

Report

Evaluation of PlaySteps: A play based parent-child interaction program

Prepared by Jan Matthews & Anastasia Pourliakas

April, 2012

Parenting Research Centre

Contents

	Pages
	<i>Tables</i> 2
	<i>Figures</i> 3
Project description	4
○ Background	4
○ Rationale	5
○ PlaySteps Program	5
○ Key Dates	5
○ Project Management and Roles	6
○ Ethics Approval	6
○ Aims	6
○ Participants	7
○ Measures	16
○ Procedure	19
○ Evaluation Design	20
Reporting Findings/Data Analysis	21
○ Comparison of ‘completers’ with ‘non-completers’	21
○ Parent-child Interaction	22
○ Parental Wellbeing	25
○ Parental Enjoyment and Confidence	28
○ Child Social-Emotional Competence	32
○ Parent Professional and Social Networks	33
Summary of findings	37
○ Improvements in parent-child interaction	37
○ Improvements in parental wellbeing	37
○ Improvements in parental enjoyment and confidence	38
○ Improvements in children’s social and emotional competence	39
○ Improvements in parent’s professional and social networks	39
Comparison with QEC PlaySteps Evaluation 2010	41
Conclusions and recommendations	43
References	44
Appendix 1	46

Tables

		Pages
Table 1	Recruitment and retention rates over the three time periods	7
Table 2	Recruitment and retention rates for each early parenting centre	8
Table 3	Comparison of NCAST scores at T1 and T2	22
Table 4	Mean and standard deviation scores for the NCAST observations at all three time points	23
Table 5	DASS scores at T1 and T2	25
Table 6	Means and standard deviations for the DASS at three time points	26
Table 7	DASS subscales: Percentage of participants in the 'clinical' range	28
Table 8	Parent Enjoyment and Confidence at T1 and T2	29
Table 9	Mean and standard deviation scores for PSOC and PPS at all three time points	30
Table 10	Total number of participants in each risk category of the ASQ:SE at all three time points	32
Table 11	Number of participants in each risk category of the ASQ:SE, excluding those with missing data (n=24)	33
Table 12	Mean and standard deviation scores for the subscales of the CLSS (n=24)	34
Table 13	Mean and standard deviation scores for the NCAST Network scales at T1 and T3	36

Figures

		Pages
Figure 1	Relationship status of PlaySteps participants	9
Figure 2	Relationship status by early parenting centre	10
Figure 3	Source of family income of participants in the PlaySteps program	11
Figure 4	Family income by early parenting centre	12
Figure 5	Educational level of PlaySteps participants	13
Figure 6	Education level by early parenting centre	14
Figure 7	Self-report caregiver wellbeing at pre-test	15
Figure 8	Self-report communication with partner at pre-test	15
Figure 9	Depression, Anxiety and Stress Scale	25
Figure 10	Mean scores for the DASS Subscales at all three time points, excluding missing data	26
Figure 11	Mean scores on the PSOC	28
Figure 12	Mean scores on the PPS	29
Figure 13	Mean scores on the PSOC Subscales across all three time points, excluding missing data	31
Figure 14	Mean scores of the PPS across all three time points, excluding missing data (n=21)	31
Figure 15	Pre to follow up change in total scores of the CLSS, excluding missing data (n=24)	34
Figure 16	Pre to follow up changes in DLC scores for all participants, excluding missing data (n=25)	35

Project Name	Evaluation of PlaySteps: A play based parent-child interaction program
Investigators	Assoc. Professor Jan Matthews Parenting Research Centre
	Anastasia Pourliakas Parenting Research Centre
Date	April, 2012

Project description

Background

PlaySteps, a parenting education program for parents of young children, was developed and delivered by the Queen Elizabeth Centre and evaluated by the Parenting Research Centre (PRC) in 2007-2010.

Following that study, the Victorian Government Department of Human Services Mental Health, Drugs and Regions Branch provided funding for the delivery and evaluation of the PlaySteps program in three Victorian early parenting centres.

Commencing in 2011, the primary purpose of this project was to further investigate the effectiveness of the PlaySteps program across the three Early Parenting Centres (EPCs) in Melbourne: the Queen Elizabeth Centre (QEC), Mercy Health O'Connell Family Centre (OFC) and Tweddle Child and Family Health Service (Tweddle). This report documents the process and outcomes of the evaluation.

The PlaySteps program is a centre-based therapeutic intervention, attended by parents and their children, designed to improve parent-child relationships and interaction through play. Members of staff of the three EPCs were involved in all aspects of service delivery, including data collection, while the Parenting Research Centre (PRC) was contracted to conduct the evaluation component of this project.

Rationale

For promoting optimal child development and behaviour and also later resilience, there is significant evidence for the importance of parenting and the parent-child relationship (e.g. Alvord & Johnson Grados, 2005; Gerull & Rapee, 2002; Lindsey, Mize & Pettit, 1997). From as early as 3-4 months of age, infants of mothers who demonstrate higher levels of maintaining attention, warm sensitivity and social responsiveness show more prosocial behaviours such as smiling, cooing and gazing towards their mothers (Legerstee & Varghese, 2001).

Although parent-child interaction has long been a focus of clinical services, the need for further empirical research looking at the specific effects of such programs on parenting enjoyment, wellbeing, and satisfaction and on child outcomes has been put forward (Kelly & Barnard, 1999). This evaluation aimed to contribute to the literature on the effectiveness of group format relationship focussed and play based parenting programs in increasing parenting capacity and producing positive outcomes for families. It is particularly significant as it aims to further build on evidence gained in the initial PlaySteps evaluation (details of which are provided at the end of this report). It was of interest to examine the extent to which the results of the initial evaluation were replicated in the implementation of PlaySteps in two additional EPCs in Melbourne. This project also represents a landmark collaboration between the three EPCs - QEC, OFC and Tweddle, the first of its kind to date.

PlaySteps Program

PlaySteps is an eight-week play-based group program for parents and their children designed to improve parent-child relationships and interaction. As part of normal delivery of the PlaySteps program, parents are routinely asked to complete a series of questionnaires and participate in a short video-taped activity session between the parents and children at the start and end of the program. Participants in this research project were asked to complete these same measures 3 months after completing the PlaySteps Program.

Key Dates

Commencement: 30/06/11

Projected completion: 30/04/12

Project Management and Roles

The PRC provided support to all three centres for the conduct of the evaluation. The PRC entered data and was responsible for data analysis, interpretation and reporting.

Specifically, the PRC

- Prepared and submitted ethics applications to the PRC Human Research Ethics Committee, and the Mercy Health Human Research Ethics Committee.
- Developed a written protocol manual with detailed instructions for data collection storage and transfer.
- Developed plain language statements, consent forms and provided hard copies of all necessary measures.
- Developed SPSS databases for data for all time points and entered data.
- Analysed data and provided reports to the steering committee.
- Prepared and delivered a presentation at the “Inaugural Early Parenting Centre Forum: A Shared Vision – Better Outcomes for Children and Families” on the 23rd March, 2012.

Ethics Approval

Ethics approval was obtained from the Mercy Health Human Research Ethics Committee on May the 26th, 2011 (R10/44) and from the PRC Human Research Ethics Committee on the 13th of April, 2011 (APP07).

Aims

Specifically, this study aimed to examine the extent to which:

1. Participating in the PlaySteps playgroup improves parent-child interaction
2. Participation in the program is associated with changes in parental wellbeing
3. Parental enjoyment and confidence in parenting increases after participating in the PlaySteps program
4. There are changes in children’s social and emotional competence.
5. Parents’ professional and personal social networks change after participating in the PlaySteps program.

Participants

Recruitment

The following table gives details of recruitment and retention across all three early parenting centres, followed by recruitment and retention for each early parenting centre individually.

Table 1

Recruitment and retention rates over the three time periods.

Number of participants initially enrolled	Number of participants at Time 1 (pre test)	Number of participants at Time 2 (post test)	Number of participants at Time 3 (follow up)
39	36	29	25
Number of participants who did not begin intervention	Number of participants who withdrew during intervention	Number of participants not contactable at follow up	Total attrition
6	4	4	14

Of the 39 participants who enrolled/expressed interest in the evaluation, three did not do the Time 1 assessment. Of the 36 participants who completed the Time 1 measures, three did not start the program and four withdrew after starting the program. Thus, 29 participants completed the Time 2 assessment. Four of those were unable to be contacted at the time follow up assessments were done. Table 2 outlines recruitment and attrition for the three EPCs separately.

Table 2

Recruitment and retention for each early parenting centre.

	Number of participants initially enrolled	Number of participants at Time 1 (pre test)	Number of participants at Time 2 (post test)	Number of participants at Time 3 (follow up)
OFC	17	15	12	9
QEC	11	11	8	7
TWEDDLE	11	10	9	9
	Number of participants who did not begin intervention	Number of participants who withdrew during intervention	Number of participants not contactable at follow up	Total attrition
OFC	2	3	3	8
QEC	2	1	1	4
TWEDDLE	1	1	0	4

Age and gender

Caregivers' ages ranged from 21 to 42 years with a mean of 31.5 years and a standard deviation of 5.79 years. Ninety seven percent were female. Children's ages ranged from one to 37 months with a mean of 12.5 months and a standard deviation of 9.30. Sixty three percent were female.

Relationship status

Twenty-two caregivers (61.1%) were in partnered relationships (either married - 38.9%, or defacto – 22.2%). Of those not partnered, 11 (30.6%) were single, while 2 (5.6%) were divorced. One participant (2.8%) was a widow.

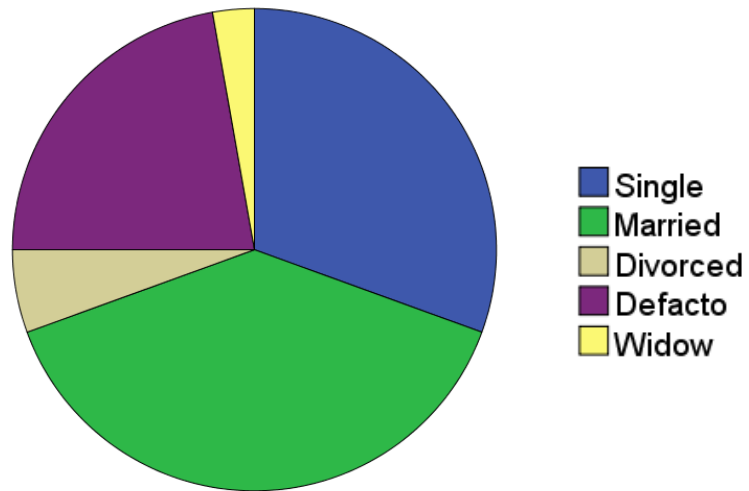


Figure 1. Relationship status of PlaySteps participants.

Of those who attended QEC, 6 were either married or defacto (54.6%), 2 were single (18.2%), 2 were divorced (18.2%) and one participant was a widow. Of those who attended OFC, 9 were either married or defacto (60%) and 6 were single (35.3%). No participant who attended OFC were divorced. Of those who attended Tweddle, 7 were either married or defacto (70%), 3 were single (30%) and no participants were divorced.

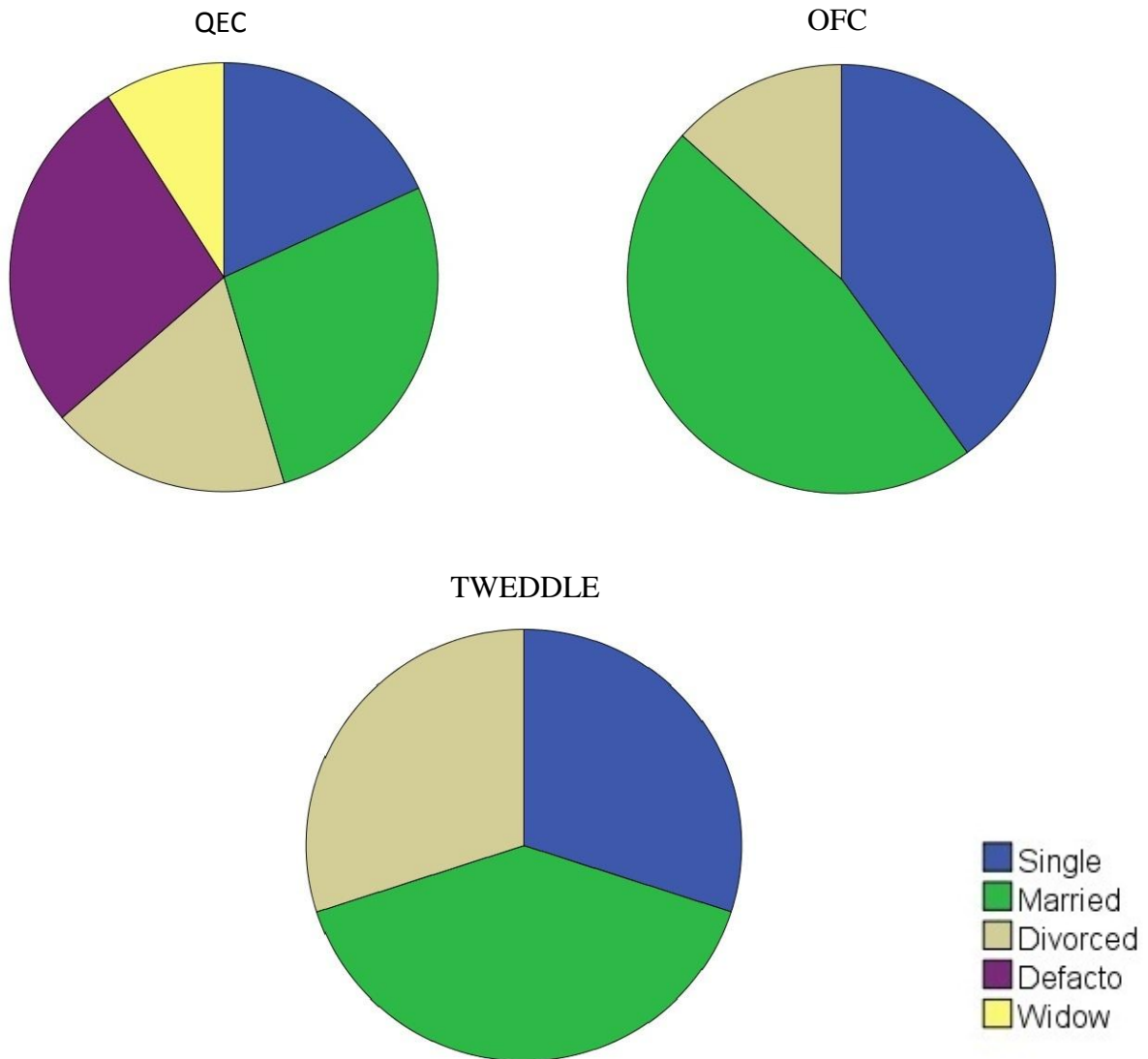


Figure 2. Relationship status by early parenting centre.

Family income

Source of family income for caregivers in this study included 11 who were employed (32.4%), 22 receiving a benefit or pension (64.7%) and 1 participant who was both employed and received a benefit or pension (2.9%). Two participants did not respond to this question.

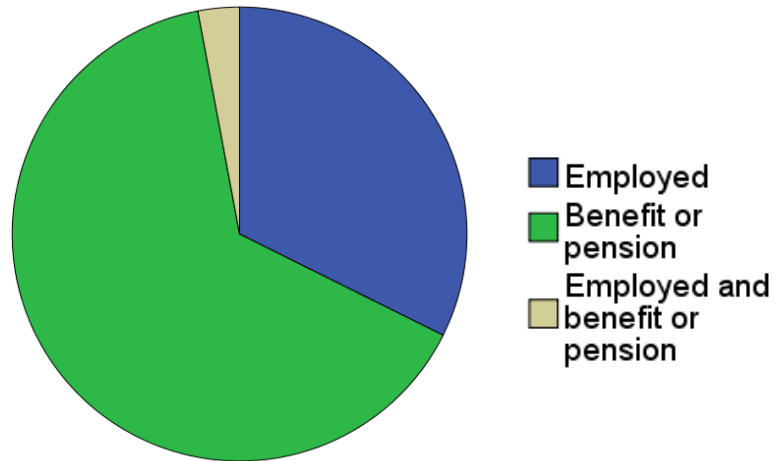


Figure 3. Source of family income of participants in the PlaySteps program.

Of those who attended QEC, one parent was employed (11.1%) and 8 were receiving a benefit or pension (88.9%). Of those who attended OFC, 7 parents were employed (46.7%) and 8 were receiving a benefit or pension (53.3%). Of those who attended Tweddle, 3 parents were employed (30%), 6 were receiving a benefit or pension (60%) and one parent was both employed and receiving a benefit or pension (10%).

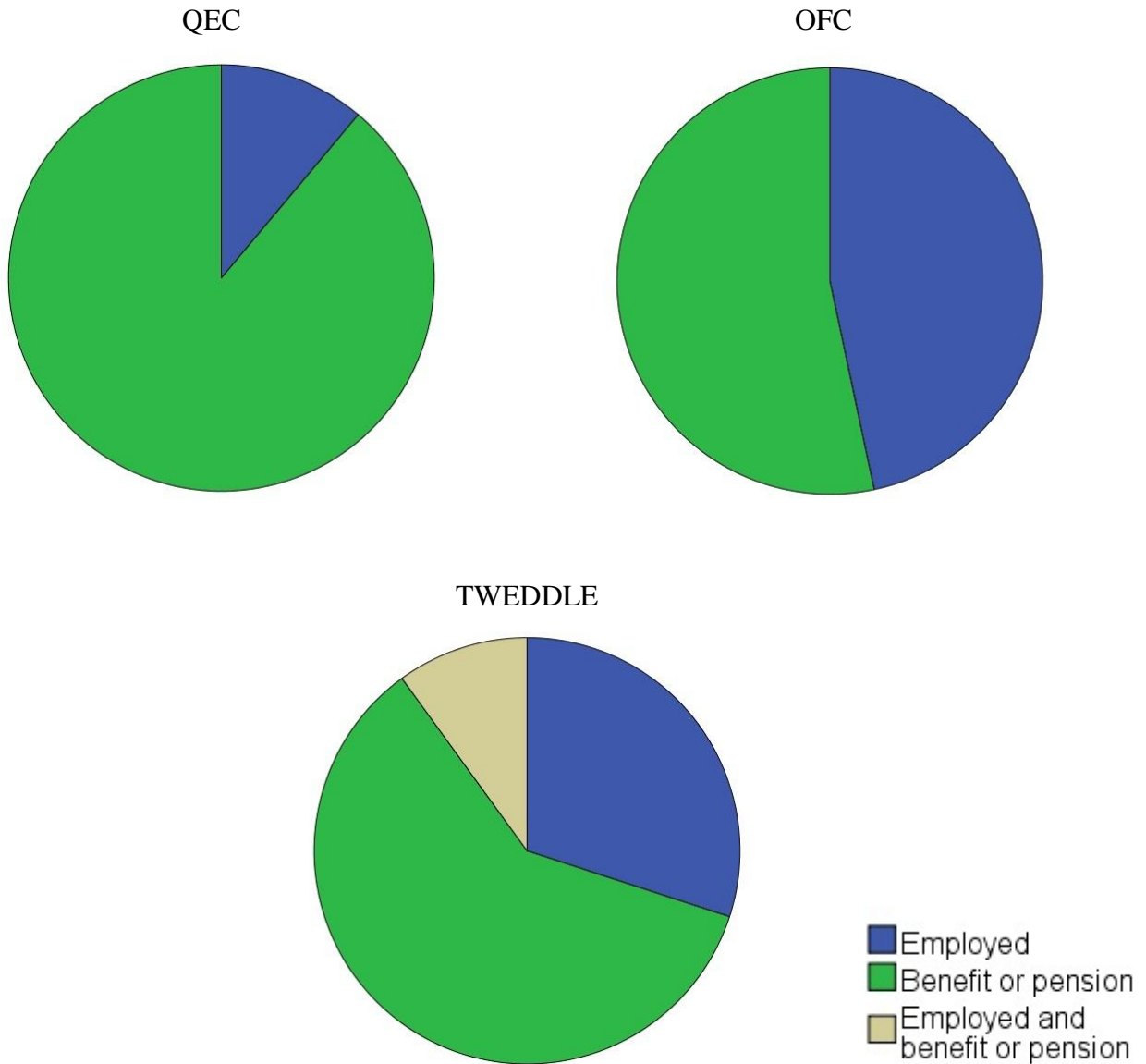


Figure 4. Family income by early parenting centre

Education

The educational level of the caregivers ranged from Year 9/10 completed (11 – 30.6%) to Graduate degree (2 – 5.6%). Twelve (33.3%) had completed year 12 or equivalent, 5 others had an undergraduate degree (13.9%) and 4 had a TAFE qualification (4 – 11.1%). The remaining percentages were those who described their educational level as 'other' or did not respond to this item.

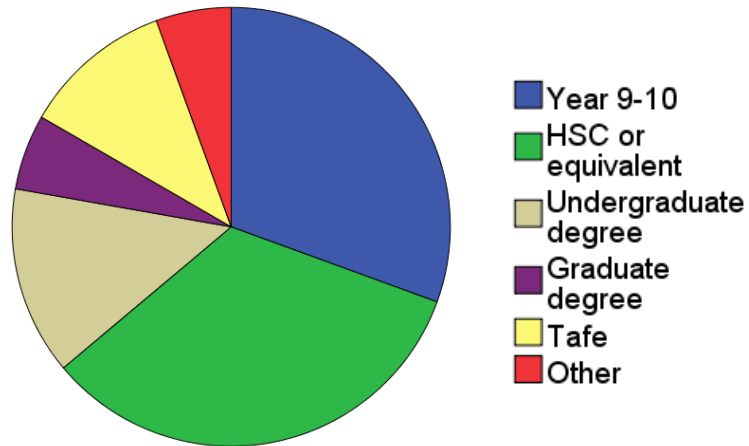


Figure 5. Educational level of PlaySteps participants.

Of those who attended QEC, 5 completed Year 9/10 (45.5%), 4 completed Year 12 or equivalent (36.4%) and one parent completed an undergraduate degree. Of those who attended OFC, 5 completed Year 9/10 (33.3%), 5 completed year 12 or equivalent (33.3%), one completed an undergraduate degree and one parent completed a graduate degree. Of those who attended Tweddle, 3 completed Year 12 or equivalent (27.3%), 3 completed an undergraduate degree (27.3%), one parent completed year 9/10, one parent completed a graduate degree and one parent complete a TAFE qualification.

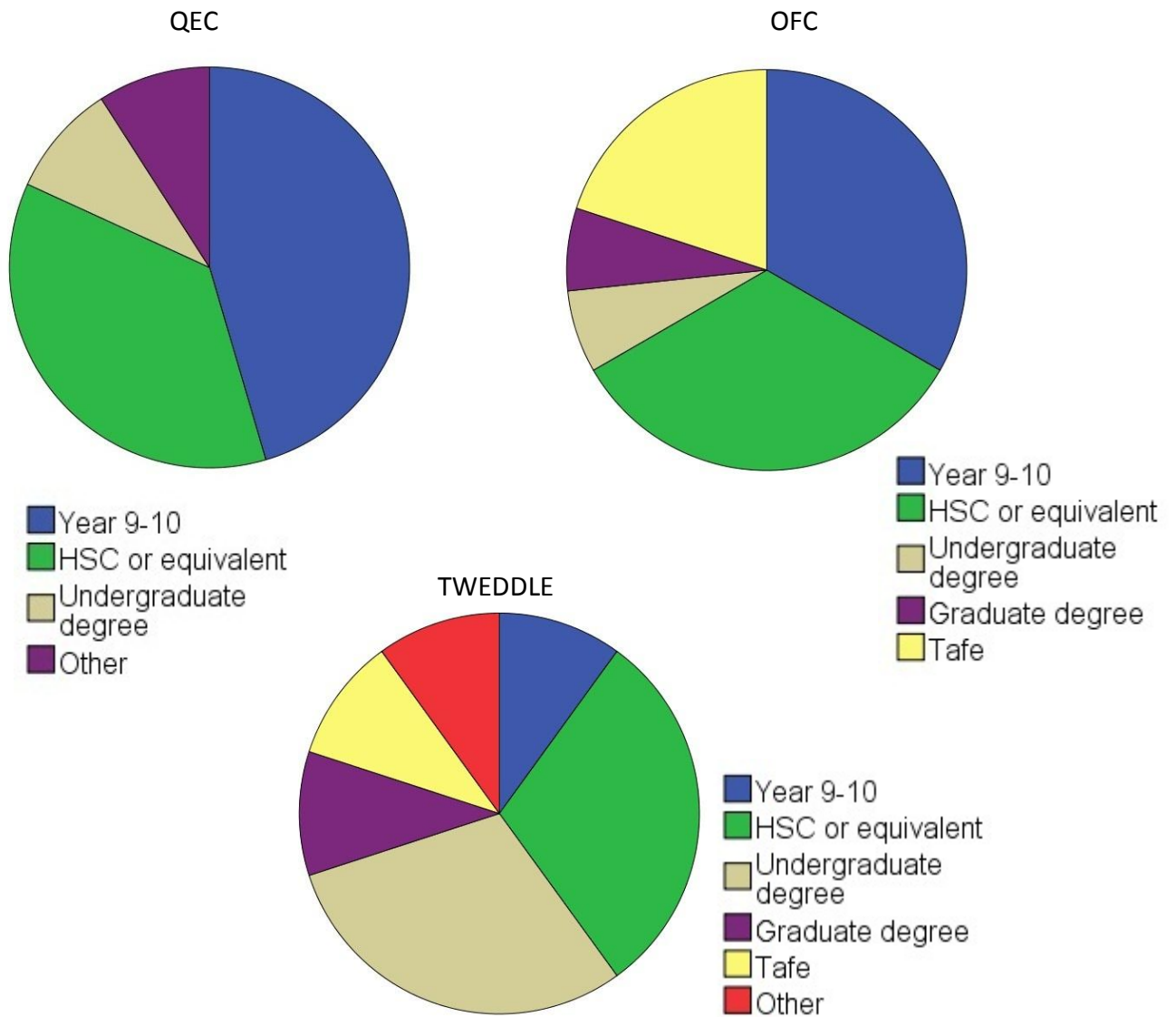


Figure 6. Education level by early parenting centre.

Other characteristics

English was the primary language spoken at home by 92.3% of participants, and 83.8% were born in Australia. Fifty three percent were previous clients of the early parenting centres.

Parent wellbeing

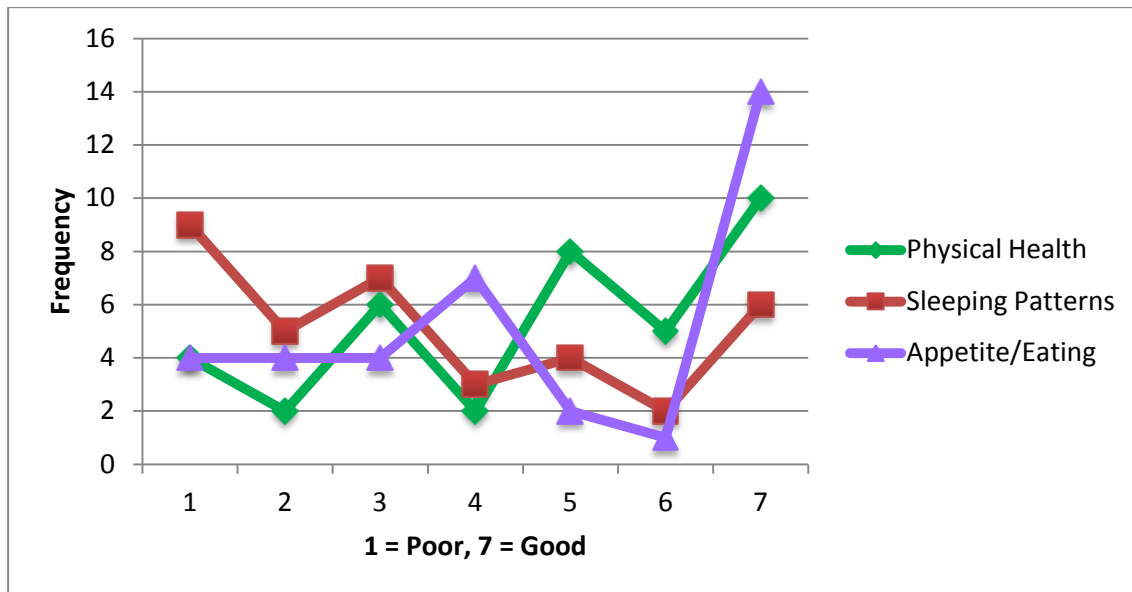


Figure 7. Self-report caregiver wellbeing at pre-test.

As can be seen in Figure 7, caregivers in the PlaySteps program reported, at the commencement of the program, that their physical health and appetite/eating habits were quite good, with majority answering 7 on a 1-7 scale. Sleeping patterns however showed the opposite trend, with most reporting sleep to be poor.

Communication with partner



Figure 8. Self-report caregiver communication with partner at pre-test.

As can be seen in Figure 8, most caregivers in the PlaySteps program said their communication with their partner was fair to excellent, with two participants stating it was poor.

Measures

Demographic Questionnaire (QEC): This is based on the QEC demographics questionnaire. It asks a series of questions about the family's health and background details. Among others, it includes education level, marital status, source of family income, whether the caregiver uses any substances, how they find their own health and details of their pregnancies. This measure is completed by parents with a staff member during the first assessment session.

NCAST Teaching Scale (Barnard, 1994): is a 73-item scale with six subscales, used to rate observations of parent-child interactions. For this project, parent-child interactions were videotaped at all three time points. They were scored by an independent rater from QEC who has been trained to criterion in NCAST assessment. On all scales, higher scores indicate more optimal parent-interaction behaviour.

Four subscales describe the parent's part in the interaction (*Sensitivity to Cues*, *Response to Distress*, *Social-Emotional Growth Fostering* and *Cognitive Growth Fostering*) and two subscales describe the child's role in the interaction (*Clarity of cues* and *Responsiveness to Caregiver*).

- The *Sensitivity to Cues* subscale reflects parental recognition and responsiveness to the child's cues.
- *Response to Distress* reflects the parent's ability to recognize distress in their child and soothe them in an effective and timely way.
- The *Social-Emotional Growth Fostering* subscale measures how well the parent creates a warm and supportive environment for the child, and encourages social exchanges.
- *Cognitive Growth Fostering* concerns the learning experiences provided by the parent for the child, and whether they use a reciprocal style of communication that encourages the child's vocalizations.
- *Total Parent* and *Total Child* subscale scores, as well as *Contingency* scores are also calculated. *Contingency* scores reflect how parent and child behaviours affect each other and shape the interaction.

Described as an approximation to a normative sample, normative scores for this scale were derived from 2,100 teaching cases provided by health-care professionals. These professionals were

establishing inter-rater reliability in their certification towards use of the scale (Sumner & Spietz, 1994).

The scale has been found to be a valid discriminator between low-risk and high-risk populations (Farel et al, 1991; Pridham et al, 2010) and is widely used in research and clinical practice with families with young children. It has also been found to correlate well with other measures of child development (Sumner & Spietz, 1994). The *Teaching Scale* demonstrates high internal consistency for the *Total score* (.87), *Caregiver total score* (.87) and *Child total score* (.81), and acceptable levels for the subscales (Sumner & Spietz, 1994). The *Caregiver total score* has good test-retest reliability whereas the infant score is not as stable.

Depression, Anxiety, Stress Scale (DASS) (Lovibond & Lovibond, 1995): is a 42-item self-report instrument with three subscales, *Depression*, *Anxiety*, and *Stress* and was administered at all three time points in this study. It measures current symptoms as experienced over the past week, on a scale of 0 (*did not apply to me at all*) to 3 (*applied to me very much*). The properties of this measure have been found to be satisfactory, with alpha coefficients for the subscales being .91, .84 and .90 respectively, and test-retest reliabilities ranging from .71 to .81 (Antony, Bieling, Cox, Enns, & Swinson, 1998; Lovibond, 1998). Higher scores indicate higher levels of distress.

Parenting Sense of Competence Scale ("Being a Parent" or *PSOC*) (Johnston & Mash, 1989): is a 16-item self-report scale with two subscales; *Satisfaction* and *Efficacy* and was administered at all time points in this study. Statements about parenting are responded to on a 6-point Likert scale from 1 = *Strongly Agree* to 6 = *Strongly Disagree*. This measure has established psychometric properties including satisfactory reliability and validity (Lovejoy, Verda, & Hays 1997; Ohan, Leung, & Johnston, 2000.) Alpha coefficients for these subscales are .75 (*Satisfaction*) and .76 (*Efficacy*). Higher scores relate to greater satisfaction and parental self-efficacy.

Pleasure in Parenting Scale (Fagot, 1995): is a 10-item Likert scale that examines how parents feel about some of the tasks of parenting (e.g. bathing the child), and was developed to tap into parental warmth towards their child. It was administered at all time points in the current study. Items are rated on a 5-point Likert scale with 1 = *dislike* and 5 = *enjoy very much*. Higher scores indicate a greater pleasure in parenting. There is evidence for test-retest reliability (.70), and evidence for the internal consistency and the validity of the scale. Alpha coefficients have been found to be .81 for mothers and .76 for fathers (Fagot, 1995).

NCAST Network survey (NET) (Brandt, 1989): this two-part survey assesses the amount and quality of both personal and professional support available to the parent. In the current study, it was

administered at pre-test (Time 1) and follow up (Time 3). It asks caregivers to list individuals they can count on to help them and their family, describe what they provide help with, rate how helpful they are and how much trouble it is for them to ask this person for help.

NCAST Community Life Skills Scale (CLSS): is a 33 item 'yes' or 'no' response scale, measuring an individual's use of community resources. Administered at pre-test and follow up, it is assessed as an interview and has six separate areas of assessment; *Transportation, Budgeting, Support Services, Support-Involvement, Interests, Hobbies and Regularity of Routines*. It has been found to be valid in predicting mothers in most need of practical help, with an alpha coefficient in the range of .63 to .69 (Bernard et al, 1999; LeCuyer-Maus, 2003). Scores range from 0-33, with higher scores indicating a higher need of help.

NCAST Difficult Life Circumstances Scale (DLC): is a 28 item yes or no response scale administered at pre-test and follow up. This scale asks individuals whether they are experiencing particular problems in their life. These problems range from financial difficulties, to health problems and whether they have been victims of physical, sexual or emotional abuse. Scores range from 0-28, with higher scores meaning more difficulties. A cut off score of 6 and above has been suggested to identify families at risk for maladaptive outcomes. Validity has been established with measures of maternal depression, physical concerns, social support, ineffective family coping and nursing diagnoses of potential for family violence (Barnard, 1994). Test re-test correlations have been found to range from 0.4 to 0.7 (Barnard, 1994) with one study finding an alpha coefficient of .69 (Le-Cuyer-Maus, 2003).

Ages and Stages Questionnaire: Social-Emotional (ASQ:SE) (Squires, Bricker & Twombly, 2004): is designed to identify infants, toddlers and preschool-age children's level of vulnerability for social-emotional problems. It was administered at all time points in this study. It can be administered at 8 age intervals (6, 12, 18, 24, 30, 36, 48 and 60 months) and is filled out by the caregiver regarding their child. Caregivers are asked to identify whether the child partakes in the described behaviour "most of the time", "sometimes" or "never or rarely". In addition, parents are asked to identify whether this behaviour is a concern to them. All items are simply worded, and range from 19 questions on the 6 months questionnaire, to 33 on the 48 and 60 month questionnaires. Cut off scores are given at each age range. Scores above the cut off are indicative of concerns that require follow-up evaluation. Scores below the cut off indicate a child who is behaving as expected for their age.

The authors of the *ASQ:SE* cite the following psychometric information. Alpha coefficients for the *ASQ:SE* overall were found to be .82. For the separate age intervals alpha coefficients were as follows: 6 months (.69), 12 months (.67), 18 months (.81), 24 months (.80), 30 months (.88), 36 months (.89), 48 months (.91) and 60 months (.91). To assess test-retest reliability, the same children were tested one to three week intervals. Percentage agreement between the two scores was found to be 94%. The rater was the same for both time points. The *ASQ:SE* was also found to have good concurrent validity, with sensitivity at 78% and specificity at 95%.

As participants were to be tested over a 3 month period in the current study, a decision needed to be made regarding how to utilize scores from participant's children who moved between age intervals on the *ASQ:SE* during this time period. For example, this would need to be done if a child was 8 months old at Time 1 and assessed using the 6 months *ASQ:SE*, and at Time 3 was 11 months and required to be assessed using the 12 month *ASQ:SE*. To address this, participants were categorized as either 'no risk', 'near risk' or 'at risk' at all time points, using the cut off scores indicated by the *ASQ:SE* authors. Analyses were conducted to determine whether participants had moved from either of the risk groups to the no risk group.

Procedure

The project aimed, initially, to recruit a minimum of 48 families attending the PlaySteps program at the QEC, OFC and Tweddle, with 16 at each site. Thirty-nine families were enrolled in the evaluation and 36 participants completed the Time 1 measures.

Two groups at each EPC were running each week, with one for parents of infants (0-1 years) and one for parents of toddlers. The children attended with their parents.

Participation in the research was voluntary, and eligible families were able to attend PlaySteps whether or not they agreed to take part in the study. Inclusion criteria for parents to attend PlaySteps were: identified difficulty with relating to their child/children, which may be due to factors such as post natal depression, drug and substance abuse, disability, mental health issues, limited family support, multiple births, or domestic violence.

The three time periods for the project were as follows:

Time 1 Pre-group assessment (week 0)

All parents attended an assessment session prior to the first group session. Parents were asked to complete assessment measures at this point routinely, and were invited to participate in the research project.

Weeks 2 to 9: Participants attended the eight sessions of the PlaySteps program. Each session ran for 3 hours per week.

Time 2 Post-group assessment (week 7-9)

All parents completed a second assessment after the program finished.

Time 3 Follow up 1 (week 20-24/ 3 months after program). Study participants attended a follow-up assessment session

Families returned to the EPCs (or received a home visit) for staff to administer the NCAST Teaching Scale and complete the questionnaire booklet. An incentive voucher, with a value of no more than \$20, was offered for each follow up assessment session the parents and children attended.

Evaluation Design

A true experimental design could not be implemented because of ethical restrictions on randomising participants to experimental or wait list conditions. In particular, it was thought inappropriate for parents to wait for an extended period before receiving the intervention. Therefore, this is described as a quasi-experimental design, a single group study with pre-post and follow-up comparisons.

It was expected that there would be positive change from Time 1 (Pre) to Time 2 (Post), which would be maintained or improved at Time 3 (follow up).

Reporting Findings/Data Analysis

The results for Parent-Child Interactions, Parental Wellbeing, Parental Enjoyment and Confidence, Child Social-Emotional Competence and Parent Professional and Social Networks are presented in the following report. Each of the major assessment areas will be presented in two ways: a data summary followed by statistical analyses.

- First there are graphs showing mean scores for all of the participants who completed each assessment at the three time periods. The number of participants providing data at each timepoint varied from 36 to 24.
- Statistical analysis of the extent of change from pre-test (Time 1) to post-test (Time 2) is presented next, with approximately 28 parents completing assessments at both time points with no missing data.
- This is followed by a statistical analysis of data from the 24 participants who completed assessments at pre, post, and follow up periods.
- Parent Professional and Social Networks are analysed comparing pre-test with follow up test results only.

The next section presents a comparison between those who completed assessments at all three time points (designated ‘completers in this report), and those who did not complete assessments at all three time point (‘non-completers’).

Comparison of ‘completers’ with ‘non-completers’

The terms ‘completers’ and ‘non-completers’ refer to completion of the assessment tasks, rather than completion of the intervention itself. An examination of Time 1 scores of those who provided data at the 3 assessment periods with those who did not complete all three assessment period tasks showed some differences for the *DASS scales*, *NCAST Teaching Scales* and the *DLC*. Completers were found to have higher *DASS scores* - indicating less distress, higher on *NCAST Teaching Scales* - indicating more optimal parent-child interactions scores, and lower on the *DLC* - indicating fewer difficult life circumstances. The means and standard deviations for these differences can be seen in Appendix 1. However, any noticeable differences were not found to be significant when analysed using a one way between subjects Analysis of Variance (ANOVA), with one exception; the *child-caregiver contingency score* of the *NCAST Teaching Scale*. There was a significant difference between completers and non-completers in the *child-caregiver contingency score*, $F(1,31) = , p = .049$, partial $\eta^2 = .12$, indicating a small but significant effect.

Parent-Child Interaction

Parent-child interactions were examined using the NCAST Teaching Scale. First there is an analysis of the pre-post differences, and then an analysis of pre-post and follow up differences for all subscales.

Pre-Post differences for the NCAST Teaching Scale

Table 3 shows the results from the 28 parents who completed this assessment task at Time 1 (Pre) and Time 2 (Post). Due to the within-subjects nature of the design, for all analyses, cases missing values on the subscales at any stage were list-wise excluded. Means, standard deviations and confidence intervals are reported for both time points.

Table 3

Comparison of NCAST scores at T1 and T2

NCAST	M (SD)	T1		M (SD)	T2	
		95% CI for mean Lower	Upper		95% CI for mean Lower	Upper
Parent behaviour						
Sensitivity to cues	10.00 (0.90)	9.65	10.35	10.18 (0.67)	9.92	10.44
Distress response	9.82 (1.31)	9.32	10.33	9.61 (1.13)	9.17	10.05
Emotional growth fostering	8.96 (1.07)	8.55	9.38	8.61 (0.83)	8.29	8.93
Cognitive growth fostering	11.46 (2.77)	10.39	12.54	12.43 (2.29)	11.54	13.31
Caregiver total	40.25 (4.25)	38.60	41.90	40.50 (3.27)	39.23	41.77
Caregiver-child total	58.04 (5.12)	56.05	60.02	57.89 (3.49)	56.54	59.25
Child behaviour						
Clarity of Cues	8.54 (0.96)	8.16	8.91	8.68 (0.91)	8.33	9.03
Responsiveness	8.71 (1.95)	7.96	9.47	8.71 (1.94)	7.96	9.47
Child total score	17.25 (2.63)	16.23	18.27	17.39 (2.59)	16.39	18.40
Contingent scores						
Caregiver Contingent	14.89 (3.21)	13.65	16.14	15.61 (2.11)	14.79	16.42
Caregiver-child contingent	22.96 (3.41)	21.65	24.28	23.75 (2.35)	22.84	24.66
Child contingent	8.07 (1.89)	7.34	8.80	8.14 (1.72)	7.48	8.81

One way repeated measures analyses of variances (ANOVAs) were conducted to examine differences between *NCAST Teaching Scale* scores at Time 1 and Time 2. The measure of effect size used was partial eta squared, with partial $\eta^2 = 0.01$, 0.13 , and 0.26 representing small, medium, and large effect sizes, respectively (Bakeman, 2005). The ANOVAs conducted revealed no significant differences between pre and post-test *NCAST Teaching Scores* for any of the subscale scores (all $ps > .16$). All effect sizes were also found to be small (all partial $\eta^2 = < .07$).

Pre-Post and Follow up differences for the NCAST Teaching Scale.

Table 4 shows results of the NCAST Teaching assessments for the 24 participants who completed this measure at all three time points. These data were used in subsequent statistical analyses. The asterisk indicates a significant finding. Normative scores have been taken from the NCAST manual (Barnard, 1994).

Table 4

Mean and standard deviation scores for the NCAST observations at all three time points

	NCAST Means (SD) (n = 24)			
	Population mean (SD)	T1	T2	T3
<u>Parent Behaviour</u>				
Sensitivity to cues	9.72 (1.30)	10.08 (0.7)	10.17 (0.7)	10.17 (0.7)
Response to distress	10.31(1.53)	9.79 (1.4)	9.62 (1.2)	10.04 (1.2)
Emotional growth fostering	9.56 (1.37)	8.83 (1.0)	8.63 (0.9)	8.71 (0.9)
Cognitive growth*	13.82 (2.64)	11.46 (2.7)	12.29 (2.3)	13.08 (2.0)
Caregiver total	43.41 (5.09)	40.17 (4.1)	40.71 (2.7)	42.00 (3.0)
Caregiver-child total	59.56 (6.95)	58.13 (5.4)	57.96 (3.5)	59.88 (3.9)
<u>Child Behaviour</u>				
Clarity of cues	8.20 (1.34)	8.63 (1.0)	8.63 (0.8)	8.67 (0.7)
Responsiveness to parent	7.95 (2.88)	8.71 (2.0)	8.62 (1.9)	9.21 (1.8)
Child Total score	16.15 (3.84)	17.33 (2.7)	17.25 (2.5)	17.88 (2.2)
<u>Contingency Scores</u>				
Caregiver contingency	17.35 (2.90)	15.00 (3.2)	15.63 (2.1)	16 (2.5)
Caregiver-child contingency	N/A	23.13 (3.5)	23.67 (2.5)	24.42 (2.8)
Child contingency	7.21 (2.68)	8.13 (2.0)	8.04 (1.7)	8.42 (1.5)

A one-way repeated measures ANOVA was used to analyse each measure of parent and child interaction across the three time-points. If the assumption of sphericity was found not to be met, a Huynh-Feldt correction was used to adjust for this.

Parent Behaviour: There was no significant effect of time found for *parent sensitivity to cues* ($F(2, 46) = 0.10, p = .90, \text{partial } \eta^2 = .00$), *parent response to distress*, ($F(2, 46) = .080, p = .46, \text{partial } \eta^2 = .03$), *observed emotional growth fostering*, ($F(2, 46) = 0.27, p = .76, \text{partial } \eta^2 = .01$),

caregiver total score ($F(1.67, 38.45) = 1.75, p = .19, \text{partial } \eta^2 = .07$, with a Huynh-Feldt correction), or *caregiver-child total score* ($F(1.52, 34.89) = 1.30, p = .28, \text{partial } \eta^2 = .05$, with a Huynh-Feldt correction).

However, a significant effect of time of time was found for the *cognitive growth fostering subscale* ($F(2, 46) = 3.25, p = .048, \text{partial } \eta^2 = .12$), indicating that parents improved in their ability to pick up and encourage cognitive skills in their children, with a small to medium effect. Post-hoc comparisons with a Bonferroni adjustment revealed no significant difference between Time 1 and Time 2 (mean difference = -0.83 , 95% CI $[-2.80, 1.14]$, $p = .85$), or Time 2 and Time 3 (mean difference = -0.79 , 95% CI $[-2.11, 0.53]$, $p = .40$), but a significant difference between Time 1 and Time 3 (mean difference = -1.625 , 95% CI $[-3.21, -0.041]$, $p = .04$).

Child Behaviour: There was no significant effect of time found for *child observed clarity of cues* ($F(2, 46) = 0.02, p = .97, \text{partial } \eta^2 = .00$), *child responsiveness to parent* ($F(2, 46) = 0.67, p = .52, \text{partial } \eta^2 = .03$), or *child total score* ($F(2, 46) = 0.49, p = .62, \text{partial } \eta^2 = .02$).

Contingency Scores: There was also no significant effect found across time for *caregiver contingency score* ($F(1.64, 37.80) = 0.88, p = .41, \text{partial } \eta^2 = .04$, with a Huynh-Feldt correction), *child contingency score* ($F(2, 46) = 0.31, p = .73, \text{partial } \eta^2 = .01$), or *child-caregiver contingency score* ($F(1.57, 36.05) = 1.08, p = .34, \text{partial } \eta^2 = .05$, with a Huynh-Feldt correction).

Parental Wellbeing

The mean scores for the *DASS Scale* are represented in Figure 9 for all the participants who filled in this measure at the various time points.

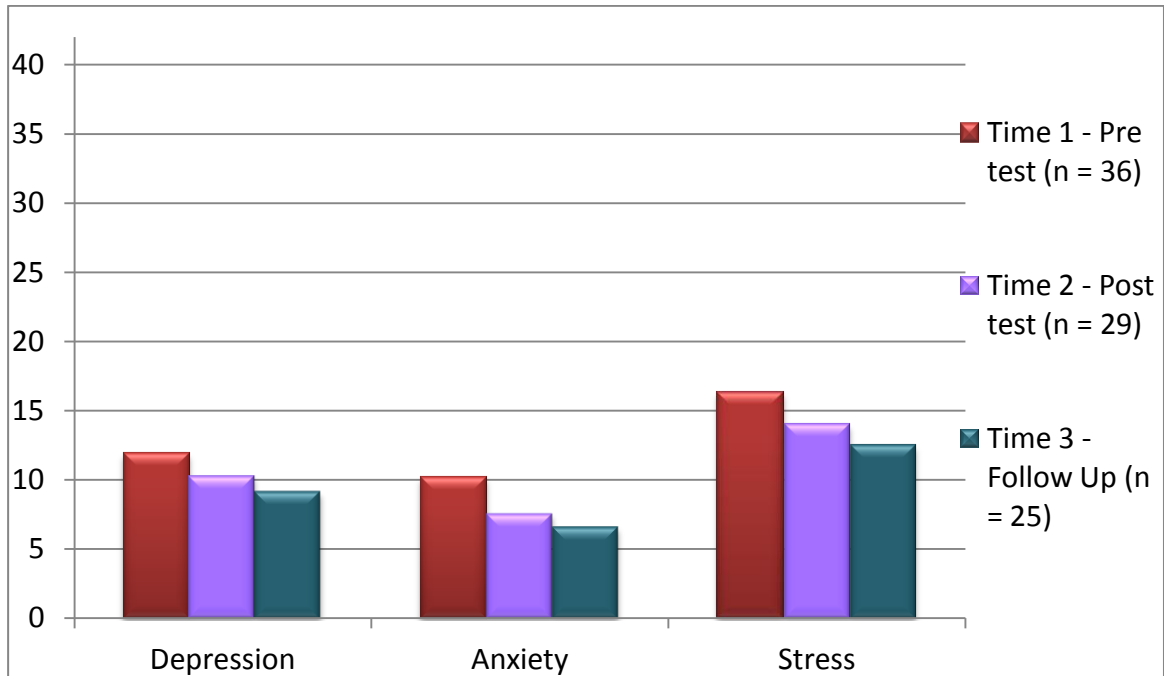


Figure 9. Depression, Anxiety and Stress Scale.

Pre-Post differences for the DASS Scale

Table 5 shows the means, standard deviations and confidence intervals for data provided by the 29 parents (n = 28 for the the *Depression* subscale) who filled in the *DASS* at Time 1 and Time 2.

Table 5

DASS scores at T1 and T2.

	T1			T2		
	M (SD)	95% CI for mean Lower	Upper	M (SD)	95% CI for mean Lower	Upper
Depression (n = 28)	13.45 (9.9)	9.68	17.22	9.90 (8.1)	6.81	12.98
Anxiety (n = 29)	11.17 (7.03)	8.50	13.85	7.55 (6.23)	5.18	9.92
Stress (n = 29)	17.48 (10.26)	13.58	21.39	13.97 (8.64)	10.68	17.25

There was a significant decrease found in parents' self-reported *Anxiety* from Time 1 to Time 2, $F(1, 28) = 17.20$, $p < .01$, partial $\eta^2 = .38$, and a significant decrease in parents' self-reported *Stress*

from Time 1 to Time 2, $F(1, 28) = 5.04$, $p = .03$, partial $\eta^2 = .15$. The decrease in *Depression* from Time 1 to Time 2 was found not to be significant, $F(1, 28) = 3.72$, $p = .06$, partial $\eta^2 = .12$.

Pre-Post and Follow up for DASS

Table 6 presents the means and standard deviations of *DASS* subscale scores of participants who filled in this scale at all three time points. Statistical analyses were conducted on these data to test for maintenance of intervention effects. With a repeated measures design, cases missing values on the subscales at any stage were list-wise excluded from analysis. This resulted in 24 completed measures of *Depression*, 25 for *Anxiety* and 24 for *Stress*. The asterisk refers to a statistically significant long term effect.

Table 6

Means and standard deviations for the DASS at three time points

	Norms	T1	T2	T3
Depression*	6.34 (6.97)	14.21 (9.96)	9.92 (8.42)	9.00 (8.75)
Anxiety*	4.70 (4.91)	10.76 (6.94)	7.28 (6.49)	6.60 (6.55)
Stress*	10.11 (7.91)	17.46 (10.35)	13.21 (8.73)	12.46 (9.04)

Norms refer to the average scores obtained from a population sample, as provided by the developers of the scale.

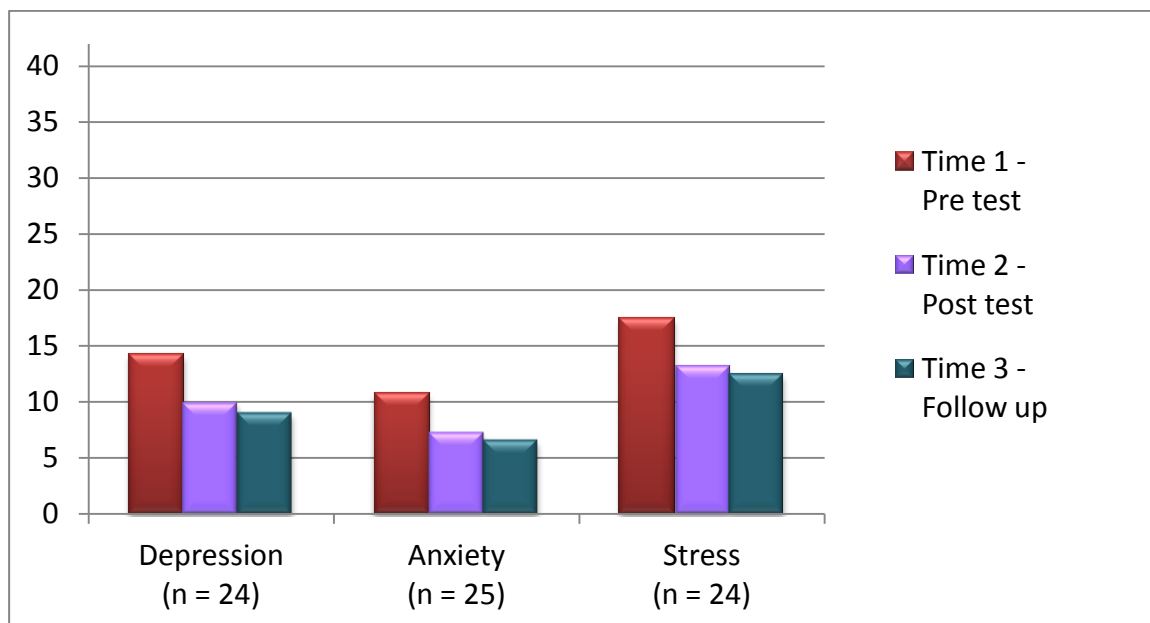


Figure 10. Mean scores for the DASS Subscales at all three time points, excluding missing data.

Depression

A one way repeated measures ANOVA revealed significant differences for the *Depression* subscale of the *DASS* over time. As the assumption of sphericity was not met, the Huynh-Feldt correction was used, $F(1.63, 37.46) = 4.81, p = .02$, partial $\eta^2 = .17$. Post hoc comparisons with a Bonferroni correction revealed there were no differences between Time 1 and Time 2 (mean difference = 4.29, 95% CI [-0.92, 9.50], $p = 0.13$). Differences between Time 1 with Time 3 were found to be approaching significance (mean difference = 5.21, 95% CI [-0.03, 10.44], $p = 0.51$).

Anxiety

A significant effect of time was found for the *Anxiety* scale of the *DASS*. Huynh-Feldt correction was once again used as the assumption of sphericity was not met, $F(1.68, 40.30) = 9.35, p < .001$, partial $\eta^2 = .28$. Post hoc comparisons using the Bonferroni correction indicated significant differences in Anxiety scores between Time 1 and Time 2 (mean difference = 3.48, 95% CI [1.27, 5.69], $p < .001$) and also between Time 1 and Time 3 (mean difference = 4.16, 95% CI [0.90, 7.42], $p < .01$).

Stress

The one way repeated measures ANOVA revealed significant differences for the *Stress* subscale of the *DASS* over time. As the assumption of sphericity was not met, the Huynh-Feldt correction was once again used, $F(1.54, 35.36) = 5.88, p = .01$, partial $\eta^2 = .20$. Post hoc comparisons, with a Bonferroni correction, revealed there were significant differences between Time 1 and Time 2 (mean difference = 4.25, 95% CI [0.22, 8.28], $p = 0.04$) while differences between Time 1 and Time 3 were approaching significance (mean difference = 5.00, 95% CI [-0.03, 10.03], $p = .052$).

Clinical significance

Clinical significance of changes in parental wellbeing were examined using scores on the *DASS*. A score above or within the 'moderate' severity range for each subscale is considered to be in the 'clinical' range. For *Depression*, this is 14-20, for *Anxiety*, 10-11 and *Stress* 19-25. A participant was then classified as being in or out of the clinical range based on these cut off scores. Table 7 represents these findings at pre-test, post-test and follow up. The percentage of participants in the clinical range is calculated by dividing the number in the 'clinical' range by the number completing the measure at that time point. For example, 11 out of 24 participants (46%) were in the 'clinical' range for the *Depression* subscale of the *DASS* at Time 1 and 8 out of 24 (33%) were in the 'clinical' range for the *Depression* subscale at Time 2.

Table 7

DASS subscales: Percentage of participants in the 'clinical' range

Subscale	Time 1	Time 2	Time 3
Depression (n = 24)	n = 11 (46%)	n = 8 (33%)	n = 6 (25%)
Anxiety (n = 25)	n = 16 (64%)	n = 8 (32%)	n = 6 (24%)
Stress (n = 24)	n = 12 (50%)	n = 6 (25%)	n = 3 (13%)

Parental Enjoyment and Confidence

This section reports on the findings of the *Parenting Sense of Competence (PSOC) Scale* and the *Pleasure in Parenting Scale (PPS)*.

Figure 11 shows the mean scores for *Efficacy* and *Satisfaction* on the *PSOC* for the participants who completed this measure at the three time points.

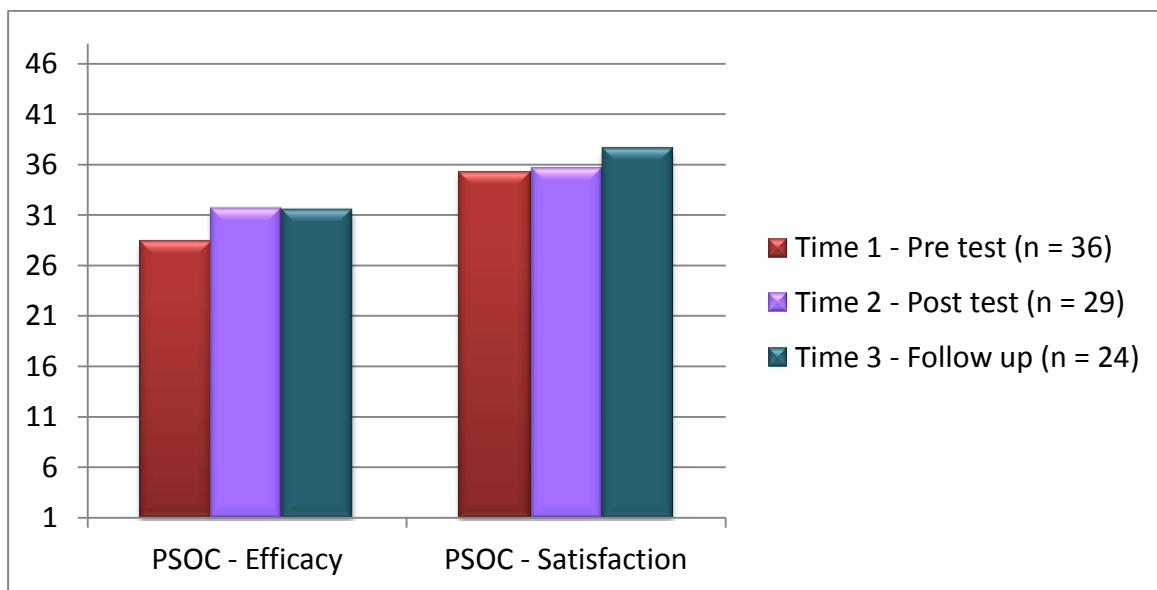


Figure 11. Mean scores on the PSOC.

Figure 12 below shows the mean scores for all participants who filled in the *PPS* over the period of data collection. The number of participants who completed the *PPS* at each of the time points is presented in brackets.

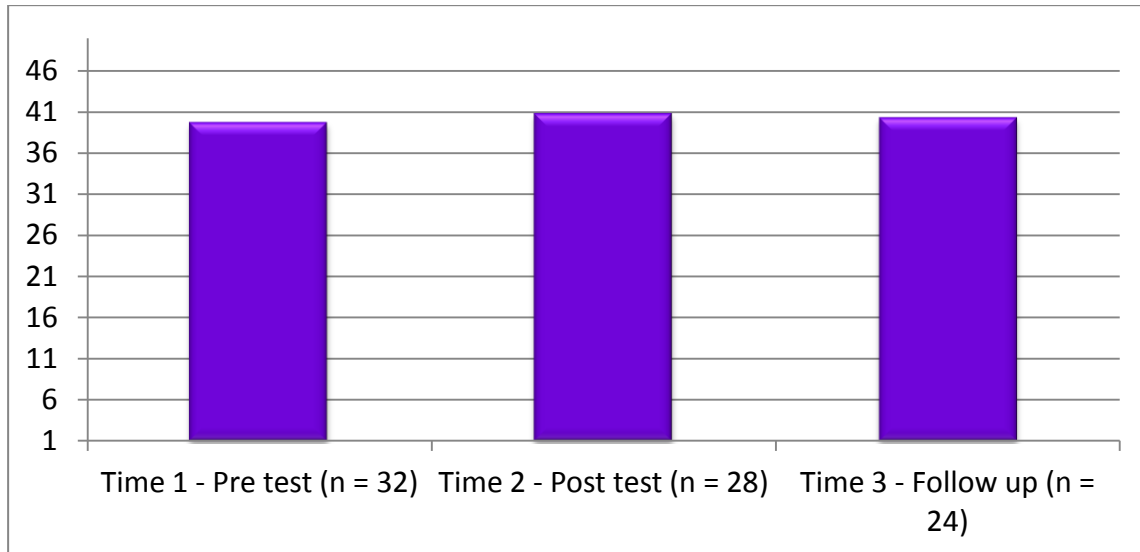


Figure 12. Mean scores on the *PPS*.

Pre-Post differences for the *PSOC* and *PPS*

In Table 8, the means, standard deviations and confidence intervals are reported for participants who completed the parenting measures at Time 1 and Time 2. Once again, cases missing values on the subscales at any stage were list-wise excluded from analysis.

Table 8

Parent Enjoyment and Confidence at T1 and T2

	M (SD)	T1		M (SD)	T2	
		95% CI for mean Lower	Upper		95% CI for mean Lower	Upper
<i>PSOC</i> Efficacy (n = 28)	28.11 (6.31)	25.66	30.56	31.61 (4.99)	29.67	33.54
<i>PSOC</i> Satisfaction (n = 29)	34.45 (9.38)	30.88	38.02	35.59 (8.01)	32.54	38.63
<i>PPS</i> (n = 25)	39.40 (5.14)	37.28	41.52	40.64 (4.21)	38.90	42.38

Analysis of variance revealed a significant increase in overall *PSOC Efficacy* scores from Time 1 to Time 2, $F(1, 27) = 9.18, p < .01$, partial $\eta^2 = .25$. No significant increase was found in *PSOC Satisfaction* scores at Time 1 and Time 2, $F(1, 28) = 1.43, p = .24$, partial $\eta^2 = .05$, or total scores on the *PPS*, $F(1, 24) = 1.53, p = .23$, partial $\eta^2 = .06$.

Parent Enjoyment and Confidence - Pre-Post and Follow up

Table 9 shows the results for the parenting measures for those participants who filled in these measures at all three time points. Asterisks indicate a statistically significant effect.

Table 9

Mean and standard deviation scores for PSOC and PPS at all three time points.

	Means (SD)		
	T1	T2	T3
PSOC Efficacy* (n = 23)	28.96 (5.24)	31.83 (5.33)	31.70 (6.01)
PSOC Satisfaction (n = 24)	35.75 (9.47)	36.29 (7.98)	37.54 (7.35)
PPS (n = 21)	39.76 (5.09)	40.81 (4.52)	39.76 (5.09)

Parents' confidence and satisfaction (PSOC)

A one-way repeated measures ANOVA revealed that *PSOC Efficacy* scores across time were approaching a significant increase, $F(2, 44) = 3.17$, $p = .052$, partial $\eta^2 = .13$. Analysis using a Bonferroni correction showed significant differences between Time 1 and 2 (mean difference = -2.87, 95% CI [-5.61, -.13], $p = .04$), but no significant difference between Time 1 and 3 (mean difference = -2.87, 95% CI [-6.39, 0.91], $p = .19$).

The Huynh-Feldt correction was used to assess changes in *PSOC Satisfaction*, as the assumption of sphericity was not met. No significant difference was found in *PSOC Satisfaction* scores across time, $F(1.67, 38.34) = 0.97$, $p = .37$, partial $\eta^2 = .04$.

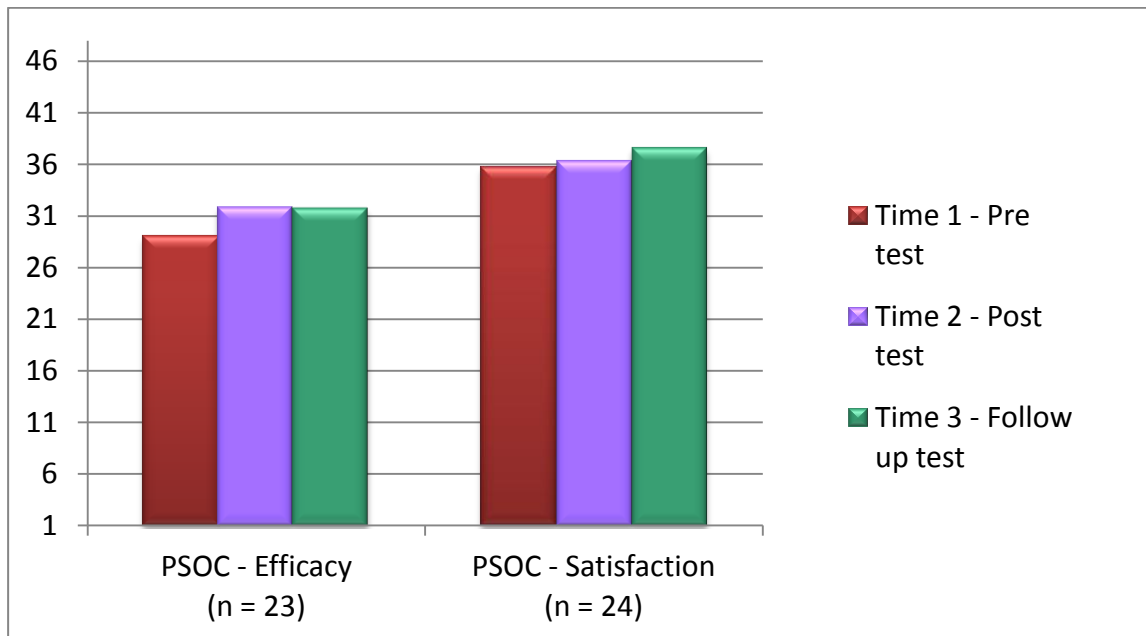


Figure 13. Mean scores of the PSOC subscales across all three time points, excluding missing data.

Pleasure in Parenting

Figure 14 presents the mean scores of 21 participants who completed the measures on three occasions. No significant change in PPS scores were observed across all three time points, $F(2, 40) = 0.56$, $p = .56$, partial $\eta^2 = .03$.

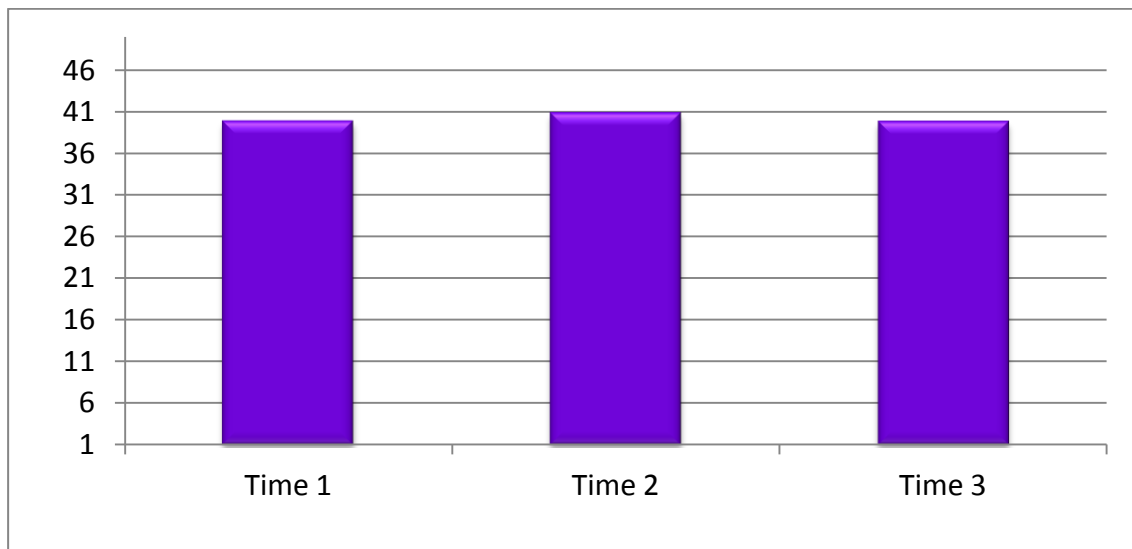


Figure 14. Means scores of the PPS across all three time points, excluding missing data (n = 21).

Child Social-Emotional Competence

Child social-emotional competence was assessed using the *ASQ:SE*. The total score of each child at each time point was categorised as either 'no risk' 'near risk' or 'at risk'. Cut off scores are provided for each time point of the *ASQ:SE* to identify which children may be at risk and in need of further assessment. The 'near risk' category was developed for this study, modelled after the original Ages and Stages Questionnaire. It indicates a score up to ten points below the cut off score for the risk category. For example, the cut off score for the 6 month *ASQ:SE* indicating an 'at risk' child is 45, therefore, near risk scores were considered to be between 35 and 44 on the 6 month *ASQ:SE*. Table 10 below shows the number of participants in each category at all three time points.

Table 10

Total number of participants in each risk category of the ASQ:SE at all three time points.

Risk Category	Pre test (n = 34)	Post test (n = 29)	Follow up (n = 25)
Normal	n = 20 (58.8%)	n = 26 (89.7%)	n = 20 (80%)
Near Risk	n = 3 (8.8%)	n = 1 (3.4%)	n = 3 (12%)
At Risk	n = 11 (32.4%)	n = 2 (6.9%)	n = 2 (8%)

Pre to post changes in ASQ:SE

A chi-square test of independence was performed to assess whether there were any changes in the risk category of the children in this study, for social-emotional problems between Time 1 and Time 2. The relationship between these variables was significant, $\chi^2 (4, N = 22) = 18.73, p < .001$.

ASQ:SE across all time points

A chi-square test of independence was once again performed to assess whether there were any changes in the risk category of the children in the study for social-emotional problems across all three time points. The relationship between these variables was once again found to be significant, $\chi^2 (4, N = 24) = 14.53, p = .006$.

Changes in the risk category of the children in the current study were also assessed for pre-test (Time 1) and follow up test (Time 3). The relationship between these variables was not significant, $\chi^2 (4, N = 24) = 1.66, p = .80$, indicating no difference in risk category between Time 1 and Time 3.

Table 11 below shows the number of participants in each category at all three time points, excluding those with missing data.

Table 11

Number of participants in each risk category of the ASQ:SE, excluding those with missing data (n = 24).

Risk Category	Pre test	Post test	Follow up
Normal	n = 19 (79.2%)	n = 22 (91.7%)	n = 19 (79.2%)
Near Risk	n = 2 (8.3%)	n = 1 (4.2%)	n = 3 (12.5%)
At Risk	n = 3 (12.5%)	n = 1 (4.2%)	n = 2 (8.3%)

Parent Professional and Social Networks

Parent professional and social networks were assessed using three different measures, the *NCAST Community Life Skills Scale (CLSS)*, the *NCAST Difficult Life Circumstances scale (DLC)* and the *NCAST Network scale*. These scales were only assessed at pre-test (Time 1) and follow up (Time 3) as the previous PlaySteps study found no differences in pre to post test in these measures, but some differences at pre-test to follow up. It was thought that this would enable an appropriate amount of time to pass for any changes in these areas to take effect.

Pre to follow up changes in the CLSS

A one-way repeated measures ANOVA was used to assess any changes in the *total CLSS* scores of participants. A significant increase was found in the *CLSS* scores of participants from pre-test (M = 25.99, SD = 3.6, 95% CI [24.34, 27.41]) to follow up test (M= 27.92, SD = 4.07, 95% CI [26.20, 29.63]), $F(1, 23) = 5.80, p = .02, \text{partial } \eta^2 = .20, n = 24$.

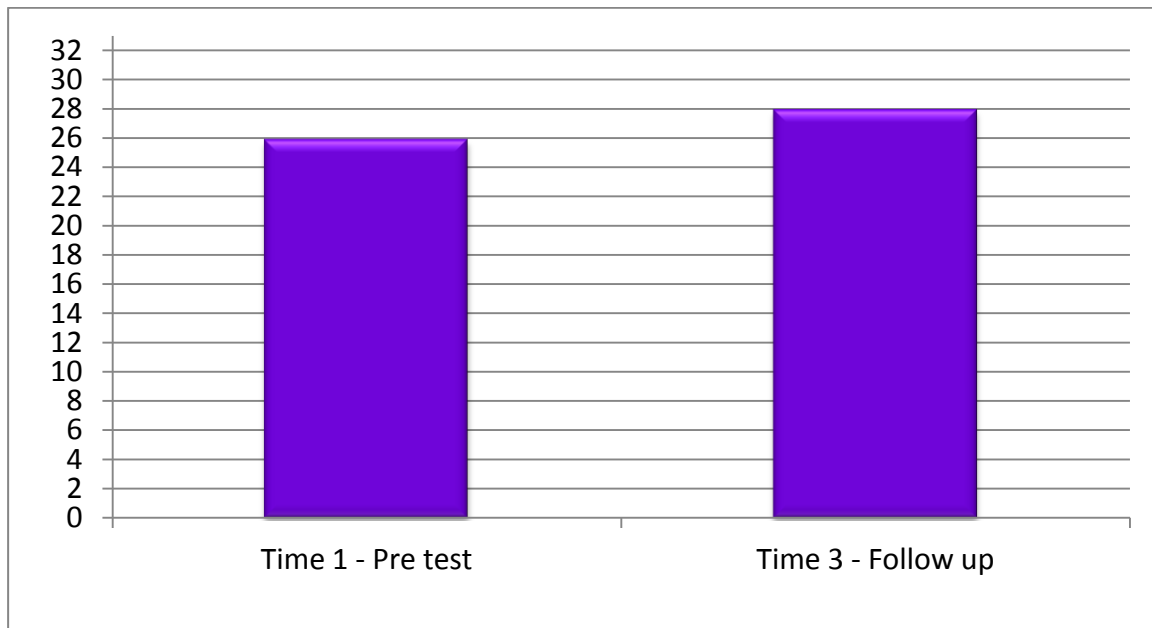


Figure 15. Pre to follow up changes in total scores of the CLSS, excluding missing data (n = 24).

CLSS Subscales. The CLSS consists of 6 subscales, use of: *transportation, budgeting, support services, support-involvement, interests-hobbies, regularity-organisation-routines*. Table 12 shows the mean and standard deviations for participants on these subscales, with an asterisks indicating a significant result.

Table 12

Mean and standard deviation scores for the subscales of CLSS (n=24).

CLSS Subscale	Means (SD)	
	Time 1	Time 3
Transport	3.75 (0.44)	3.71 (0.55)
Budgeting	4.25 (0.79)	4.38 (0.71)
Support Services	4.08 (0.83)	4.42 (0.72)
Support Involvement*	4.42 (1.50)	5.25 (1.62)
Hobbies-Interests*	2.50 (1.02)	3.04 (1.08)
Regularity-Organisation-Routines	6.88 (1.12)	7.21 (0.98)

There were no significant differences found across time for the subscales of *transportation* ($F(1, 23) = 0.19, p = .66, \text{partial } \eta^2 = .01$), *budgeting* ($F(1, 23) = 1.20, p = .27, \text{partial } \eta^2 = .05$), *support services* ($F(1, 23) = 3.54, p = .07, \text{partial } \eta^2 = .13$), or the *regularity-organisation-routines* subscale ($F(1, 23) = 2.42, p = .13, \text{partial } \eta^2 = .10$).

However, significant increases in subscales scores were found for the *support-involvement subscale* ($F(1, 23) = 4.83, p = .04, \text{partial } \eta^2 = .17$), and the *hobbies-interests subscale* ($F(1, 23) = 5.41, p = .03, \text{partial } \eta^2 = .19$), indicating parents increased their use of social contacts and their involvement in hobbies and interests outside of family life, from Time 1 to Time 3.

Pre to follow up changes in the DLC

A one-way repeated measures ANOVA was used to assess any changes in difficult life circumstances, assessed with the *DLC*. No significant differences in *DLC* scores were found from pre-test ($M = 15.17, SD = 11.21, 95\% \text{ CI } [10.54, 19.78]$) to follow up ($M = 15.75, SD = 11.33, 95\% \text{ CI } [11.07, 20.42]$), $F(1, 24) = .10, p = .75, \text{partial } \eta^2 = .00, n = 25$.

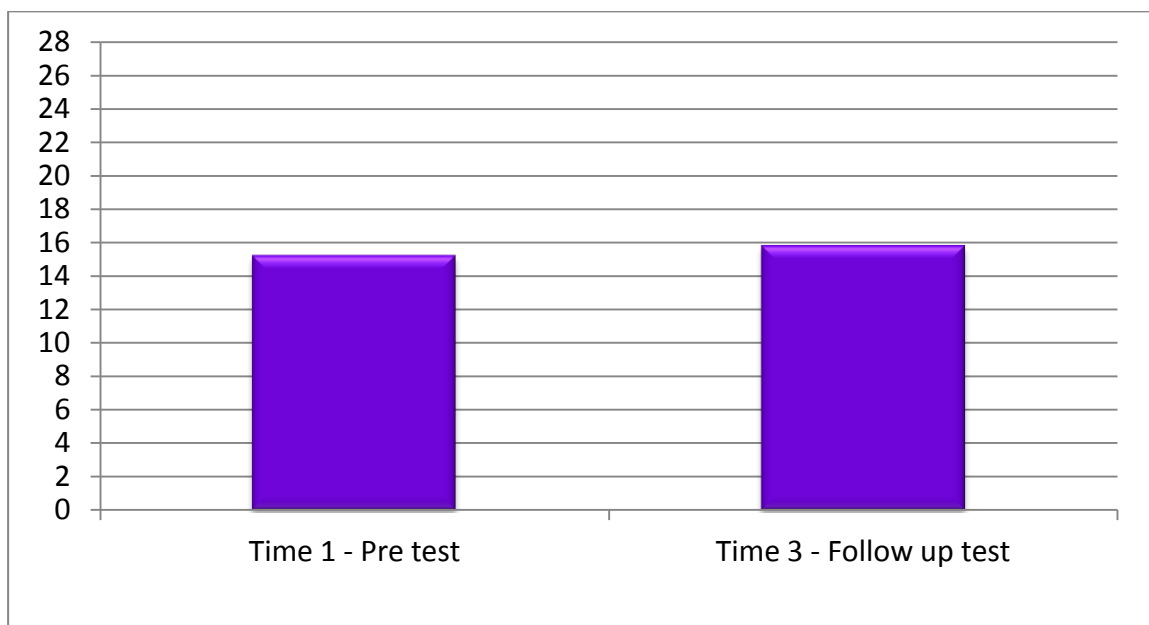


Figure 16. Pre to follow up changes in DLC scores for all participants, excluding missing data ($n = 25$).

Pre to follow up changes in the NCAST Network Scale

Changes in the amount and quality of professional and personal networks were assessed from pre-test (Time 1) to follow up (Time 3) using a one-way repeated measures ANOVA. For both personal and professional contacts, three variables were examined – the number of contacts for each participant, how helpful the participant found these contacts and how much trouble it was (or the difficulty) for participants to gain help from these contacts. The means and standard deviations for these variables are displayed in Table 13 below, an asterisk indicates a significant result.

Table 13

Means and standard deviation scores for the NCAST Network scale at T1 and T3

	Mean (SD)	
	T1	T3
Personal Contacts		
How many? (n = 23)	3.61 (1.87)	3.65 (2.39)
How helpful? (n = 23) *	1.57 (0.59)	1.83 (0.39)
How much trouble? (n = 22)	1.86 (1.28)	1.50 (1.22)
Professional Contacts		
How many? (n = 22)	2.64 (1.40)	1.95 (1.36)
How helpful? (n = 19)	1.89 (0.32)	1.89 (0.32)
How much trouble? (n = 18)	1.33 (1.28)	1.06 (1.16)

There were no significant difference found in the average *number of personal contacts* after completing the program ($F(1, 22) = .007, p = .94, \text{partial } \eta^2 = .00$). However, from pre-test to follow up a significant increase was found in the average rating of *how helpful* participants believed their personal contacts are, $F(1, 22) = 5.35, p = .03, \text{partial } \eta^2 = .20$. No difference was found in the amount of difficulty in using these contacts, $F(1, 21) = 0.91, p = .35, \text{partial } \eta^2 = .04$.

No difference was also found in the amount of *professional contacts* from pre-test to follow up ($F(1, 21) = 2.95, p = .10, \text{partial } \eta^2 = .12$), in the *helpfulness* of these contacts ($F(1, 18) = 0.00, p = 1.00, \text{partial } \eta^2 = .00$) nor in the difficulty of using these contacts ($F(1, 17) = .57, p = .46, \text{partial } \eta^2 = .03$).

Summary of findings

Improvements in parent-child interaction

There were no significant changes between Time 1 and Time 2 in parent-child interaction, as measured by the *NCAST Teaching scale*. For the participants who completed assessments before intervention and at follow up, there was a statistically significant improvement in the *Cognitive Growth Subscale* of the *NCAST Teaching Scale*, with a small effect size. This indicates that upon completion of the PlaySteps program, parent's ability to communicate reciprocally with their child to encourage learning and vocalisation increased. However, there were no significant differences found across time for any of the other subscale scores or total scores for this scale.

A visual inspection of the NCAST subscales showed an increase in mean scores from Time 1 to follow up for: *Sensitivity to Cues*, *Response to Distress*, *Responsiveness to Parent*, *Caregiver Total*, *Caregiver-Child Total*, *Child Total*, *Caregiver Contingency*, *Caregiver-Child Contingency* and *Child Contingency*. However, these changes failed to reach statistical significance.

When compared with the NCAST population means, it is worth noting that the mean subscale scores of this sample were higher than the population mean at Time 1 for the subscales: *Sensitivity to Cues*, *Clarity of Cues*, *Responsiveness to Parent*, *Child Total* scores and *Child Contingency* score. Thus, it could be speculated that there may not have been much room for improvement amongst these subscales.

Improvements in parental wellbeing

Changes in the short-term, that is, in pre-post differences, were observed for the *Anxiety* and *Stress* subscales of the *DASS*, with a large effect size for *Anxiety* and a medium effect size for *Stress*. These changes were in the expected direction, indicating a decrease in perceived anxiety and stress for participants after the completion of the PlaySteps program. Although a visual inspection of depression scores indicates a decrease in participants' perceived depression, this did not reach statistical significance.

It is also of note that mean scores for participants at Time 1 were higher than the average scores cited by the developers of the scale. This indicates that levels of distress for participants in this study were higher than would normally be found in the population.

A significant decrease in all three subscale scores (*Depression, Anxiety and Stress*) was found from Time 1 to follow up. A medium effect size was observed for *Depression* and *Stress*, while a large effect size was found for anxiety. This shows that improvements in parental wellbeing upon completion of the PlaySteps program were maintained over the follow up period.

Improvement in parental wellbeing was also shown by an examination of 'clinical' status as revealed by the *DASS* scores. At Time 1, the number of parents in the 'clinical' range was high for *Anxiety* and *Stress*, with over half the group scoring in or above the moderately severe range in these subscales. Forty-six percent of participants were found to be in the 'clinical' range for *Depression* at Time 1. Inspection of these scores post-intervention and at follow up showed a decrease in the percentage of participants in this 'clinical' range.

Overall, the results for parent wellbeing show improvements in the desired direction for *Depression, Anxiety and Stress* in the short and longer term, with strong effects for anxiety and stress.

Improvements in parental enjoyment and confidence

Parental enjoyment and confidence was measured in two ways, with the *Parenting Sense of Competence Scale* (PSOC) and the *Pleasure in Parenting Scale* (PPS). From Time 1 to Time 2, a significant increase was found for the PSOC *Efficacy* subscale, with a medium effect size, indicating that participants' perceived confidence in their parenting ability increased upon completion of the PlaySteps program. No significant differences were found in PSOC *Satisfaction* or PPS scores between Time 1 and Time2.

From Time 1 to follow up, a similar effect was found. There was a significant increase in PSOC *Efficacy* scores with a small effect size, indicating maintenance of improvement over the longer term. Although visual inspection of the *PSOC Satisfaction* scale showed an increase in scores across all three time points, these did not reach statistical significance. There was also no significant difference found in *PPS* scores for parents.

Taken together, the results for the *PSOC* and *PPS* scales indicate initial and sustained improvements in parenting efficacy, but not in parenting enjoyment.

Improvements in children's social and emotional competence

Children's social and emotional competence was assessed at all three time points via parental report with the *Ages and Stages Questionnaire: Social Emotional (ASQ:SE)*. From Time 1 to Time 2, improvements in social and emotional competence were shown upon examination of the 'risk' status of children. Over half of the child participants were found to be in the 'normal' risk range at Time 1, this increased to almost 90% at Time 2. However, there was a slight increase in the number of participants in the 'near risk' and 'at risk' categories from Time 2 to Time 3. . The number of those in the 'normal' risk range at follow up was found to be 80%.

Overall, after completion of the PlaySteps program, for some of the children in the categories of 'near risk' and 'at risk', parental reports of the social and emotional competence of their children reflected a positive change. However, for a small percentage of these children, the changes may not have been maintained in the long term.

Improvements in parent's professional and social networks

Changes in parent's professional and social networks were assessed at Time 1 and follow up using the *Community Life Skills Scales (CLSS)*, the *Difficult Life Circumstances Scale (DLC)* and the *NCAST Network scale*.

A significant increase was found in the total score of *CLSS* from pre-intervention to follow up, with a medium effect size. This indicates an increase in participants' use of community resources and their involvement and participation in community life, after their involvement in the PlaySteps program.

Specifically, when examining the differences in subscale scores, a significant increase with a medium effect were found for the *support-involvement subscale* and the *hobbies-interests subscale*, both with medium effects. No significant differences were found for the subscales of *transport, budgeting, support services, or regularity-organisation-routine*. Visual inspection showed an increase in mean scores for the budgeting, support services and the regularity-organisation-routine subscales, however, these did not reach statistical significance.

No differences were found in the *DLC* scale scores for participants from Time 1 to follow up, indicating no reported changes in challenging life events after involvement in the PlaySteps program.

Finally, the *NCAST Network* scale showed no significant difference in the number *professional contacts*, or the *perceived helpfulness* or *difficulty* in using these contacts. Visual inspection showed a decrease in the number of *professional contacts* used, and also in the *perceived difficulty* of using these professional contacts. However, these differences did not reach statistical significance. In regards to *personal contacts*, there was no significant difference found in the *number of personal contacts* or in the *perceived difficulty* of using these contacts. However, there was a statistically significant increase in *the perceived helpfulness of personal contacts* from Time 1 to follow up, with a medium effect. This indicates that after involvement in the PlaySteps program, participants perceived the personal contacts to be more helpful in assisting them with their day to day troubles. Although visual inspection showed a decrease in the difficulty of using these personal contacts, it did not reach statistical significance.

Overall, after involvement in the PlaySteps program, there was no reported change in the difficult life circumstances of caregivers who participated. The number of personal and professional contacts for participants and the difficulty in utilizing these contacts did not change. However, after involvement in the PlaySteps program, participants reported an increase in their involvement in community life, in the use of community resources and also to perceive their personal relationships as being more helpful to them.

Comparison with QEC PlaySteps evaluation 2010

The current evaluation used many of the measures that were collected in the earlier study conducted at the QEC. Below is a brief comparison of findings from common measures.

Participants

There were no apparent differences in participant relationship status and parent age between the two studies. There were minor differences in child age. In the previous study, the average age of the children was 14 months compared to 12 ½ months in the later study.

There was a difference between the two studies in source of family income. In the current study, 32.4% of parents were in paid employment compared with 54% in the 2010 evaluation.

The educational level of parent participants in the previous study was lower with 44% of participants completing year 12 education or less, compared with 27% who had completed year 12, or who had left school earlier, in the current study. No participants in the previous study had a TAFE qualification, compared with 11% with TAFE qualifications in the later study.

Parent-child interactions

Visual inspection of the data at Time 1 (pre-intervention) shows some differences between the earlier and later studies in the NCAST Teaching Scale findings. The average before-intervention scores were higher in the 2011 study than the earlier study for 4 out of 6 of the measures of parent behaviour: *Sensitivity to Cues*, *Response to Distress*, *Caregiver Total*, and *Caregiver –Child Total*. For three of these scale scores, the differences were small. However, in the 2011 study, the average score for the *Sensitivity to Cues* subscale was more than one standard deviation higher than the average score for the earlier study. Slightly higher average scores were found for the 2011 participants for *Clarity of Cues*, *Caregiver Contingency*, *Caregiver-Child Contingency*, and *Child Contingency*.

In the previous study all average scores in the parent domain were slightly below the population means, whereas in the later study, only two of the average scores were lower. Taken together, this indicates that the 2010 participant group was demonstrating marginally lower levels of optimal parent-child interactions, before engaging in PlaySteps, than were participants in the later study.

With reference to program outcomes, parents in the earlier study showed statistically significant improvements in NCAST Teaching Scale responses: *Sensitivity to Cues*, *Emotional Growth Fostering*, *Cognitive Growth Fostering*, *Caregiver Total*, *Care-giver-Child Total*, *Caregiver Contingent* and *Caregiver-Child Contingent Scores*. In the current study, only *Cognitive-Growth Fostering* showed a statistically significant difference from Time 1 to follow up.

The differences in outcomes for the two studies could be partially explained by the pre-program scores' differences. There may have been less room for improvement in the later study, given that most of the pre-intervention scores were similar to the population averages.

Parent wellbeing

Both studies showed appreciable improvement in self-reports of parent wellbeing. The main difference between the two studies is in the area of depression. In the earlier project, the largest effect size was for change in depressive symptoms from Time 1 (Pre) to Time 2 (Post). However in the 2011 study there was no statistically significant difference between Time 1 and Time 2 for *Depression* and there was a small effect size.

Parent enjoyment and confidence

The 2010 findings showed statistically significant Pre-Post differences in both *Efficacy* and *Satisfaction*, according to the *Parenting Sense of Competence Scale*. The 2011 study showed a difference for *Efficacy* only. With regard to *Pleasure in Parenting*, there was a significant Pre-Post difference for the 2010 participant group. No significant difference was found for the later study.

As with the parent-child interaction scores, the average pre-intervention scores for the parent enjoyment and confidence measures were higher in the current study than in the earlier one. The greatest difference was for the average score for *Pleasure in Parenting* in the 2011 study which was more than one standard deviation higher than the average 2010 *Pleasure in Parenting* score.

Professional and Personal Social Networks

Only the *NCAST Network Scale* was used in both studies and similar results were found. There were no differences pre and post intervention for majority of the *NCAST Network Scale*. However, both studies found improvements in helpfulness of personal contacts.

Conclusions and recommendations

Visual inspection of the demographic data from the three early parenting centres shows different patterns of characteristics across relationship status, family income, and education. This suggests that the centres may be tapping into different client groups for this program. It would be of interest to investigate the effects of the program according to the centre the participants attended and their personal characteristics. However, the small sample size in the current evaluation does not permit meaningful outcome comparisons. Future research with a larger sample size could focus on the association between participant variables and outcomes.

This project demonstrated improvements in a number of parenting measures and promising results for changes in the risk status of children's social and emotional competence. For some of the improvements found, the effect sizes were highly acceptable. Particularly strong results were found for parental mental health and sense of efficacy in the parenting role. These results were consistent with the findings of the earlier study.

It is interesting to speculate on the mechanisms that bring about changes in parental mental health and confidence; that is, to hypothesize what components of the PlaySteps program contribute to these improvements. Future research could focus on an analysis of the mechanisms for change.

As with many 'real world' evaluation projects, there are limitations to which rigorous research design can be applied. In this evaluation, the use of a quasi-experimental design with no control or comparison condition, limits the confidence with which conclusions can be made about the effects of intervention: that is, whether the intervention alone was responsible for improvements. However, the results are promising, particularly as many of the measures replicated the findings of an earlier study. Future research could possibly address the design issue with a study that incorporates a control/comparison group. That is, participants could be allocated to PlaySteps or another program that has different aims and content.

References

- Antony, M., Bieling, P.J., Cox, B.J., Enns, M.W., & Swinson, R.P. (1998). Psychometric properties of the 42 item and 21 item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176–181.
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379-384.
- Barnard, K. (1994). *NCAST Teaching Scale*. Seattle: NCAST Publications, University of Washington, School of Nursing.
- Barnard, K., Hilsinger, G., Patteson, D., Snyder, C., Solchany, J., & Shangle, M. (1999). *Parent Protective Factors Project (1995-1999), Final Report*. Seattle, WA: University of Washington.
- Brandt, P.A. (1989) *Network Survey*. Seattle, WA: NCAST, University of Washington.
- Fagot, B.I. (1995). Development of a pleasure in parenting scale. *Early Development and Parenting, 4*, 75–82.
- Farel, A. M., Freeman, V. A., Keenan, N. L., & Huber, C. J. (1991). Interaction between high-risk infants and their mothers: The NCAST as an assessment tool. *Research in Nursing & Health, 14*(2), 109-118.
- Johnston, C., & Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology, 18*, 167-175.
- Kelly, J.F. & Barnard, K.E. (1999). Parent education within a relationship-focused model, *Topics in Early Childhood Special Education, 19*(3), Fall, 151-157.
- Kelly, J. F., Buehlman, K., & Caldwell, K. (2000). Training personnel to promote quality parent-child interaction in families who are homeless. *Topics in Early Childhood Special Education, 20*, 174-185.
- LeCuyer-Maus, E. A. (2003). Stress and coping in high-risk mothers: Difficult Life Circumstances, psychiatric-mental health symptoms, education, and experiences in their families of origin. *Public Health Nursing, 20*(2), 132-145.
- Lovejoy, M.C., Verda, M.R., Hays, C.E. (1997). Convergent and discriminant validity of measures of parenting efficacy and control. *Journal of Clinical child Psychology, 26*, 366-376.
- Lovibond, P.F. (1998). Long-term stability of depression, anxiety and stress syndromes. *Journal of Abnormal Psychology, 107*(3), 520–526.

- Lovibond, P.F., & Lovibond, S.H. (1995). The structure of negative emotional states: Comparison of the Depression, Anxiety, Stress Scales with the Beck Depression and Anxiety Inventories. *Behavior Research & Therapy*, 33(3), 335–343.
- Ohan, J. L., Leung, D. W., & Johnston, C. (2000) The Parenting Sense of Competence Scale: Evidence of a Stable Factor Structure and Validity, *Canadian Journal of Behavioural Science*, 32, 251-261.
- Pridham, K. A., Lutz, K. F., Anderson, L. S., Riesch, S. K., & Becker, P. T. (2010). Furthering the understanding of parent–child Relationships: A nursing scholarship review series. Part 3: Interaction and the parent–child relationship—Assessment and intervention studies. *Journal for Specialists in Pediatric Nursing*, 15(1), 33-61.
- Squires, J., Bricker, D., & Twombly, E. (2004). Parent-completed screening for social emotional problems in young children: The effects of risk/disability status and gender on performance. *Infant Mental Health Journal*, 25(1), 62-73.
- Sumner, G., & Spietz, A. (1994). *NCAST Caregiver/Parent–Child Interaction Teaching Manual*. Seattle: NCAST Publications, University of Washington, School of Nursing.

Appendix 1

Comparison of mean scores (M) and standard deviation scores (SD) for completers and non-completers (an asterisks indicates significant findings).

<i>Measure</i>	Completers M (SD)	Non-completers M (SD)
DASS – Depression	13.80 (9.96) (n = 25)	7.63 (8.54) (n = 11)
DASS – Anxiety	10.76 (6.94) (n = 25)	8.82 (8.01) (n = 11)
DASS – Stress	17.96 (10.43) (n = 25)	12.55 (10.29) (n = 11)
Parenting Sense of Competence – Efficacy	28.32 (5.49) (n = 25)	28.63 (8.62) (n = 11)
Parenting Sense of Competence – Satisfaction	35.64 (9.28) (n = 25)	34.30 (9.29) (n = 10)
Pleasure In Parenting Scale (PPS)	39.96 (5.22) (n = 23)	38.89 (3.55) (n = 9)
Difficult Life Circumstances Scale (DLC)	15.17 (11.20) (n = 25)	17.72 (10.23) (n = 11)
Community Life Skills Scale (CLSS)	25.88 (3.64) (n = 24)	25.64 (3.38) (n = 11)
<i>NCAST Teaching Scale</i>		
Sensitivity to cues	10.12 (0.73) (n = 25)	9.50 (1.07) (n = 8)
Responsiveness to distress	9.84 (1.34) (n = 25)	9.38 (1.19) (n = 8)
Emotional growth fostering	8.92 (1.08) (n = 25)	9.00 (1.07) (n = 8)
Cognitive growth fostering	11.64 (2.78) (n = 25)	10.00 (2.51) (n = 8)
Caregiver total score	40.52 (4.39) (n = 25)	37.88 (2.59) (n = 8)
Child-caregiver total score	58.32 (5.33) (n = 25)	54.62 (3.48) (n = 8)
Child clarity of cues	8.56 (1.00) (n = 25)	8.50 (0.93) (n = 8)
Child responsiveness to parent	8.64 (2.02) (n = 25)	8.25 (2.12) (n = 8)
Child total score	17.20 (2.75) (n = 25)	16.75 (2.38) (n = 8)
Caregiver contingency score	15.20 (3.27) (n = 25)	13.00 (2.07) (n = 8)
Child contingency score	8.04 (1.97) (n = 25)	7.50 (1.69) (n = 8)
Child-caregiver contingency score*	23.24 (3.48) (n = 25)	20.50 (2.56) (n = 8)