 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 equal across the two samples (29%) from randomisation to the twelve month interview, with the majority of attrition/disengagement occurring prior to the six month interview.

Figure 4.1 - Twelve Month Consort Diagram

Note: Percentages from t1 forward are based on active participants only. Dropout participants include both voluntary and involuntary dropouts.

4.1.2 | Analysis of Attrition/Disengagement before Twelve Months

It is important to examine whether attrition and disengagement/missed interviews has led to systematic differences between those who remained in the programme. The analysis below compares the baseline characteristics of participants who completed a twelve month interview to those who did not complete a twelve month interview. Thus, the analysis of 'attritors' includes those who have officially dropped out of the programme between baseline and twelve months and those who missed the twelve 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 and low treatment groups are analysed separately.

Table 4.1 reports the baseline characteristics of the high treatment group by attrition status and tests for significant differences between the attrition group (29%) and the non-attrition group (71%) based on each characteristic. It shows that of the 21 maternal characteristics examined, statistically significant differences were found for three measures. Specifically, high treatment group mothers who did not complete a twelve month interview were less likely to be employed at baseline ($p<.05$), had lower levels of cognitive resources ($p<.10$), and were less likely to have drank alcohol during pregnancy ($p<.05$).

Table 4.1 - 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	22	20.63	7.83	82	21.84	7.88	ns
Mother's age	22	23.95	5.06	82	25.87	6.01	ns
Partnered	22	0.64	0.49	82	0.82	0.39	ns
Married	22	0.09	0.29	82	0.16	0.37	ns
Living with parent(s)	22	0.68	0.48	82	0.54	0.50	ns
First time mother	22	0.64	0.49	82	0.51	0.50	ns
Low education	22	0.45	0.51	82	0.30	0.46	ns
Mother employed	22	0.14	0.35	82	0.43	0.50	*p*<.05
Saves regularly	22	0.36	0.49	82	0.50	0.50	ns
Social housing	21	0.62	0.50	82	0.54	0.50	ns
Cognitive Resources (WASI) †	10	76.10	10.65	81	83.11	12.60	*p*<.10
Vulnerable attachment (VASQ)	22	19.14	3.26	82	18.00	3.87	ns
Self efficacy (Pearlin)	22	2.69	0.76	82	2.79	0.56	ns
Self esteem (Rosenberg)	22	12.23	2.89	82	12.98	2.63	ns
Knowledge of infant development (KIDI)	22	71.95	9.56	82	72.33	7.04	ns
Positive parenting attitudes (AAPI)	22	5.24	1.86	82	5.25	1.23	ns
Physical Health Condition	22	0.73	0.46	82	0.76	0.43	ns
Mental Health Condition	22	0.27	0.46	82	0.28	0.45	ns
Smoking during pregnancy	22	0.50	0.51	82	0.51	0.50	ns
Drinking during pregnancy	22	0.09	0.29	82	0.29	0.46	*p*<.05
Drugs ever used	22	0.18	0.39	82	0.12	0.33	ns

Note: *N*=sample size, *M*=mean, S*D*=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. † Measured at 3 months postpartum.

Table 4.2 reports the baseline characteristics of the low treatment group by attrition status and tests for significant differences between the attrition group (30%) and the non-attrition group (70%) based on each characteristic. It shows that of the 21 maternal characteristics examined, statistically significant differences were found on four measures. Specifically, low treatment group mothers who did not complete their twelve month interview had lower levels of education ($p<.10$), lower levels of cognitive resources ($p<.05$), less knowledge of child development ($p<.05$), and less favourable approaches to parenting ($p<.01$).

Table 4.2 - 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	19	22.05	6.74	82	21.17	7.03	ns
Mother's age	19	26.00	5.95	82	25.13	6.02	ns
Partnered	19	0.89	0.32	82	0.83	0.38	ns
Married	19	0.26	0.45	82	0.16	0.37	ns
Living with parent(s)	19	0.42	0.51	82	048	0.50	ns
First time mother	19	0.53	0.51	82	0.49	0.50	ns
Low education	19	0.79	0.42	82	0.30	0.46	$p<.01$
Mother employed	19	0.26	0.45	82	0.43	0.49	ns
Saves regularly	19	0.37	0.50	82	0.55	0.50	ns
Social housing	19	0.58	0.51	82	0.55	0.50	ns
Cognitive resources (WASI) †	9	69.22	10.26	83	81.54	12.75	$p<.05$
Vulnerable attachment (VASQ)	19	19.05	4.35	82	17.54	3.86	ns
Self efficacy (Pearlin)	19	2.72	0.11	82	2.91	0.07	ns
Self esteem (Rosenberg)	19	12.79	3.01	82	12.78	2.84	ns
Knowledge of child development (KIDI)	19	66.02	6.63	82	70.70	8.30	$p<.05$
Positive parenting attitudes (AAPI)	19	4.20	1.42	82	5.33	1.34	$p<.01$
Physical Health Condition	19	0.53	0.51	82	0.65	0.48	ns
Mental Health Condition	19	0.16	0.37	82	0.26	0.44	ns
Smoking during pregnancy	19	0.53	0.51	82	0.46	0.50	ns
Drinking during pregnancy	19	0.32	0.48	82	0.26	0.44	ns
Drugs ever used	19	0.21	0.42	82	0.13	0.34	ns

Note: *N*=sample size, *M*=mean, *SD*=standard deviation. *p*-values were obtained from two-sided t tests based on permutation testing with 1000 replications. '$p<.01$', '$p<.05$' and '$p<.10$' indicate that the test is statistically significant at the 1%, 5%, and 10% level respectively. † Measured at 3 months postpartum.

Table 4.3 reports the baseline characteristics of the comparison group by attrition status and tests for significant differences between the attrition/disengaged group (17%) and the non-attrition/engaged group (83%) based on each characteristic. It shows that of the 21 maternal characteristics examined, four statistically significant differences were found. Specifically, comparison group mothers who did not complete their twelve month interview had higher levels of education ($p<.01$), lower levels of self-efficacy ($p<.10$), less knowledge of child development ($p<.05$), and less favourable approaches to parenting ($p<.01$).

Table 4.3 - 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	17	28.18	6.82	82	24.54	6.08	ns
Mother's age	17	27.59	5.90	82	27.22	6.30	ns
Partnered	17	0.82	0.39	82	0.88	0.33	ns
Married	17	0.29	0.47	82	0.16	0.37	ns
Living with parent(s)	17	0.35	0.49	82	0.32	0.47	ns
First time mother	17	0.41	0.51	82	0.41	0.50	ns
Low education	17	0.24	0.44	82	0.26	0.44	*p<.01*
Mother employed	15	0.47	0.52	82	0.43	0.50	ns
Saves regularly	15	0.53	0.52	81	0.57	0.50	ns
Social housing	17	0.53	0.51	81	0.41	0.49	ns
Cognitive resources (WASI) †	6	84.67	18.38	82	88.40	13.59	ns
Vulnerable attachment (VASQ)	17	16.88	4.18	82	16.91	3.49	ns
Self-efficacy (Pearlin)	17	2.53	0.86	82	2.91	0.65	*p<.10*
Self-esteem (Rosenberg)	17	12.29	3.04	82	13.20	3.10	ns
Knowledge of child development (KIDI)	17	72.35	8.22	82	73.03	8.84	*p<.05*
Positive parenting attitudes (AAPI)	17	5.02	1.68	81	5.86	1.30	*p<.01*
Physical Health Condition	17	0.71	0.47	82	0.66	0.48	ns
Mental Health Condition	16	0.31	0.48	82	0.38	0.49	ns
Smoking during pregnancy	17	0.24	0.44	82	0.37	0.48	ns
Drinking during pregnancy	17	0.24	0.44	82	0.32	0.47	ns
Drugs ever used	17	0.06	0.24	82	0.17	0.38	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. † Measured at 3 months postpartum.

| 4.1.3 | Key Findings |

The level of official attrition from *PFL* between baseline and twelve months was quite low at 15% across the whole sample. Importantly, the level of official attrition was minimal between the six and twelve month interview rounds, at 1% across the high and low treatment group and zero in the comparison group. Overall, official attrition between programme intake and twelve months is higher among the high treatment group (20%) than among the low treatment group (14%) who were less intensively engaged in the *PFL* programme. As the high treatment group were more regularly in contact with the *PFL* programme staff they had more opportunities to officially inform the staff of their desire to dropout from the programme. Indeed, a greater proportion of the low treatment group (15%) was classified as disengaged or missed their twelve month interview when compared with the high treatment group (9%). Thus it is possible that many of the participants who were missed during the twelve month survey represent participants who are less engaged with the programme and more inclined to dropout.

For the purposes of analysis we compared the baseline characteristics of those who participated in the twelve month survey to those who did not. Overall, there is weak evidence that there are systematic differences between these groups. However, there is no definite pattern as to whether the more advantaged or disadvantaged participants were difficult to contact. For example, in the low treatment group, those who did not participate in this survey had lower levels of education, lower levels of cognitive resources, and less favourable approaches to parenting as measured by the AAPI scale. Mothers in the high treatment group who did not participate in the survey were less likely to be employed and had lower levels of cognitive resources, but they were less likely to have consumed alcohol during pregnancy. For the comparison group, mothers who did not complete a twelve month survey had lower levels of self-efficacy, lower scores on the KIDI scale, and less favourable approaches to parenting. On the other hand, they had higher levels of education. It is possible that these differences are simply the result of random differences in the data. It will be important to examine the attrition rate at each subsequent survey wave in order to understand whether truly systematic patterns emerge.

| 4.2 | **Participant Engagement up to Twelve 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 and determinants of engagement in the *PFL* sample between baseline and the twelve month survey.

| 4.2.1 | Instruments |

Information on participant engagement within *PFL* was gathered from two sources – the *PFL* database and paper files maintained by the *PFL* mentors and survey responses from participants at the twelve 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 twelve months old, b) the percentage of prescribed home visits delivered between intake and twelve 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 twelve months. For each of these measures, we examined programme engagement from programme entry until birth, between birth and six months, between six and twelve months, and for the whole period. 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, even though they may have subsequently dropped out during 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.

4.2.2 Participant Engagement from Mentor Records

Table 4.4 provides a summary of participant engagement in the *PFL* programme between programme entry and twelve months of age for the high treatment group. The analysis is disaggregated into the prenatal period, birth to six months, six to twelve months, and total engagement up to twelve 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 twelve months. The *PFL* manual initially set guidelines of weekly home visits during the pre and postnatal period, however, the implementation team moved to fortnightly visits soon after the programme began as weekly visits were not feasible to the majority of the *PFL* participants. Thus, the figures below are estimated based on prescribed weekly and 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 twelve months was 72 based on a prescribed weekly visit and 36 based on a prescribed fortnightly home visit. Table 4.4 shows that on average, participants in the high treatment group received 20.7 home visits between programme entry and twelve months. The minimum number of visits received was 1 and the maximum was 42. The average number of home visits in the prenatal period was 6.2 and the average number of visits between birth and six months and between six and twelve months were quite similar at 7.6, and 6.9 visits respectively. Figure 4.2 displays the variation in the number of home visits over the entire period.

These figures were then used to calculate the proportion of prescribed home visits actually delivered. Table 4.4 shows that based on a weekly prescribed visit, only 28.8% of visits were delivered on average. The proportion is relatively similar in the different phases of the programme, however a greater proportion of visits were delivered in the prenatal period. When a prescribed fortnightly visit is used as an indicator, the proportion of home visits delivered is 57.7%. Thus, over half of all prescribed visits were delivered.

Table 4.4 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. On average, each visit was 56.3 minutes long, with the shortest visit lasting 5 minutes and the longest visit lasting an hour and a half. The duration of home visits was similar across the different time periods. On average, the high treatment group spent 20.3 hours participating in home visits. The minimum duration spent in home visits was 5 minutes and the maximum was 45 hours in total. Figure 4.2 displays the variation in the duration of home visits over the entire period.

Table 4.4 - Participant Engagement in Home Visits in *PFL* up to 12 Months of Age

	Prenatal – birth	Birth - 6 months	6 Months – 12 Months	Total
Prescribed no. of home visits (based on weekly visits)	20	26	26	72
Prescribed no. of home visits (based on bi-monthly visits)	10	13	13	36
Delivered no. of home visits	6.2 (4.3) 0-21	7.6 (4.2) 0-19	6.9 (4.3) 0-17	20.7 (11.0) 1-52
% of prescribed home visits delivered (based on weekly visits)	33.7 (22.7) 0-175	29.0 (16.0) 0-73	29.5 (16.6) 0-65	28.8 (14.7) 1-64
% 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	57.7 (29.3) 3-128
Mean duration of home visits (mins)	55.1 (17.6) 5-110	59.1 (11.9) 33-91	57.8 (12.5) 15-90	56.3 (11.7) 5-86
Total duration of home visits (hours)	5.8 (4.1) 0-18	7.62 (4.6) 0-19	6.8 (4.4) 0-18	20.3 (11.5) 0.1-45
N	96	96	96	

Note: The table presents the mean, standard deviation in parentheses, and the minimum and maximum. These statistics were calculated for participants who received at least one home visit during the prenatal to 12 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. Therefore, the mean duration during the birth to 6 month period is based on 92 participants as four participants received zero visits during this time period. For the birth to 6 months period the mean duration is calculated based on 91 participants, and for 6 to 12 months period, the mean duration was calculated for 86 participants.

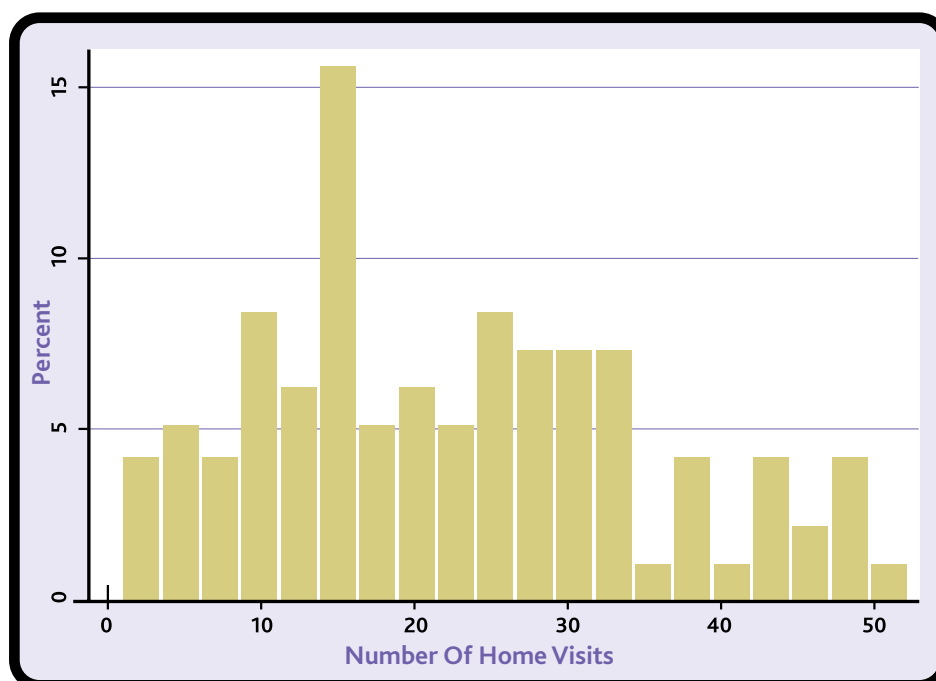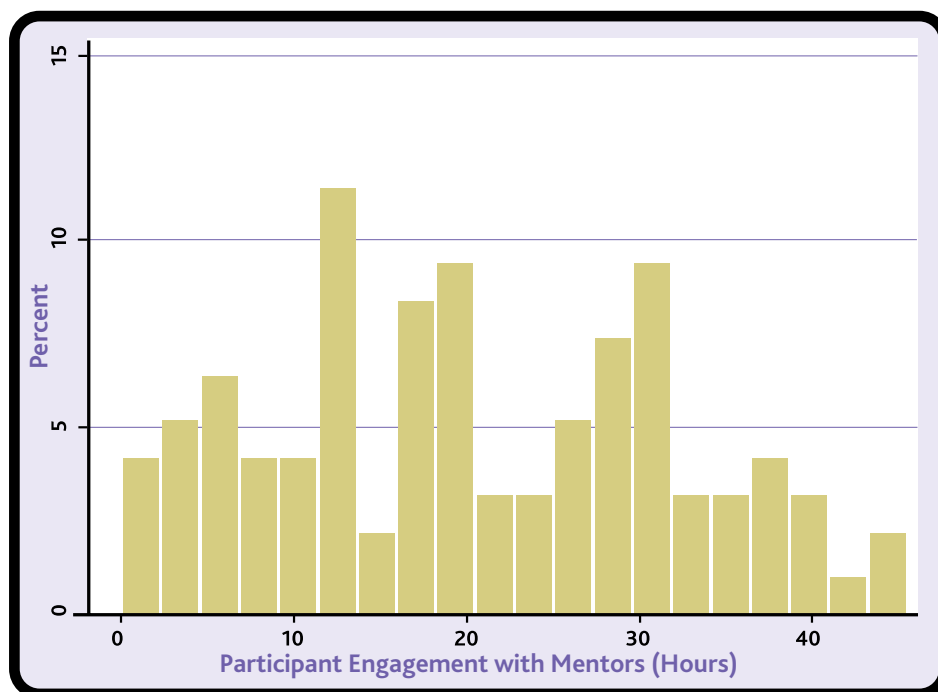**Figure 4.2 - Variation in Number of Home Visits from Programme Entry to 12 Months**

65

Figure 4.3 - Variation in Duration of Home Visits from Programme Entry to 12 Months

4.2.3 Participant Engagement from Participant Interviews

HIGH TREATMENT GROUP

Based on participant responses to the twelve month interview, 16% of participants in the high treatment group reported meeting with their mentor once a week, 61% reported meeting twice a month, 13% reported meeting once a month, and 10% 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 no participants received weekly visits, 7% received fortnightly visits, 54% received monthly visits and approximately 39% received home visits less than once a month on average. However, it should be noted that the sample used in the mentor database includes participants who received at least one home visit during the prenatal to twelve month period and therefore includes participants who have subsequently dropped out of the programme. This may explain the differing figures across the two data sources.

LOW TREATMENT GROUP

Based on participant responses to the twelve month interview, 16% of participants in the low treatment group reported meeting the information officer (IO) more than once a month, 57% reported meeting less than once a month, and 10% reported meeting their IO once or twice a year. Finally, 17% of participants reported either never meeting their IO, or only meeting their IO at the start of the programme. This corresponds to the *PFL* manual which indicates that the low treatment group does not receive any scheduled meetings. Rather, participants may schedule a meeting with the Information Officer at their discretion.

4.2.4 | 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 birth (prenatal) and birth to twelve month (postnatal). 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 the characteristics of participants who received less home visits and whether the factors associated with engagement differed in the pre- and postnatal periods.

Table 4.5 reports the relationship between maternal characteristics measured at baseline and the total number of home visits the high treatment group participated in the prenatal and postnatal periods. Within the prenatal period, it shows that 4 of the 23 maternal characteristics had a significant impact on the frequency of home visits. Specifically, mothers who joined the programme earlier in pregnancy ($p<.01$), mothers with more vulnerable attachment style during pregnancy ($p<.10$), and mother with higher self-esteem ($p<.05$) had more home visits during this period, while those who smoked during pregnancy ($p<.05$) received less home visits between programme entry and birth. Within the postnatal period, only one of the 23 characteristics was associated with engagement between birth and twelve month - mothers with higher cognitive resources ($p<.05$) had more home visits in the postnatal period.

Table 4.5 - OLS Regression Model of Frequency of Home Visits Between Programme Entry and 12 Months

Dependent Variables	Frequency of visits	Frequency of visits
	Prenatal	Postnatal (0-12)
Weeks in pregnancy at programme entry	**-0.34*** (0.05)**	-0.02 (0.11)
Mother's age	0.12 (0.10)	0.30 (0.24)
Partnered	-0.21 (1.00)	2.24 (2.32)
Married	0.06 (1.23)	-3.44 (2.85)
Living with parent(s)	0.57 (0.90)	0.92 (2.09)
First time mother	-0.32 (1.06)	2.01 (2.46)
Low education	0.92 (0.87)	2.12 (2.03)
Mother employed	0.18 (0.83)	-0.69 (1.94)
Saves regularly	-0.16 (0.79)	-1.75 (1.83)
Social housing	-1.17 (0.78)	0.57 (1.82)
Cognitive resources (WASI) †	0.02 (0.04)	**0.18** (0.09)**
Mental well-being (WHO-5)	-0.03 (0.09)	0.05 (0.20)
Vulnerable attachment (VASQ)	**0.20* (0.12)**	0.14 (0.27)
Self-efficacy (Pearlin)	-1.16 (0.86)	0.13 (1.99)
Self-esteem (Rosenberg)	**0.41** (0.20)**	0.65 (0.46)
Knowledge of child development (KIDI)	0.05 (0.06)	0.04 (0.13)
Positive parenting attitudes (AAPI)	0.04 (0.39)	-0.51 (0.92)
Physical Health Condition	0.18 (0.96)	-1.34 (2.24)
Mental Health Condition	0.64 (0.90)	2.13 (2.09)
Smoking during pregnancy	**-2.07** (0.84)**	-2.31 (1.95)
Drinking during pregnancy	-0.08 (0.84)	-0.09 (1.95)
Drugs ever used	1.25 (1.05)	2.23 (2.43)
Child is a girl	-1.09 (0.75)	-2.23 (1.74)
Constant	2.23 (6.55)	-20.00 (15.22)
N	89	89
Adj. R2	0.471	0.019

Note: Regression coefficients (Coef), standard errors (SE) and p-values obtained from an OLS regression. N= 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. † Measured at 3 months postpartum.

DURATION OF HOME VISITS

Table 4.6 shows that 5 of the 23 maternal characteristics measured at baseline had a significant impact on the duration of home visits experienced by the mothers between programme entry and birth. Specifically, mother who joined the programme earlier in pregnancy (p<.01), older mothers (p<.10), mother's with greater self-esteem (p<.05), and mothers with a male child (p<.10) spent more time in home visits during the prenatal period. While mothers who live in social housing (p<.10) spent less time in home visits. Only one participant characteristic was associated with the duration of visits in the postnatal period – mothers with higher cognitive resources (p<.05) spent more time in home visits between birth and 12 months.

Table 4.6 - OLS Regression Model of Duration of Home Visits Between Programme Entry and 12 Months

Dependent Variables	Duration of visits	Duration of visits
	Prenatal	Postnatal (0-12)
Weeks in pregnancy at programme entry	**-0.31*** (0.05)**	-0.02 (0.12)
Mother's age	**0.19* (0.10)**	0.38 (0.26)
Partnered	-0.24 (0.97)	3.04 (2.53)
Married	-0.43 (1.19)	-4.02 (3.10)
Living with parent(s)	0.41 (0.87)	0.26 (2.28)
First time mother	0.60 (1.03)	2.91 (2.68)
Low education	0.65 (0.85)	1.51 (2.21)
Mother employed	0.02 (0.81)	-0.71 (2.11)
Saves regularly	-1.17 (0.77)	-2.01 (1.99)
Social housing	**-1.29* (0.76)**	0.57 (1.98)
Cognitive resources (WASI) †	0.03 (0.04)	**0.22** (0.10)**
Mental well-being (WHO-5)	0.02 (0.08)	0.05 (0.22)
Vulnerable attachment (VASQ)	0.15 (0.11)	0.12 (0.29)
Self-efficacy (Pearlin)	-0.83 (0.83)	0.35 (2.17)
Self-esteem (Rosenberg)	**0.45** (0.19)**	0.56 (0.50)
Knowledge of child development (KIDI)	0.02 (0.06)	0.04 (0.14)
Positive parenting attitudes (AAPI)	0.12 (0.38)	-0.47 (1.00)
Physical health condition	0.15 (0.94)	-0.54 (2.44)
Mental health condition	0.59 (0.87)	2.28 (2.28)
Smoking during pregnancy	**-1.63* (0.82)**	-1.49 (2.13)
Drinking during pregnancy	-0.07 (0.81)	0.28 (2.12)
Drugs ever used	0.52 (1.02)	2.18 (2.65)
Child is a girl	**-1.41* (0.73)**	-1.98 (1.90)
Constant	-0.45 (6.36)	-25.06 (16.58)
N	89	89
Adj. R2	0.451	0.038

Note: Regression coefficients (Coef), standard errors (SE) and p-values obtained from an OLS regression. *N*= 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. † Measured at 3 months postpartum.

4.2.5 | Key Findings

The analysis of participant engagement found that families in the high treatment group received an average of 21 home visits by the *PFL* mentors between programme entry and twelve months, representing just 29% of prescribed home visits when based on weekly visits, however this figure increases to 58% when fortnightly visits is used as a guideline. While the original *PFL* manual was based on weekly visits, in reality, fortnightly visits are a more realistic measure of prescribed visits. Thus, over half of all home visits were delivered between programme entry and twelve months. This corresponds to the participant data whereby the majority of participants reported meeting their mentor twice a month (61%) and just 16% engaged in weekly home visits. In addition, it is consistent with the majority of home visits programmes which are typically based on bi-monthly visits. The average number of home visits delivered during the first six months (7.6) and second six months (6.9) were broadly similar, suggesting that a regular pattern of visits has been established between the mentors and participants. The average duration of home visits was in line with the *PFL* manual which recommended that each visits last between 30 minutes and two hours. The average duration of home visits during the six to twelve month period was just less than 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 implementation 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 twelve months indicates that relatively few individual participant characteristics were associated with the frequency or duration of home visits. In addition, the factors associated with engagement appear to differ in the pre- and postnatal periods, with individual characteristics playing a greater role in the prenatal than the postnatal period. Two factors were associated with both the frequency and duration of home visits in the prenatal period – the timing of programme entry and self-esteem. As expected, mothers who entered the programme earlier in pregnancy had more home visits and subsequently spent more time in the programme. The association between mother's self-esteem and programme engagement is rarely examined in the literature, yet it can be linked to theory. Mother's with higher self-esteem may engage in more home visits as they may have the confidence to schedule and attend visits with their mentors. They also may be more open to receiving new information and guidance on parenting practices.

The other factors which were associated with engagement in the prenatal period included mother's age, attachment style, residing in social housing, child gender, and smoking during pregnancy. Consistent with the literature, older mothers are more likely to engage by spending more time in home visits (McGuigan et al. 2003; Daro et al. 2003). While younger mothers are arguably more at-risk of poor parenting practices and thus may benefit from home visiting programmes, older mothers, through experience, may recognise the difficulties of raising young children and thus have a greater incentive to engage with the programme. The association between mother's attachment style and engagement may be related to the mentoring aspect of the programme. As mentoring is based on building a one-to-one relationship with participants, mothers with vulnerable attachment may appreciate the efforts of the mentors to engage with them and may be more likely to reciprocate this relationship. We also found evidence regarding risk status and engagement - mothers who smoke during pregnancy and live in social housing engaged less. Finally, we found that mothers with a male child spent more time in home visits. There is some precedence for this in the literature, as typically more home visiting effects are identified for girls (e.g. Anderson, 2008; Campbell et al., 2002; Heckman et al., 2010). Therefore, it may be that mothers of male children are in need of greater programme support.

Interestingly, few participant characteristics were associated with engagement in the postnatal period. One potential explanation is that as participants spend longer in the programme, and build deeper relationships with their mentors, their patterns of engagement become more established such that their individual

characteristics no longer play a role in determining their engagement patterns. The only factor related to engagement in the postnatal period was cognitive resources. 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 (Kitzman, et al., 1997; Baker et al., 1999). That mothers in the *PFL* sample with higher cognitive resources participated in more home visits and had visits of a longer duration suggest that engagement also may 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 marital status, employment status, and socio-emotional functioning were not associated with engagement in *PFL*. In addition, it is important to note that the level of engagement was not associated with socioeconomic factors, (i.e. education, employment) or parenting behaviour. This analysis was restricted to engagement from programme entry to twelve months; it is possible that the individual factors associated with engagement may change over time. However, much of these findings are consistent with the analysis of engagement reported in the six month report, with very few individual characteristics associated with engagement at either time point. The findings regarding the role of the timing of programme entry, mothers age, vulnerable attachment style, smoking during pregnancy, and cognitive resources are consistent in both the six and twelve month analyses, however, the significant findings regarding engagement and physical health and drug use during pregnancy reported at six months were no longer associated with engagement between programme entry and twelve months. Future reports will continue to analyse and monitor engagement patterns.

4.3   *PFL* Participant Satisfaction up to Twelve 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 may be 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, Buysse & Tyndall, 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.

4.3.1   *PFL* Participant Satisfaction Instruments

CLIENT SATISFACTION

Client satisfaction was measured using the Client Satisfaction Questionnaire (CSQ; Turner, Markie-Dadd, & Sanders, 1998). The CSQ (α=.88) 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 1-12 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 twelve measures except the 'improved relationship with partner' measure, as this was a rooted 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.

4.3.2 | *PFL* Participant Satisfaction Results

First, 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, with mean values of 69.07 and 61.81 for each group, respectively ($p<.01$, d=0.83). Ten of the individual 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 quality of the services received (d=0.59); whether the participant received the type of help they wanted (d=0.65); whether the programme met the child's needs (d=0.68); whether the programme met the mother's needs (d=0.66); the amount of help the mother and child received (d=0.75); whether the programme helped the mother to deal more effectively with their child's behaviour (d=0.76); whether the programme helped the mother to deal more effectively with family problems (d=0.80); whether the programme improved the mother's relationship with their partner (d=0.41); overall satisfaction with the programme (d=0.57); and whether the programme helped to develop skills that can be applied to other family members (d=0.84). There was one non-hypothesised result such that mothers' reports of improvements in child behaviour (d=0.33) were greater in the low treatment group than the high treatment group. Finally, there was no difference between the high and low treatment group regarding the mothers' feelings about their child's progress (d=0.03).

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

Table 4.7 - Participant Satisfaction: High and Low Treatment Groups

Variable	N(nHIGH/ nLOW)	$M_{HIGH\ PRE}$	(SD_{HIGH})	$M_{LOW\ PRE}$	(SD_{LOW})	Individual Test p^1	Effect Size d
Total Satisfaction	82/83	69.07	(6.65)	61.81	(10.56)	$p<0.01$	0.83
Quality of services received	82/83	6.43	(0.85)	5.76	(1.37)	$p<0.01$	0.59
Received type of help wanted	82/83	6.56	(0.82)	5.78	(1.50)	$p<0.01$	0.65
Programme met child's needs	82/83	6.26	(1.03)	5.28	(1.76)	$p<0.01$	0.68
Programme met mother's needs	82/83	6.22	(0.99)	5.36	(1.54)	$p<0.01$	0.66
Amount of help received	82/83	6.73	(0.55)	5.93	(1.41)	$p<0.01$	0.75
Help deal with child's behaviour	82/83	6.16	(1.11)	5.27	(1.25)	$p<0.01$	0.76
Help deal with family problems	82/83	5.88	(1.21)	4.94	(1.15)	$p<0.01$	0.80
Improved relationship with partner*	82/83	4.65	(1.04)	4.16	(1.37)	$p<0.01$	0.41
Overall satisfaction	82/83	6.76	(0.62)	6.19	(1.25)	$p<0.01$	0.57
Helped develop skills	82/83	6.32	(0.99)	5.10	(1.81)	$p<0.01$	0.84
Child's behaviour	82/83	5.10	(1.48)	5.55	(1.34)	s~	0.33
Child's progress	82/83	6.67	(0.77)	6.65	(0.74)	ns	0.03

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'. [1] 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.

4.3.3 *PFL* Participant Satisfaction Key Findings

Overall, participant satisfaction with the programme between six and twelve 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 the six month findings, both groups were generally satisfied with the whole programme, their child's progress and the type of help they receive from the programme. However, both groups reported being least satisfied with how the programme has improved relationships with their partner. This is consistent with the six month findings and may reflect the goals of the programme which are focused on improving child outcomes rather than family relationships more generally.

4.4 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 information is more readily transferable. Contamination is particularly undesirable in experimental evaluations as it may bias the results by reducing the mean differences between the treatment and control groups (Torgerson, 2001). Thus the reliability of the evaluation results, which are based on observations from a contaminated control group, may be questionable.

4.4.1 Measuring Contamination in *PFL*

The aim of this section is to discuss and measure potential contamination across the high and low *PFL* treatment groups between programme intake and twelve 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 twelve 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. The indirect and direct measures used to gauge contamination provided an indication of whether contamination occurred during this period.

4.4.2 Indirect Measures of Contamination

Information to track contamination indirectly was collected from participants during the twelve month interview. Specifically, participants reported how many people they know with babies the same age as their child and how many people they know in the *PFL* programme, with possible response options of 0, 1-3, 4-6, 7-10, or 10+. These questions were used to create yes/no binary variables. A binary variable was also used to indicate whether participants in the high and low treatment groups share their *PFL* materials with anyone else. All participants, regardless of which group they were in, were also asked whether they have ever been given booklets/guides on parenting or have received professional advice on parenting outside of the *PFL* programme. The results of this analysis are presented in Table 4.8 below.

Table 4.8 - Comparison of Indirect Contamination Responses across Groups

Variable	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	M_{LFP}	(SD_{LFP})	High – Low		High – LFP		Low – LFP	
							p	*d*	*p*	*d*	*p*	*d*
Knows neighbours with same age child	0.67	(0.47)	0.66	(0.48)	0.57	(0.50)	ns	.00	ns	.17	ns	.16
Knows neighbours taking part in *PFL*	0.58	(0.49)	0.69	(0.47)	0.10	(0.30)	*p*<.10	.22	*p*<0.01	1.15	*p*<0.01	1.48
Shares *PFL* material with others	0.74	(0.44)	0.56	(0.50)	~		*p*<.01	.41	~	~	~	~
Received booklets/leaflets/guides about parenting	0.33	(0.47)	0.28	(0.45)	0.61	(0.49)	ns	.11	*p*<.01	.71	*p*<.01	.59
Received professional advice on parenting	0.09	(0.28)	0.06	(0.24)	0.07	(0.26)	ns	.10	ns	.05	ns	.05
N	82		83		82							

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. 's~'indicates that the variable was significant in a one-tailed left-sided test. '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 indirect measures of contamination show that about two-thirds of the *PFL* participants and over half of the comparison group knew neighbours with the same age children as their child, and there were no significant differences across the groups. In addition, the low treatment group knew significantly more people in the programme than the high treatment group (*p*<.10; *d*=.22), and the comparison group knew significantly less people taking part in the programme than the high (*p*<0.01; *d*=1.15) or low (*p*<0.01; *d*=1.48) 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. The analysis also shows while a large proportion of the high (74%) and low (56%) treatment groups share their *PFL* materials with others, a significantly greater proportion of the high treatment group share the information they receive as part of the *PFL* programme with other parents in the community (*p*<.01; *d*=.41). 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 suggests that there was a potential for contamination across the groups, although we do not know who they shared their material with.

Another potential source of contamination may occur if the control groups seek out treatment from other sources, in this case, information on parenting. Table 4.10 shows that there was no difference between the high treatment group and the low treatment group regarding whether they had received booklets or guides on parenting from non-*PFL* sources or received non-*PFL* professional advice on parenting. Indeed, more of the high treatment group reported receiving external material (33%) and advice (9%) on parenting than the low treatment group (28% and 6% respectively), suggesting that the low treatment group is not actively seeking out parenting material elsewhere. However, a significantly greater proportion of the comparison group (61%) received external parenting materials compared to the high treatment group (*p*<.01; *d*=.71) and the low treatment group (*p*<.01; *d*=.59). As this group is not aware of the experimental programme or its content, this is not indicative of systematic contamination. In addition, there were no differences across the three groups regarding receiving non-*PFL* professional advice on parenting.

4.4.3 Direct Measures of Contamination

While the indirect measures only provide an indication of the likelihood of contamination, they cannot be used to directly determine whether contamination occurred. A contamination question was included in the twelve month survey to elicit a direct measure of contamination. This question asked participants from the high and low treatment groups if they have heard of the Triple P Parenting Programme, which is one of the *PFL* supports which is delivered to parents in the high treatment group only when their children are at least two years old. However, prior to the twelve month interview, the mentors would regularly discuss the Triple P programme with parents during the home visits, thus the high treatment group should have knowledge about Triple P at the time of the twelve month interview. On the other hand, the low treatment group should not report knowing about Triple P as it is not one of the services they receive. Thus if a high proportion of the low treatment group, relative to the high treatment group, report knowledge of Triple P it is suggestive evidence that the high treatment group shared information or materials on the Triple P Programme with the low treatment group.

Table 4.9 - Comparison of Direct Contamination Responses across Groups

Variable	M_{HIGH}	(SD_{HIGH})	M_{LOW}	(SD_{LOW})	*High – Low*	
					p	*d*
Have you heard of Triple P?	0.45	(0.50)	0.04	(0.19)	*P*<.001	1.10
N		82		82		

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. 's~'indicates that the variable was significant in a one-tailed left-sided test. '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 4.9 indicates that the high treatment group reported a significantly greater knowledge of Triple P than the low treatment group (*p*<.001; *d*=1.10). While 45% of the high treatment group stated that they had heard of Triple P, only 4% of the low treatment group stating hearing about it. This suggests that minimal contamination occurred between the high and low treatment groups between intake and twelve months.

4.4.4 Key Findings

Overall, the contamination analysis revealed three important findings. First, 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* catchment area and stated that they regularly share material from the programme with each other. It is important to note that there were no statistical differences across the high and low treatment group regarding the proportion knowing neighbours in the community and reporting sharing of *PFL* information. Interestingly, the proportion of participants reporting knowing neighbours with children the same age as their children and knowing neighbours who are taking part in *PFL* has remained constant between six and twelve months in the high treatment group, yet has increased in the low treatment group. While the proportion reporting sharing their *PFL* materials with others has increased between six and twelve months among the high treatment group and fallen slightly in the low treatment group.

Second, while the conditions for contamination were present, the analysis of the direct measures of contamination using the Triple P reference suggests that contamination between the high and low treatment group is minimal at twelve months. Very few participants in the low treatment group reported knowledge of Triple P which is one of the key components of the *PFL* programme. Thus, the overall level of contamination up to twelve months is quite low and is not a concern regarding biasing the twelve-month outcome results. However, it is important to note that only one measure of contamination was included in the twelve month assessment, thus it is possible that the low treatment group are implementing some of the other parenting strategies and advice provided by the mentors to the high treatment group.

Third, the indirect measures of contamination validate the use of the matched comparison group as a safeguard against contamination as a relatively small proportion of the comparison group stated they knew other people in the programme. These results are as expected as there were no opportunities for the comparison group to interact with the *PFL* treatment groups. However, there is some evidence to suggest that the comparison group sought out advice and materials on parenting from other sources to a much greater extent than either the high or low treatment groups. This finding is perhaps unsurprising as this group are not receiving any parenting supports from *PFL*. However, this finding does not preclude the comparison group as a suitable control group for evaluating the impact of the *PFL* programme, as participants in this group are not aware that the other groups are receiving parenting support. Thus, their actions can be interpreted as 'services as usual' rather than actively attempting to access the *PFL* treatment.

Contamination will continue to be monitored in all future evaluation reports.

Chapter Five

Report Summary & Conclusion

Overview

This report presented the results on the effectiveness of the *Preparing For Life* programme between programme entry and when the *PFL* child was approximately twelve months of age. It included an analysis of the quantitative information derived from interviews with *PFL* participants and implementation data from *PFL*'s database. Overall, the twelve month evaluation suggests that the programme is progressing regarding the retention of participants and programme satisfaction and the results are in line with evaluations of other home visiting programmes, which typically identify few significant effects at twelve months. Although there were less significant differences reported between the high and low *PFL* treatment groups than at six months, measures which focused 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 home environment.

At twelve months, 6% of the variables measured were statistically significant in the non-hypothesised direction, such that the low treatment group were outperforming the high treatment group on these measures. As the report adopted an acceptance level of 10% regarding statistical significance, it is possible that the 8% of positive findings and the 6% of negative findings were random. The discrepancy in the proportion of significant differences between six and twelve month results is difficult to interpret at this stage and trends in results may become clearer as data collection continues. However, overall these findings are consistent with previous evaluations of home visiting programmes that report limited results at twelve months (Gomby, Curloss & Behrman, 1999).

The majority of the relationships in the analysis were in the hypothesised direction, with the high treatment group reporting somewhat better outcomes than the low treatment group. As outlined in Chapter 1, we hypothesised that treatment effects would be found in the domains of child health, parenting and maternal health. The results suggest partial support for our hypotheses. There were some significant findings in the domains of child health and maternal health. However, there were no significant results found in the parenting domain at twelve months. Moreover, the programme appeared to have little significant impact on the home environment and household factors and SES. Counter to our hypotheses, significant treatment effects were found in the domains of child development and social support. This is noteworthy as previous studies of home visiting programmes do not report effects in these domains at twelve months. There were findings in the non-hypothesised direction in all domains except social support. Although positive effects for parenting and the home environment were found at six months, no significant effects for child development were found at six months. One potential explanation for the differences in the findings at six and twelve months is that many of the measures which were significant at six months were not included in the twelve month survey (e.g. the HOME). While some measures (ASQ, WHO-5 and satisfaction with father involvement, for example) were present in both the six and twelve month surveys, there were a number of measures which were used at six months and not at twelve months and vice versa. In the parenting domain for example, parental locus of control, attachment and stress were measured at six months, whereas at twelve months, the parenting domain consisted of parenting knowledge and parenting attitudes. In addition, in the twelve month report, the childcare analysis took account of the differences between the groups, however as the literature regarding appropriateness of childcare at twelve months was inconclusive; we did not hypothesise which outcome was preferable. Therefore, comparing changes in significance on the same measures over time may be a more appropriate means of monitoring changes in the effectiveness of the programme rather than the percentage of significant results at each time point.

As such, the dynamic analysis included in this report captures changes in the outcomes of the participants over time. If the programme is effective, the high treatment group should experience a greater improvement in scores across time compared to the low treatment group. Overall, the results of the dynamic analysis reported few significance differences between the high and low treatment groups regarding changes in outcomes between baseline and twelve months and six and twelve months. While 7% of the results were

significant in the hypothesised direction, a further 7% were significant in a non-hypothesised direction, overall indicating few improvements across time in parent or child outcomes during the first year of life. Again, few studies identify such significant dynamic effects during this early stage of programme delivery.

With regards to dosage and timing, participants, on average, received 7 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, averaging just over 1 visit per month over the postnatal period. In total, the participants in the high treatment group received an average of 21 visits between recruitment at the 22nd week of pregnancy and when the infant turned twelve months, which represents just over half of all prescribed home visits planned, based on a fortnightly visits. However, there was minimal attrition between six and twelve months (2 participants) and participant satisfaction was high, indicating that while engagement among participants is relatively low, they are satisfied with the level of support they are receiving and they are choosing to remain in the programme.

The results comparing the high and low treatment groups to the comparison community can be interpreted as confirming the main treatment results, as well the integrity of the RCT design. The comparison of the high treatment and comparison groups to some extent mirrored the findings of the high and low treatment groups, indicating that there were slightly more treatment effects for the high treatment group compared to the comparison group (17%) than when the low treatment group were compared to the comparison group (12%), although there were also effects in the non-hypothesised direction. That there were more differences between the high treatment group and the comparison group than the main high versus low treatment group analysis, suggests that some of the supports provided to the low treatment group may be effective. In addition, in some cases different effects were found in the main analysis and the community comparison analysis. For example, both the results of the high treatment group and comparison group analyses and the low treatment group and comparison group analyses suggest that parents in the entire *PFL* sample read more to their children than those in the comparison group, while no significant difference was found in reading between the high and low treatment groups. Similarly, both the high and low treatment groups reported better child cognitive functioning than those in the comparison group, while no difference was detected between the high and low treatment groups. These findings suggest that some common programme components, such as the developmental and reading packs may have an impact on both the high and low treatment group participants.

However, as there were a number of results in the non-hypothesised direction in the comparison of the low treatment and comparison groups this suggest that the *PFL* programme is not having a significant impact on most of the outcomes for the low treatment group. This finding echoes the results of the contamination analysis which suggest that despite the high risk of contamination within the community between the high and low treatment groups, contamination was not a significant issue at twelve months into the study. As the programme progresses, the evaluation team will continue to test for potential contamination between the treatment groups. In addition, further work on the comparison group analysis, regarding the properties of the conditioning set, is on-going and may further inform the main treatment results.

The purpose of this concluding chapter is to discuss and interpret the main results comparing the high and low *PFL* treatment groups in the context of the larger report content. As such this chapter uses all the information contained within this report such as the dynamic results, relevant research literature, implementation results, attrition and engagement analyses and differential interactionsi effects to explain the main findings. The sections below discuss the main results for each of the eight key domains.

5.2 | Child Development

A far richer set of measures of child development were included in the twelve month assessment compared to the six month assessment, however the evaluation reported only two significant treatment effects within the child development domain at twelve months. While there were no effects for the majority of the child development variables, these results are promising as few evaluations of early childhood home visiting interventions report programme effects for child development at twelve months. Evaluations of

Healthy Families America, ProKind, Parents as Teachers and the Nurse Family Partnership did not identify any treatment effects at this time point and several of these studies utilised the same measures as the *PFL* evaluation, such as the ASQ and BITSEA. This dearth of results in child development is to be expected as there is huge variation in the acquisition of new skills in infancy.

Treatment effects were found for fine motor skills and developmental competence. The advantage in fine motor skills is further supported by findings in the dynamic analysis which indicate that there is a significant difference in the changes in fine motor skills between six and twelve months for the high treatment group. Fine motor skills become more deliberate as infants develop (Meggitt, 2007) and parental awareness of this may lead to parents actively encouraging activities and play that promotes strong fine motor skill development. The second finding, that the high treatment group were less at risk of socio-emotional and behavioural delay was robust and was upheld by the step-down test, which is a stringent measure of treatment effects. These changes may be a result of the *PFL* Tip Sheets. Between six and twelve months the Tip Sheets focused on child development; including cognitive development, movement, hand-eye co-ordination and social-emotional development. It is possible that the information provided increased parental awareness of developmental milestones, or indeed encouraged parents to act as role models for their infants emerging developmental skills. This would provide some tentative support for the manualised approach, and future reports will track the relationship between the timing of information provided in the *PFL* Tip Sheets and the emergence of coinciding treatment effects. Further, this may indicate that parental behavioural change in some realms may be possible in the first year of the mentor relationship, even though large scale behavioural changes have been found to be challenging to achieve in the first year of programme implementation.

There was also an effect in the non-hypothesised direction, indicating that infants in the low treatment group had more communicative gestures than infants in the high treatment group. This finding may be of concern, as early use of gestures is associated with future language ability, (Eadie et al., 2010). Furthermore, emerging communicative skills can have an effect on other areas of development, such as personal-social development (Meggitt, 2007). However, it is too early in the evaluation to assume that such a finding is a trend. We will continue to measure child communication at each data collection point.

Findings in the differential effects analysis indicate that infants of first time mothers and partnered mothers and mothers with low emotional well-being may benefit more in the areas of child development. Furthermore, both the high and low treatment groups report better child cognitive functioning than those in the comparison group although no difference was detected between the high and low treatment groups.

5.3 │ Child Health

Child Health is an important element of school readiness, as health problems can affect child development physically as well as limit the child's exposure to everyday experiences. The results for the child health domain at twelve months are comparable to the findings of other home visiting evaluations where results are typically mixed. The REACH and EIP programmes reported a number of positive impacts on child health including the number of illnesses in the first year of life (Barnes-Boyd et al., 1996), hospitalisation, and immunisation rates (Koniak-Griffen et al., 2002). Healthy Families America (Duggan et al., 1999; Mitchell-Herzfeld et al., 2005) and Parents as Teachers (Wagner et al., 1996) did not report effects on child health. Child Health effects were found at twelve months for *PFL* regarding the mother's health decisions for her child, in relation to feeding choices and immunisation rates, and as such, mirror the findings observed at six months.

The effect of the programme on immunisation rates is consistent with findings at six months and is similar to results found in other home visiting interventions (Barnes-Boyd, Fordham & Nacion, 2001; Johnston, Heuber, Anderson, Tyll & Thompson, 2006, Koniak-Griffen et al., 2002). The Tip Sheets delivered by the mentors provide information about the benefits of vaccination and this is one area that parents can easily take action as immunisations are provided free of charge in the community. Similarly, the effects on appropriate eating may be attributed to the information provided on the Tip Sheets regarding appropriate

weaning and feeding practices. The evidence from other interventions regarding feeding is mixed (Bull et al., 2004; Haire-Joshu et al., 2008; Mitchell-Herzfeld et al., 2005). Programme effects within *PFL* were found for feeding at six months, indicating that both immunisations and feeding practices are two areas of child health that are consistently being enhanced by the programme. While there were no significant effects found on infant weight, it is expected that feeding practises may affect weight as the children continue to mature.

While at six months infant breathing difficulties were more likely to be reported by mothers in the high treatment group, at twelve months they were less likely to report that their child had chest infections. The high incidence of breathing problems at six months may have been due to maternal awareness of infant breathing, while chest infections are usually diagnosed by GPs. This could be construed that mothers were more aware of breathing problems at six months, yet the actual instance of breathing problems was less than that of the low treatment group. Thus, breathing problems that were not identified at six months among the low treatment group could have resulted in the development of chest infections between six and twelve months. Without objective measures of child health however it is difficult to reliably explain the juxtaposition between the six and twelve month findings for breathing difficulties. There is little evidence that home visiting has an impact on infant breathing rates. Asthma rates will be measured when the children are older, as this condition is not normally diagnosed until the child is school age (www. hse.ie, as accessed on 26th July, 2012). However, chest infections may be an early indication of later bronchial ailments, such as asthma. Given the high rates of smoking during pregnancy in this sample, as demonstrated at the baseline and six months findings, this is one area that the *PFL* evaluation will continue to monitor.

There was a difference in the non-hypothesised direction for sleeping location. There was a Tip Sheet relating to sleep during the six and twelve month time period, which indicated that children should be sleeping in their own cot/bed/room. Despite this, the high treatment group were more likely to share a bed with their baby. This was also the case at six months, yet this finding may have been explained by the high percentage of mothers living with the child's grandparent and subsequent space issues. At twelve months however, there was no difference between the high and low treatment groups with regards to living with the child's grandparent. Consistency in sleeping arrangements however has been identified as an important factor in the appropriateness of sleeping location (Taylor et al., 2008) and perhaps the relatively higher rates of mothers sharing their bed with their child in the high treatment group is due to their earlier experience of residing with the child's grandparent. These practices may reflect habits that began in early infancy and may have been difficult to change as the child developed.

5.4 | Parenting

Studies of early childhood interventions which evaluate parenting when infants are twelve months old, typically report mixed results. There were no treatment effects on parental risk (as measured by the AAPI) and parental knowledge of child development in this study, which is consistent with findings from other home visiting interventions on similar measures. Culp et al., (2004), for example, found positive effects on the AAPI measures at twelve months, whereas Mitchell-Herzfeld et al., (2005), Wagner and Siker (2001), Wagner, et al. (1996) and Barlow et al. (2007) did not. Moreover, there were no effects found for the KIDI measure in this study, which is consistent with other studies at twelve months, such as Mitchell- Herzfeld et al. (2005) and Wagner et al. (2001).

At six months the most notable parenting effects were found in reduced parenting stress and improvement in parent–child interactions. It is important to note that the measures of parenting at twelve months are measuring different aspects of parenting. At eighteen months many of the measures of parenting are similar to those used at six months. Parenting factors were assessed at both baseline and twelve months by the KIDI and AAPI instruments. While there was a significant difference between the high and low treatment groups for KIDI scores at baseline, indicating that mothers in the high treatment group had greater knowledge about child development, there were no differences between the two groups at twelve months, nor an improvement in scores over time.

The dynamic analysis indicated that there was significant differences between the high and low treatment groups between baseline and twelve months on the AAPI measure. Relative to the low treatment group, the high treatment group performed better regarding asserting children's power and independence and worse regarding parental expectations and lack of empathy. Even though there were no significant differences in parental knowledge at twelve months, mothers in the high treatment group had significantly more knowledge about child development at baseline. Parental knowledge may affect how parents interpret child behaviours (McLoyd, 1998); therefore it may be that mothers in the high treatment group, having expected more dramatic progress at baseline, were more likely to underestimate their children's developmental progress over time.

Although there were limited findings in the parenting domain for the overall sample, interaction results indicated that there were some positive effects on parenting for particular types of parents. Most notably, parents of girls were more likely to show programme effects in this domain. Furthermore, although no significant differences were found between the high and low treatment groups in the number of mothers who read to their children, both high and low treatment mothers were significantly more likely to read to their children than those in the comparison group. Given that both high and low treatment families receive developmental book packs from *PFL*, this finding may be directly related to programme influence. We may begin to see patterns emerge as to which factors of parenting might be most affected by the programme as data collection continues.

5.5 | Home Environment

At twelve months, the home environment domain consisted of a measurement of family functioning and interactions as measured by the Family Environment Scale. There were no positive treatment effects found in this domain, which is consistent with other studies of home visiting interventions. Of the few studies that did find effects, there were improvements in 'conflict tactics' within the family (Duggan et al., 1999) and family communication (Pinkser & Geoffroy, 1981; Wood & Davidson, 1987).

There was one difference in the non-hypothesised direction. The low treatment group reported having more cohesive family relationships than those in the high treatment group. An argument could be made that maternal expectation of appropriate family functioning may have been raised by the programme, resulting is dissatisfaction with the current levels of cohesiveness.

As with the parenting measures, different measures of the home environment were used at six and twelve months, therefore a direct comparison of the impact of the programme on the home environment over time is not possible. The measures used at twelve months relate to the family characteristics rather than the physical aspects of a child's home environment. However, interaction analyses revealed that some types of participants may have benefited from the programme in the realm of the home environment. Non-partnered parents may have more programme effects in the area of family relationships in this domain. Additionally, low risk parents and those with high emotional well-being in the high treatment group had more family organisation. Furthermore, although no significant differences were found between the high and low treatment groups in the home environment domain, the high treatment vs. comparison group analyses revealed that the high treatment group reported less family conflict than those in the comparison group. As we continue to gain data over time, patterns may emerge as to what aspects of the home environment may be most affected by the *PFL* programme.

5.6 | Maternal Health & Wellbeing

The home visiting literature contains many measures of maternal health and wellbeing, however few report differences as a result of the intervention at twelve months (Mitchell-Herzfeld et al., 2005; Koniak-Griffin et al., 2002). An evaluation of Healthy Families America (Duggan et al., 1999) reported unfavourable outcomes for mothers' general mental health, whereas an evaluation of ProKind reported positive effects in the areas of contraception, parental stress and self-efficacy (Jungman et al., 2012).

Similarly there were mixed results for the *PFL* programme at twelve months. Fewer mothers in the high treatment group reported consuming alcohol in the previous six months. This finding was robust and drove the findings in the step down category, and also had the largest effect size in this domain. The reduction in substance use may be due to an increased awareness of the risk posed to infants by parental alcohol usage. Providing a safe environment for children is paramount in the education process in *PFL* and it is a focus of the Tip Sheets in the first year.

Conversely, the high treatment group were more worried about getting a good night's sleep, rated their health lower than mothers in low treatment group and reported lower self-efficacy. It is possible that the programme raised maternal awareness of the importance of maternal health and sleeping patterns, which may have caused the high treatment parents to become more concerned about these issues. The negative finding for self-efficacy however may be of concern, as a goal of the *PFL* programme is to empower mothers in their parenting practices. At baseline there were no significant differences between the groups on this measure, however the Differences-in-Differencess analysis did not find a significant difference between the changes in maternal self-efficacy over time. We will continue to track the trajectory of self-efficacy as the programme progresses.

There were few differences found in the interaction analysis, indicating that the programme is affecting mothers in a similar fashion with regards to health, regardless of risk, cognitive resources and emotional well-being.

5.7 | Maternal Social Support

Social support continues to play a protective role in the first year of a child's life, particularly in relation to maternal wellbeing, parenting and mother-child interaction. While a number of evaluations of home visiting interventions measure the impact on social support, to our knowledge, none report any positive effects at twelve months. Within *PFL* mothers in the high treatment group were more likely to meet regularly with friends, a finding that was also evident at six months. It is possible that this effect can be interpreted as an inherent feature of being in the high treatment group which may provide more opportunities for socialising with other parents.

The main treatment results were mirrored in the interaction analyses, and there was an additional effect on father involvement for not-partnered mothers. This indicates that the programme may be increasing maternal awareness of the importance of father involvement for mothers who are not in a relationship with their child's father. These results are positive as social support is linked with other outcomes of intervention such as mother-child interactions, alleviating stress (Green, Furrer & McAllister, 2007) and enhancing maternal wellbeing (Levitt et al., 1986).

There were also effects on the two voting measures in this domain, with moderate effect sizes. This result is consistent with findings at six months. As discussed in the six month report, further information is required to identify whether this result is due to social desirability, and if not, what factors of the mentoring intervention might be associated with increased democratic participation. Pathways that have been suggested include a sense of empowerment, the association with the programme and government funding or a greater sense of community participation due to engagement with the programme.

5.8 Childcare

There were differences between the high and low treatment groups' use of childcare and while this analysis did not hypothesis the direction of the results, the findings are interesting to consider in the context of the intervention. There is scant literature on the effects of home visiting intervention on childcare choices, yet there are reports of favourable outcomes for 'service use' in general. Evaluations of both Healthy Families America (LeCroy & Crysik, 2011) and CBFRS (Culp et al., 2004) report favourable outcomes for active use of community resources as a result of intervention.

The reasons for using childcare vary, depending on parental employment, the availability of childcare and family structure. The low treatment group were more likely to use childcare at twelve months compared to the high treatment group and the type of childcare was more likely to be formal childcare. Less than one-third of the high treatment group were in childcare by twelve months, compared to less than a half of the low treatment group. It is of note that there were no significant differences between maternal or paternal full or part time employment rates, or indeed in earnings, indicating that the lower rates of childcare use for the high treatment group may be due to other factors. In addition, the low treatment group stated during the baseline assessment, prior to the child's birth that they intended to use more childcare than the high treatment group. Therefore, their level of usage at twelve months is consistent with their predictions regarding childcare use. This suggests that the differences observed at twelve month may be unrelated to the programme itself.

The high treatment group were more likely to be satisfied with their childcare and this finding had the largest effect size in the whole analysis. This may be understood in light of the higher instance of informal childcare or due to the selection of more appropriate placement for their child as a result of the programme. Tip Sheets provided between six and twelve months refer to childcare in the context of returning to work and the options available. It is also possible that there was an element of social desirability as the local crèche has a working relationship with the *PFL* programme and it is possible that mother's may have felt it inappropriate to air any dissatisfaction with the crèche.

5.9 Household Factors & SES

There were few differences between the high and low treatment groups in regards to demographics and socio-economic factors. Only one demographic factor changed significantly since six months. At six months there were significantly more high treatment mothers living with their child's grandparent. At twelve months this was no longer the case. While this may have been influenced by the *PFL* programme (who encouraged mothers to seek independent accommodation), there were few other programme driven changes in household factors or SES at twelve months. The only significant finding at twelve months was that mothers in the low treatment group were more likely to be unemployed long-term than those in the high treatment group. However, there were no treatment effects found on any other financial or employment measures. This is in line with other research on the effects of early childhood intervention on household factors. Of the studies that do measure household economic variables, most do not report intervention effects (Anisfeld et al., 2004; LeCroy & Crysik, 2011; Olds et al., 2002; Wagner et al., 1996). Furthermore, unfavourable effects were found for maternal employment rates in one evaluation of Healthy Families America (Mitchell-Herzfeld et al., 2005).

In the non-hypothesised direction, families in the low treatment group were more likely to report that their financial situation had improved in the past six months and also fathers were less likely to be unemployed. At this stage of the evaluation there is no clear accounting for this finding. It is possible, however, that the programme raised expectations of those in the high treatment group which may have fostered more dissatisfaction with their financial situation.

The limited findings in the household factors and SES domain indicate that the randomisation has been, on the whole, maintained at twelve months.

5.10 | Further Work & Future Reports

While overall there were few significant treatment effects identified at twelve months, it is important to consider that programme effects may be difficult to detect at this early stage of the intervention. Future waves of data collection will not only capture treatment effects when the programme has been running for a longer amount of time, but will be able to track changes for treatment families at multiple time points. A number of measures used at six months, which were not included at twelve months, will be included in the eighteen month survey. Therefore, we will begin to form a more in-depth picture of what kinds of factors are most affected by the programme. In addition, as more data is gathered, we will be able to expand the dynamic analysis to compare the developmental trajectories of the children over time.

Future reports will continue to track the effectiveness of the *Preparing For Life* programme when the *PFL* cohort is eighteen, twenty-four, thirty-six, and forty-eight months of age. Thus, this is the second of six reports analysing the effectiveness of the programme and should be interpreted in the context of representing the early impact of the programme.

References

Please see the following website for the reference list:

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