Changing the Life Trajectories of Australia's Most Vulnerable Children

Report No. 4

24 months in the Early Years Education Program: Assessment of the impact on children and their primary caregivers

May 2019

THE UNIVERSITY OF MELBOURNE

MELBOURNE INSTITUTE
APPLIED ECONOMIC & SOCIAL RESEARCH

KIDS FIRST

24 months in the Early Years Education Program: Assessment of the impact on children and their primary caregivers

Yi-Ping Tseng	Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Brigid Jordan	Department of Paediatrics, University of Melbourne; Royal Children's Hospital; Murdoch Children's Research Institute
Jeff Borland	Department of Economics, University of Melbourne
Nichola Coombs	Department of Economics and Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Kate Cotter	
Mael Guillou	Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Alice Hill	
Anne Kennedy	Melbourne Graduate School of Education, University of Melbourne; Faculty of Arts and Education, Charles Sturt University
Jane Sheehan	Department of Economics and Melbourne Institute of Applied Economic and Social Research, University of Melbourne
Corresponding author	**Jeff Borland (jib@unimelb.edu.au)**

Acknowledgements

Our greatest debt is to the children and families who have been willing to participate in the research trial of the Early Years Education Program. We are grateful for the expertise, care and professionalism of the many staff who were involved in providing the Early Years Education Program between 2010 and 2018. The Early Years Education Program was initiated and has been substantially funded by the Children's Protection Society (now trading as Kids First). We have appreciated the opportunity to work in partnership with, and the support we have received from, the Children's Protection Society Board and its senior management. A complete list of staff from the Children's Protection Society and University of Melbourne who have been involved with delivery of the Early Years Education Program and the research trial is in Appendix 1. Funding for the trial of the Early Years Education Program has been provided by the Children's Protection Society, Commonwealth Department of Education, Employment and Workplace Relations (now Department of Education), Commonwealth Department of Families, Housing, Community Services and Indigenous Affairs (now Department of Social Services), the Victorian Department of Human Services (now Victorian Department of Health and Human Services), the Victorian Department of Education and Training, Matt and Julia Rockman, the Ian Potter Foundation, the RE Ross Trust, the Pratt Foundation, the Barr Family Foundation, the Sidney Myer Fund, VicHealth, the Antipodean Family Foundation, the Murphy-McNicol Family, the Crawford Foundation, the Brougham Family Trust, the William Buckland Foundation, and Australian Research Council Linkage Grant LP140100897. The study uses unit-record data and parts of questionnaires from Growing Up in Australia: Longitudinal Survey of Australian Children (LSAC). The questionnaires are the property of the Commonwealth Department of Social Services (DSS). LSAC is being undertaken in partnership between the Commonwealth Department of Social Services, Australian Bureau of Statistics (ABS) and Australian Institute of Family Studies (AIFS). The findings and views reported in this paper are those of the authors and should not be attributed to DSS, ABS or AIFS.

Contents

Executive summary

Overview of main findings

This report presents findings on the impact on children and their primary caregivers of 24 months of enrolment in the Early Years Education Program (EYEP). After 24 months the impact of attending EYEP on children and their families is broad and powerful. Large positive impacts of EYEP are found on children's cognitive and non-cognitive development – primarily IQ, protective factors related to resilience and social-emotional development. There is also some evidence that EYEP improves children's language skills and lowers the psychological distress of their primary caregivers.

EYEP and the research trial

EYEP is a centre-based, early years care and education program targeted at the needs of children who are exposed to significant family stress and social disadvantage, including being at heightened risk of, or having experienced, abuse and neglect. The program has a dual focus: to address the consequences of family stress on children's development and to redress learning deficiencies. The ultimate objective of EYEP is to ensure that at-risk and vulnerable children realise their full potential, and arrive at school developmentally equal to their peers and equipped to be successful learners.

Children who participate in EYEP are offered three years of care and education (50 weeks per year, five hours per day each week). Key features of EYEP are high staff/child ratios, qualified and experienced staff, inclusion of an infant mental health consultant as a member of the staff, and a rigorously developed curriculum.

The impact of EYEP is being evaluated through a Randomised Controlled Trial (RCT) as part of the Early Years Education Research Program (EYERP); otherwise referred to in this report as the 'EYEP trial'. Children for whom consent was given to participate in the EYEP trial were randomly assigned into either an intervention group who were enrolled in EYEP or to a control group. Estimates of the impact of EYEP on children and their primary caregivers are derived from comparisons of outcomes between the groups.

Characteristics of children and primary caregivers in the EYEP trial

To be eligible for the EYEP trial, children had to be aged less than 36 months at the time of entry to the trial, assessed as having two or more risk factors as defined in the Department of Human Services 2007 Best Interest Case Practice Model, be currently engaged with family services or child protection services and have early education as part of their care plan.

The eligibility criteria enabled the selection of a group of participants in the EYEP trial for whom the program was designed – children with substantial developmental delay living in families experiencing high levels of stress (Tseng et al., 2017). Compared with a general population of children, participants in the EYEP trial are highly disadvantaged on a variety of dimensions. EYEP trial participants had lower birth weight and, at the time of entry to the trial, had compromised development of IQ, language skills, motor skills and adaptive behaviour. The primary caregivers of children in the EYEP trial have lower levels of labour force engagement and family income, are more likely to experience stressful events and have higher levels of psychological distress than the general population of caregivers.

Outcomes

The main objective of the EYEP trial is to test whether the program is meeting its goal to improve children's cognitive and non-cognitive skills (Jordan et al., 2014, p. 3). Outcomes relating to children's cognitive skills (IQ and language skills) and their non-cognitive skills (within-child protective factors related to resilience and social-emotional development) are therefore examined. Possible impacts of EYEP on primary caregivers are evaluated using outcomes

relating to perceptions of the level of stress they are experiencing and the quality of home environment.

Sample size

A total of 145 children were recruited to the EYEP trial. These children come from 99 families. There are 72 children who are in the intervention group and 73 in the control group, and respectively 50 and 49 families in those groups. There are 64 girls and 81 boys in the trial.

Over time, there has been attrition from the group of 145 children who were originally recruited to the EYEP trial. The data collection at 24 months, on which the analysis in this report primarily relies, obtained responses from 104 children and their primary caregivers. In addition, it was not possible to collect a complete set of data on outcomes for all these 104 children and primary caregivers. Hence, the analysis of the impact of EYEP on outcomes at 24 months is based on samples that usually consist of 85 to 95 children or primary caregivers.

Empirical method

Initial random assignment of participants in the EYEP trial achieved balance in the characteristics of the intervention and control groups. However, due to subsequent attrition of participants from the trial, by the time data were collected on outcomes at 24 months, several characteristics were no longer balanced. Some characteristics which are unbalanced are likely to affect children's outcomes at 24 months. Hence, it is not possible to estimate the impact of EYEP simply by comparing the average outcomes for children in the intervention and control groups. Instead, for the estimated impact of EYEP to reflect only the effect of participation in the program, it is necessary to use empirical methods that can correct for differences in characteristics between the groups. For this study we have applied two alternative methods to estimate the impact of EYEP – a regression method and a propensity score matching method.

What is the 'impact' of EYEP that is estimated in this study?

The impact of EYEP is estimated for children who attended the program for at least 60 days over 24 months (out of a possible total of approximately 480 days). We restrict our attention to this sample in order that all children in the intervention group have had a level of exposure to EYEP that could conceivably be expected to affect their development. This is the same approach as was taken in estimating the impact of EYEP after twelve months (Tseng et al., 2018).

The impact of EYEP on outcomes for children after 24 months is likely to derive from two sources: first, children in the intervention group receiving a larger number of hours of early years care and education services than the control group (for example, over the first 24 months, children enrolled in EYEP received an average of 20.4 hours per week of formal early years care and education compared with 15.7 hours per week for children in the control group); and second, differences between the design and attributes of EYEP and the services received by the control group.

Main findings

Large and statistically significant EYEP impacts are found at 24 months for several outcomes for children: (i) IQ; (ii) protective factors related to resilience; and (iii) social-emotional development. There is also a relatively large impact on children's language skills, but this estimate has limited statistical significance.

The estimated impact on IQ is one-third to one-half of a standard deviation. This compares with average impacts on IQ from early years demonstration programs in the United States of about one-quarter of a standard deviation. The estimated impact on within-child protective factors related to resilience is about one-third of a standard deviation. The proportion of children enrolled in EYEP who are classified in the clinical range for social-emotional development is lower by 30 percentage points compared with the control group, a substantial impact.

Evidence that EYEP is having an impact on the stress experienced by primary caregivers is emerging at 24 months. Primary caregivers of children in EYEP show a reduced level of distress on the Kessler Psychological Distress K6 Scale (K6) of about 1.5 points (on a zero to 30 points scale), which is marginally statistically significant. Participation in EYEP is also estimated to be associated with small decreases in the frequency (one point on zero to 80 points scale) and in intensity (three points on zero to 100 points scale) of parenting daily hassles, but these estimates are not statistically significant. The estimated

impact of EYEP on the home environment is small and not statistically significant.

Major differences are apparent in the impact of EYEP on boys and girls – especially for non-cognitive skills. For children's IQ and language skills the estimated impacts are larger and have higher levels of statistical significance for boys than girls. For protective factors related to resilience a large and highly significant impact is found for boys, compared to a zero impact for girls. By contrast, the estimated impact of EYEP on social-emotional development exhibits the opposite pattern, with a much larger impact for girls than boys. Impacts on outcomes for primary caregivers at 24 months are confined to families with girls enrolled in EYEP. For the primary caregivers of girls there is a decrease in psychological distress and in the frequency of parenting hassles. There is, however, no evidence of an impact on these outcomes for the primary caregivers of boys.

The timing of the impact of EYEP over the first 24 months has varied across outcomes. The impact of EYEP on children's IQ appears to have been concentrated in the initial twelve months of the program, as the estimated impact size does not change appreciably between twelve months and 24 months. This result is consistent with evidence from previous trials of early years demonstration programs in the United States. Other outcomes for children show a more pronounced impact from EYEP after the second year of being enrolled in the program. The estimated impact on children's language skills increases from zero after twelve months to about three to four points after 24 months. For protective factors related to resilience the estimated impact size after 24 months is two to three times larger than after twelve months. While distinguishing the exact timing of impact on social-emotional development is difficult, there does seem to have been a positive impact spread across both the first and second years of the program. Where there have been positive outcomes for primary caregivers, these have been concentrated in the second year of their child's enrolment in EYEP.

1. Introduction

This report presents findings on the impact on children and their primary caregivers after 24 months of enrolment in the Early Years Education Program (EYEP). It follows the initial evaluation of the impact of EYEP after twelve months presented in Tseng et al. (2018).

EYEP is a model of early years care and education targeted at the needs of children who are exposed to significant family stress and social disadvantage. The impact of EYEP is being evaluated through a Randomised Controlled Trial (RCT) as part of the Early Years Education Research Program (EYERP); otherwise referred to in this report as the 'EYEP trial' (Jordan et al., 2014).

EYEP was initiated by the Children's Protection Society (CPS) (now trading as Kids First), an independent not-for-profit child welfare organisation based in the north-east of Melbourne. The program was designed and implemented by CPS in collaboration with Associate Professor Brigid Jordan and Dr Anne Kennedy.

The EYEP trial is being undertaken by a consortium of researchers (who are authors of this report) with support from their institutions and in partnership with CPS. Funding for the research trial has come from CPS, government departments at the Commonwealth and State levels, philanthropic organisations, individual donors, and the Australian Research Council.

Section 2 describes the motivation for the EYEP trial. Section 3 presents an overview of EYEP. Section 4 provides background information about the EYEP trial, and details on the characteristics of children and their primary caregivers who are participants in the trial. Section 5 introduces the outcome variables that are examined in this report. Sections 6 and 7 present preliminary information relevant to interpreting the impact of EYEP. Section 8 describes the empirical methods used to estimate the impact of EYEP. Section 9 presents and discusses the main findings on the effect of enrolment in EYEP after 24 months.

2. Background

Children's experiences in the years immediately after birth are a major determinant of their lifetime circumstances and well-being. Early life experiences have a fundamental influence on brain architecture, gene expression, and physiology. Critical aspects of children's early experiences are the interactions they have with the people around them and the degree of stress they live with. Having relationships with adults that are 'reciprocal and dynamic' and a lack of excessive stress are regarded as essential to healthy development (Center on the Developing Child at Harvard University, 2016b, pp. 7–8).

The impact of the early years is especially pronounced for children who experience neglect, abuse and toxic stress. Prolonged exposure to physical, emotional and/or sexual abuse and traumatic experiences early in life have been established to cause profound long-term adverse effects on brain and physiological development.

The impacts of trauma and abuse on brain development include negative effects on self-regulation capacities and the ability to cope with stress (Perry, 2002; Evans et al., 2008; Shonkoff, 2012; Center on the Developing Child at Harvard University, 2016a, pp. 7–12).

Disruption to brain development in turn affects the ability to learn, with recent studies, for example, showing that self-regulation is linked to the development of literacy and numeracy skills (Koenen et al., 2003; Raver et al., 2011). When children fall behind in their development of cognitive and social skills early in life, this disadvantage can become entrenched in later years. By missing out at an early age, children may lack the necessary building blocks and foundation for subsequent learning (Cunha and Heckman, 2007; Heckman, 2008; Heckman and Mosso, 2014; and for an overview see Tough, 2016, pp. 48–52). Deficiencies in cognitive and social skills before the age of five therefore can persist into later life, and become the basis of problems such as low education attainment, unemployment, teenage pregnancy, and involvement in crime (Knudsen et al., 2006; Caspi et al., 2016).

Early adversity has also been linked to physiological disruptions such as alterations in immune function (for example, Bierhaus et al., 2003; Currie and Spatz-Widom, 2010; Nicholson et al., 2012); to an increased risk of lifelong physical and mental health problems, including major depression, heart disease and diabetes (Center on the Developing Child, 2016b, p. 12; Campbell et al., 2014; Hughes et al., 2017); and to a variety of health-threatening behaviours in adolescence and adulthood (for example, Rothman et al., 2008; Ford et al., 2011; Caspi et al., 2017).

Children who experience adversity and abuse can also have a negative impact on their peers. For example, exposure in primary school to a classmate who has experienced family violence has been found to cause large negative impacts on contemporaneous and long-term outcomes including behaviour, academic achievement, university attendance and earnings (Carrell and Hoekstra, 2010 and Carrell et al., 2018).

Addressing the problem of inequality in skill development for children who are exposed to significant family stress is widely agreed to require a different type of education and care than is available from mainstream early childhood services. In a review article in *Science* the renowned educationalist Jack Shonkoff (2011, p. 982) argued that whereas existing programs for children from disadvantaged backgrounds mainly focus on providing enriched learning experiences for children and parenting education for mothers, a better approach for redressing inequalities in skill development would come from 'linking high-quality pedagogy to interventions that prevent, reduce, or mitigate the disruptive effects of toxic stress on the developing brain.'

Having a model that addresses the developmental delay of at-risk children is a critical policy issue in Australia. First, the size of the at-risk population of children in Australia is substantial. It has been estimated, for example, that in 2016–17 there were 53,277 pre-school children receiving child protection services (Australian Institute of Health and Welfare, 2018, table S3). Second, at-risk children in Australia currently seem to be the group least likely to be

able to access early years care and education (Biddle et al., 2017). Third, while evidence from trials of demonstration programs such as Perry Preschool and Abecedarian provide insights into the potential impact of early years programs, they were undertaken in the United States, and the populations covered were largely African–American living in small cities in the 1960s (Schweinhart et al., 2005; Campbell and Ramey, 1994). The relevance of this existing evidence to Australia is uncertain – causing, for example, the Productivity Commission to argue (2014, p. 155): '…it is unclear whether or not such programs would generate as significant benefits in a different cultural context and where the general quality of ECEC services and schooling is different from that of the United States'. Australian policy-makers are therefore seeking evidence which is both current and derived from practice in Australia.

This set of considerations motivated CPS to create and trial a new early years program, EYEP. CPS brought together a multi-disciplinary team of researchers in 2009 to undertake the EYEP trial. A pilot was conducted in 2010 to refine the service model, the survey and measurement methods, and the research process. Enrolment of children into the EYEP trial commenced in early 2011 and concluded in early 2016. Provision of EYEP to children in the intervention group was completed at the end of 2018. The EYEP trial is approved by the University of Melbourne Human Research Ethics Committee (HREC 1034236). At the time it commenced, the EYEP trial was the first RCT of a centre-based early years care and education intervention in Australia (Tapper and Phillimore, 2012).

3. The Early Years Education Program

EYEP is an innovative Australian inter-disciplinary centre-based early years care and education program. It is designed to meet the educational and developmental needs of infants and young children who are living with significant family stress and social disadvantage, including being at heightened risk of, or having experienced, abuse and neglect. The program has a dual focus: first, addressing the consequences of significant family stress on children's brain development and emotional and behavioural regulation; and second, redressing learning deficiencies. It involves direct intervention with a child to address his or her identified needs, reverse developmental delays, and reduce the impact of risk factors and adverse events. The program seeks to build children's cognitive and non-cognitive skills recognising the critical role that both types of skills play in subsequent development and lifetime outcomes (for example, Kautz et al., 2014). The ultimate objective of EYEP is to ensure that at-risk and vulnerable children can enter their first year of formal schooling developmentally equal to their peers and with the knowledge, skills and attributes needed for ongoing successful learning.

The foundation of EYEP is a holistic model of care and education that draws on the knowledge and skill base from the field of infant mental health – including neuroscience, developmental psychology, attachment theory and findings from studies of the impact of emotional trauma on young children. A full description of EYEP and the underpinnings of its design is presented in Jordan and Kennedy (2019).

Children who participate in EYEP are offered three years of care and education (50 weeks per year and five hours per day each week from Monday to Friday). Key features of EYEP are high staff/child ratios (1:3 for children under three years, and 1:6 for children over three years), qualified and experienced staff, a rigorously developed curriculum, and the use of relationship-based pedagogy.

The basis for **care** in EYEP is an attachment-focused, trauma-informed, primary-care model which recognises the significance of respectful and responsive relationships for every child's learning and development. The purpose of the primary care model is to encourage the fostering of supplementary significant and secure attachment relationships for children who are likely to be experiencing disrupted and compromised attachment relationships in their home environments.

The **education** model in EYEP is a pedagogically-driven reflective teaching model that is child-focused and designed to align with the *National Early Years Learning Framework* of 'Belonging, Being and Becoming' (DEEWR, 2009) and the National Quality Standard (ACECQA, 2011). Each child is provided with individual learning goals developed in partnership with families. Educators plan a curriculum using play-based approaches and intentional teaching to support each child's learning and development across outcomes in the *Early Years Learning Framework*.

An innovative feature of EYEP is a multi-disciplinary model with an in-house infant mental health consultant as an integral team member, and family support and early childhood curriculum consultants. The infant mental health consultant conducts an assessment with each child as they commence in EYEP and this understanding of the individual child's emotional functioning, behavioural regulation and the parent-child attachment relationship contributes to the individualised learning plan and the relational pedagogical strategies developed for the child. Emphasis is placed on supporting children at points of transition – such as when they arrive at and depart from the centre each day, move into a new room at the centre (based on their age), or commence at and leave EYEP. Ensuring children have adequate nutrition while they are at the centre is also a key element of EYEP. An in-house qualified cook provides 75 per cent of children's daily nutritional needs via a healthy eating policy.

The EYEP model requires that only full-time educators are employed. This is intended to allow the educator to develop a strong and consistent relationship with children for whom they are responsible. New educators receive introductory professional development based on attachment theory and key infant mental health concepts.

Each educator receives weekly, formal, and individual reflective supervision from a member of the EYEP leadership team. A critical part of EYEP is for educators to have eight to ten hours out of the classroom each week to undertake activities such as curriculum planning, reflective supervision, peer consultations, and liaison with families.

The EYEP model actively engages with parents to encourage their continued participation in the program, as well as to enhance their usage of all health, educational and social services available in the community that could improve outcomes for their children. Although EYEP is a child-focused intervention, not intended to directly affect parenting behaviours, development of a sustained partnership with parents is a core principle. The orientation and attendance plan for a child enables the primary caregiver to gradually build a trusting relationship with the educator of their child. Meetings between parents and family support/child protection workers and the early years educators (primary care worker for the child) take place every twelve weeks.

EYEP also addresses a variety of barriers that might otherwise exist for families taking advantage of support services – such as affordability, where families' beliefs place low priority on early education services, and inter-personal barriers including attitudes on the part of service providers that might compromise engagement (Centre for Community Child Health, 2011; see also Turnbull et al., 2000).

Meetings on 'Program Logic' were held quarterly or bi-monthly throughout the EYEP trial. These meetings were to ensure that EYEP was being implemented with fidelity and to address unanticipated issues, as well as to provide support for the EYEP leadership team. The Program Logic meetings were attended by the designers of EYEP (Brigid Jordan and Anne Kennedy), the EYEP Manager, EYEP Coordinator and Pedagogical Leader, the infant mental health consultant, and the CPS executive staff member to whom the EYEP Manager reported at the time.

4. The EYEP trial and participants

Eligibility criteria and characteristics of EYEP trial participants

Criteria for eligibility for the EYEP trial were chosen with the aim of evaluating its impact on children exposed to significant family stress and social disadvantage. Children were required to be aged less than 36 months, assessed as having two or more risk factors as defined in the Victorian Department of Human Services 2007 *Best Interest Case Practice Model*, and be currently engaged with family services or child protection services and have early education as part of their care plan. The list of risk factors consists of 24 'Child and family risk factors' and nine 'Parent risk factors'. Risk factors include having teenage parents, parental substance abuse, parental mental health difficulties, and the presence of family violence. A full list of the risk factors is included as Appendix 2. Referrals of potential EYEP participants were made by caseworkers from clients of child welfare services including (but not exclusively from) Child FIRST and Child Protection within the Victorian Department of Health and Human Services.

Information on the children for whom consent was given to participate in the EYEP trial is presented in Table 1. There are 145 children who were recruited into the EYEP trial when aged less than 36 months. There are 64 girls and 81 boys, and the children come from 99 families.

Table 1: Key descriptive information on children in the EYEP trial

	Number	Per cent
Children – By group		
EYEP	72	49.7
Control	73	50.3
Families – By group		
EYEP	50	50.5
Control	49	49.5
Children – By gender		
Female	64	44.1
Male	81	55.9

**Note:* In the initial report on the EYEP trial (Tseng et al., 2017) it was incorrectly stated that 97 families were included in the trial. This error did not affect any other information presented in that report.

In an earlier report, we presented a detailed overview of the main characteristics of children in the EYEP trial and their primary caregivers (Tseng et al., 2017). That report confirms that the eligibility criteria achieved the selection of a group of participants in the EYEP trial for whom the program was designed – children with substantial delays in development living in families experiencing high levels of stress. This was evident in several ways.

First, at the time of entry to the EYEP trial, most children had many more than the minimum number of two risk factors. About 30 per cent of children had two or three risk factors, 35 per cent had four or five risk factors, and 35 per cent had six to nine risk factors. The most frequent 'Child and family risk factors' for participants were 'attachment/relationship issues', 'mental health issues', and 'family violence, current or past'; and the most frequent 'Parent risk factor' was 'harsh, inconsistent discipline, neglect or abuse'. The existence of multiple risk factors for children in the EYEP trial is noteworthy – being consistent with evidence that it is this feature which primarily identifies children who are living in environments likely to adversely affect their long-term development (Fergusson and Horwood, 2003, p. 130; Hughes et al., 2017).

Second, at their time of entry to the EYEP trial, the children had relatively low birth weights, even compared to children of the same age living in the bottom quartile of households in Australia ranked by socio-economic status (SES). They also exhibited compromised development in the areas of IQ, language skills, motor skills, and adaptive behaviour. This can be seen in Table 2 (Panel A) which presents summary information on the birth weights and development of children in the EYEP trial.

Third, primary caregivers for children in the EYEP trial are more likely to be young parents, have fewer financial resources, and not be participating in the labour force. The number of stressful life events beyond their control at

Table 2: Characteristics of children in the EYEP trial and their primary caregivers

Panel A: Children in the EYEP trial

	EYEP	LSAC – Low SES households	General population
Very low birth weight (Less than 1500g) (per cent)	6.0	0.9	
Average score at time of entry to EYEP trial:			
Cognitive development	92.3		100
Language	87.7		100
Motor skills	88.8		100
Social and emotional development	99.5		100
Adaptive behaviour	88.8		100

Panel B: Primary caregivers of children in the EYEP trial

	EYEP	LSAC – Low SES households
Severe psychological stress (K6 equal to 19 or greater) (per cent)	25.8	4.4
Had a major financial crisis - Past 12 months (per cent)	32	18.8
Had problems with the police and a court appearance – Past 12 months (per cent)	15.3	4.0
Labour force status: Unemployed and not in the labour force (per cent)	89.0	70.7
Disposable family income: Per cent less than $250 per week ($ 2016 qtr. 1)	27.4	12.9

the time of entry to the trial was extraordinarily high. Many primary caregivers for children in the EYEP trial had severe levels of psychological distress. Summary information on primary caregivers of children in the EYEP trial is shown in Table 2 (Panel B).

The randomised controlled trial

Families with children who were eligible and consented to participate in the EYEP trial were randomly assigned into either an intervention group enrolled in EYEP or to a control group. There are 72 children in the intervention group and 73 in the control group, and respectively 50 and 49 families in these groups. In families with multiple children participating in the trial, all those children were assigned to either the intervention group or control group.

The intervention group remained enrolled in EYEP for three years, or until school entry if that time was reached prior to completion of the three years. At the time of consent to participate in the trial children were required to be young enough to be able to attend EYEP for three years before reaching school commencement age. For some children, however, factors such as delay in commencing attendance at EYEP meant that school entry occurred without them completing three years of attendance at EYEP.

The control group received 'usual care', a mix of parental and guardian care as well as care and education provided by other childcare centres or kindergartens. The usual care was determined by the choice of the child's primary caregiver(s) without any direction from the EYEP research trial.

Children in the control group are not enrolled in EYEP, however it is still possible that their outcomes may be affected by participating in the trial. For example, there was an ethical obligation on the researchers to report to primary caregivers of children in the control group on any specialist assistance believed necessary for those children. Hence, outcomes for the control group may be better than if they had not been involved in the trial. In future research, we plan to investigate this issue by comparing outcomes for the control group against a matched sample from the LSAC.

Data collection

Data are being collected on an extensive set of outcome measures for participants in the EYEP trial at five points in time: at entry to the trial, at yearly intervals for three years after entry to the trial (at 12, 24 and 36 months), and six months after beginning the first year of school. Data collection at the time of entry to the trial encompassed two stages. First, for all children for whom consent was given to participate in the EYEP trial, data on risk factors for eligibility and a small set of demographic characteristics were collected. Second, at a subsequent appointment, detailed data were collected on child development and on demographic and other characteristics of children and their primary caregivers. Analysis in this report is based on data from the time of entry to the trial and at twelve and 24 months after entry to the trial.

5. Overview of outcome variables

The purpose of the EYEP trial is to test whether the program achieves the goals of improving children's cognitive and non-cognitive development (Jordan et al., 2014, p. 3). Given these multiple goals, seeking to represent the effect of EYEP through any single outcome measure would miss much of its intended impact. Hence, we investigate a range of outcomes relating to children's development (Craig et al., 2008). In addition, the impact of EYEP on primary caregivers is evaluated using outcomes relating to perceptions of the level of stress they are experiencing; and the impact on the quality of children's home environments is investigated.

A list of the six outcomes and associated measures is presented in Table 3, and a brief description of each measure is provided below. In selecting measures our general approach has been to choose a single measure for each outcome. This is done to minimise the relatedness of the measures reported and the scope for 'cherry-picking' findings. The set of outcomes and measures reported on at 24 months is the same as at twelve months. More details on the measures are provided in Tseng et al. (2018, Appendix 3).

- **Child development – IQ and language skills:** These aspects of child development are measured using standardized tests: the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley 2006); and the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI) (Wechsler, 2002). These are the most widely applied measures of the development of infants and toddlers in clinical and research settings. They capture both fluid intelligence (the rate of learning) and crystallized intelligence (acquired knowledge) (for more detail on these concepts, see Kautz et al., 2014, p. 7).

Our analysis uses the Bayley Scales for children aged up to 42 months, and WPPSI for children aged 43 months and above. Age-adjusted composite scores can be calculated for the IQ and Language domains of development for both measures. Both measures are scaled with a mean of 100 and standard deviation (SD) of 15. A score of 100 defines the average performance of a given age group, and scores of 85 and 115 are one standard deviation below and above the mean respectively. A score between 70 and 85 is defined to identify a delay in child development, and a score below 70 a significant delay in development.

Since the Bayley Scales and the WPPSI are scaled equivalently against population norms, in our analysis we simply integrate the scores from these measures. This means that if a child was assessed using the Bayley Scales at the time of entry to the trial and WPPSI at twelve months, the scores from each test are treated as being directly comparable.

Table 3: Outcomes and measures of the impact of EYEP

	Outcome	Measure
1	Child development - IQ	Bayley Scales of Infant and Toddler Development III (BSID); Wechsler Preschool and Primary Scale of Intelligence (WPPSI)
2	Child development – Language skills	Bayley Scales of Infant and Toddler Development III (BSID); Weschler Preschool and Primary Scale of Intelligence (WPPSI) – Verbal IQ score
3	Child development – Protective factors related to resilience (initiative, self-regulation, attachment/ relationships, behavioural concerns)	Devereux Early Childhood Assessment Program (DECA)
4	Child social-emotional development	Brief Infant Toddler Social Emotional Assessment (BITSEA); Child Behaviour Checklist (CBCL)
5	Parent psychological distress	K6; The Parenting Daily Hassles Scale
6	Home environment	Home Observation and Measurement of Environment (HOME)

- **Child development – Within-child protective factors related to resilience:**
  This aspect of development is measured by the Devereux Early Childhood Assessment (DECA) (Mackrain et al., 2007; LeBuffe and Naglieri, 2012). It is a parent response measure.

DECA-I is used to assess infants aged from one month to less than 18 months, DECA-T is used for toddlers from ages 18 months to less than 36 months, and DECA-P2 is used for children aged three to five years. Reponses from each instrument on items relating to children's attachment/ relationships, initiative, and self-regulation are integrated into a Total Protective Factors Scale. This Scale is reported as age normalised T scores and percentile rankings against a norm population. The T score has mean of 50 and SD of 10, and ranges from 28 to 72. A score of 40 or below is defined as signifying an area of need.

- **Child social-emotional development:**
  These aspects of child development are measured using the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) (Briggs-Gowan and Carter, 2006); and the Child Behavior Checklist (CBCL) (Achenbach and Rescorla, 2000). Both are parent response measures. We have used BITSEA for children up to 35 months and used the CBCL for children three years and older.

The BITSEA Parent Response Form is a tool for identifying children aged less than 36 months who may have social-emotional and behavioural problems and/or delays, or deficits in social-emotional competence. In this report, we focus on the instrument for identifying socio-emotional and behavioural problems. The problem score from BITSEA ranges from 0 to 62. A percentile ranking based on age-based population norms can be assigned to each problem score.

The CBCL is a parent response index of behavioural, social, and emotional functioning intended for children from 18 months up to five years. The total score on the CBCL ranges from 0 to 200. A percentile ranking based on age-based population norms can be assigned to each score (although scores below the 50th percentile are aggregated).

The BITSEA and CBCL instruments are integrated to obtain a consistent measure of problems with child emotional and social development by using as the outcome measure from each instrument the proportion of children classified as having development problems in the clinical range; that is, with a score below the population norm age-based 10th percentile cut-off.

- **Parent psychological distress:**
  Parent stress is measured using the Kessler Psychological Distress K6 Scale (K6) (Kessler et al., 2002); and the Parenting Daily Hassles Scale (Crnic and Greenberg, 1990).

The K6 scale is a widely used measure of psychological distress, including in the 1997 Australian National Survey of Mental Health and Wellbeing (Furukawa et al., 2003). The scale has six questions about feelings over the last four weeks. A K6 score is derived from summing the responses of the primary caregiver to these questions. The score can range from 6 to 30, with individuals scoring 6 to 13 being classified as exhibiting 'low' psychological distress, 14 to 18 classified as 'medium' psychological distress, and 19 to 30 classified as 'severe' psychological distress.

The Parenting Daily Hassles Scale aims to assess the frequency and intensity/impact of 20 experiences that can be a 'hassle' to parents. The frequency score can range from 0 to 80 and the intensity score from 0 to 100. Scores above (respectively) 50 and 70 are considered to show high frequency and significant intensity of pressure on parents.

- **Home environment:**
  Home environment is assessed using the Home Observation and Measurement of Environment (HOME) (Caldwell and Bradley, 2003).

HOME is a home-based rating of the home environment by an assessor/observer. It is designed to achieve systematic measurement of the environment based on observation of interaction between the primary caregiver and their child, and interview data on significant aspects of the family's interpersonal and physical environment. The Infant-Toddler instrument is used for children aged up to three years; and the Early Childhood instrument for children aged above three years. For our report, we rescale the scores from the instruments so that both have scales from 0 to 100. Higher scores signify a higher rated home environment.

6. Description of data and analysis of randomisation and attrition

Background

Analysis of the impact of enrolment in EYEP on outcomes for children and their primary caregivers for this report has drawn primarily on data collected after 24 months in the EYEP trial. As well, data collected at the time of entry to the trial are used to control for the potential impact of sample attrition; and data on outcomes at twelve months are used when investigating the timing of the impact of EYEP over the first 24 months of enrolment.

Data on the risk factors for eligibility and basic demographic characteristics are available for all 145 children for whom consent to participate in the trial was given. More detailed data on child development and demographic characteristics was collected prior to commencement in EYEP for 134 children. By the time of the data collection at 24 months further drop-out resulted in data being available for a maximum of 104 children and their primary caregivers. Details of the evolution of the maximum sample size for the intervention and control groups are shown in Figure 1.

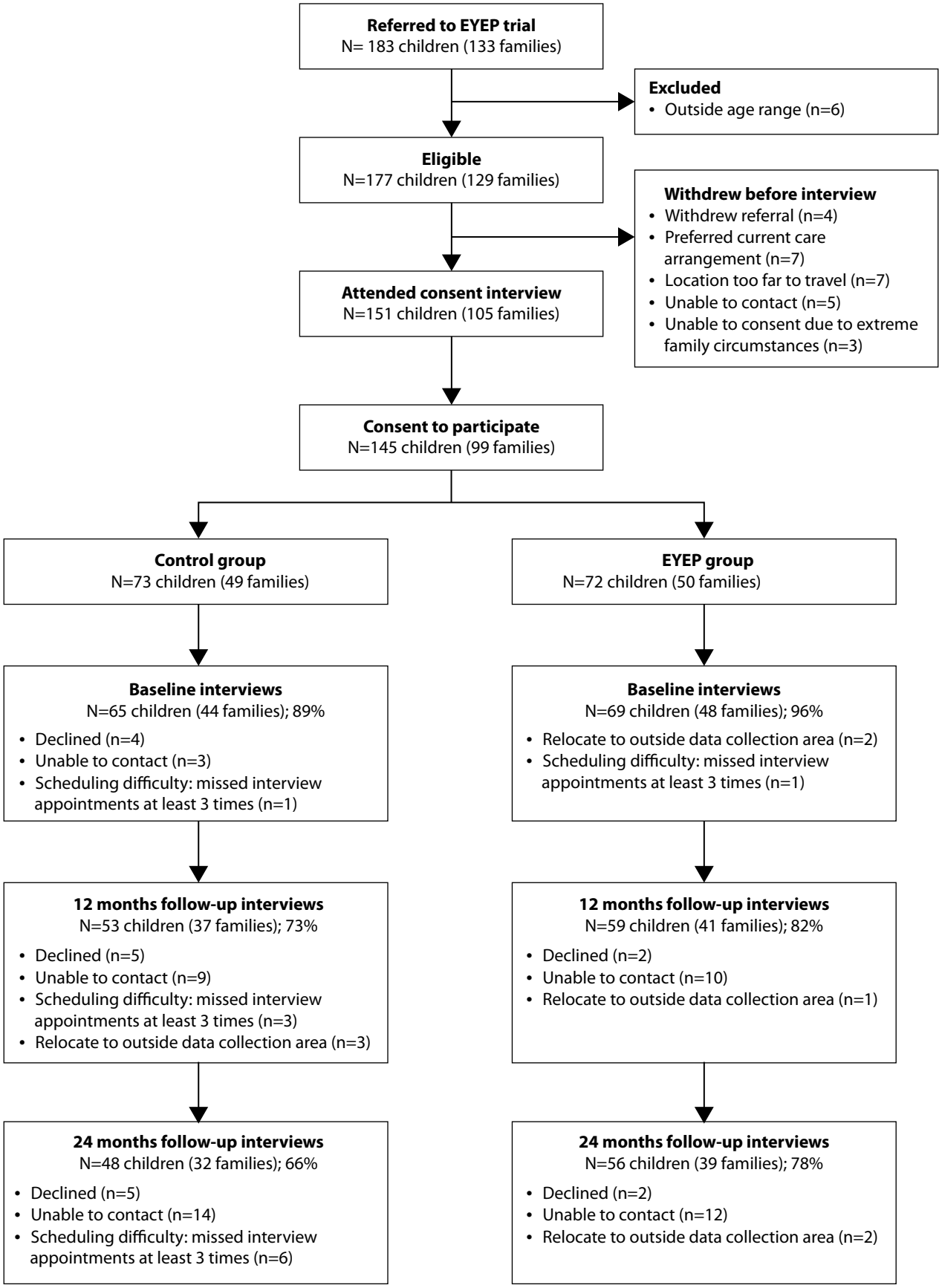
For children and primary caregivers from whom data were collected at 24 months there is also some extra non-response. For example, data for a child might have been collected on their IQ and language development, but not collected on their home environment. The main reason for non-response on specific variables was scheduling issues and time constraints. For example, some tests must be completed in a specified time period relative to months since entry in order to be valid. Other studies on populations of children and families with high levels of disadvantage have experienced similar difficulties in collecting complete information for all children (for example, St. Pierre et al., 2005; US Department of Health and Human Services, 2010, p. 2–19).

Table 4 provides information on the number of responses available for each outcome measure and for the various samples used in the analysis in this report. Numbers of

Table 4: Outcome variables – Sample sizes

	At time of entry to the EYEP trial			At 24 months after entry to the EYEP trial		
	EYEP group		Control group	EYEP group		Control group
	All	Attendance at least 60 days (first 24 months)	All	All	Attendance at least 60 days (first 24 months)	All
Child development – IQ	68	54	56	53	50	44
Child development – Language	68	54	56	53	50	44
Child development – Protective factors related to resilience	67	53	55	50	46	44
Child social-emotional development				51	47	41
Parental psychological distress	68	53	61	51	47	48
Parenting daily hassles	63 (frequency); 61 (intensity)	48 (frequency); 46 (intensity)	57 (frequency); 57 (intensity)	46 (frequency); 47 (intensity)	42 (frequency); 43 (intensity)	42 (frequency); 41 (intensity)
Home environment	64	49	55	31	30	35

Figure 1: Flow chart of EYEP trial participation and attrition



responses are shown for the time of entry to the trial and at 24 months; and separately for the intervention and control groups. On most outcomes at 24 months data are available for between 46 to 53 children who were enrolled in EYEP, and for 41 to 48 children in the control group. However, data for the HOME outcome are available only for 31 children enrolled in EYEP and 35 children in the control group.

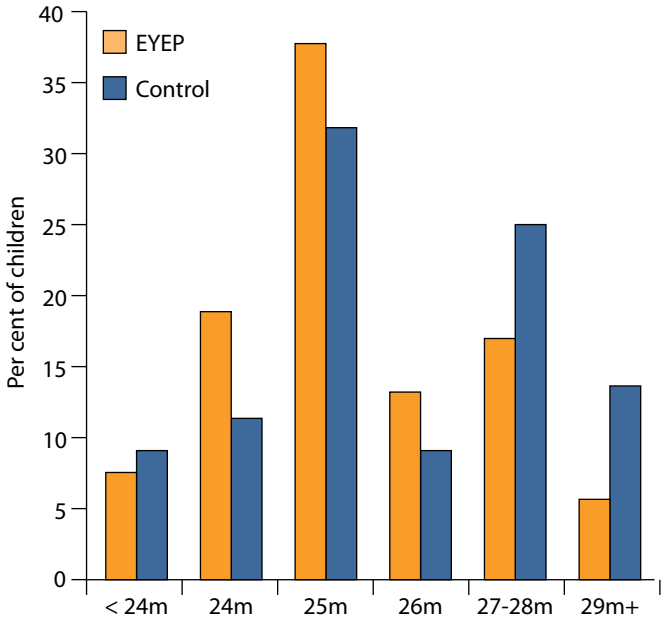
Timing of data collection

Figure 2 presents summary information on the timing of data collection for the sample of children whose outcomes are studied in this report. The summary information is for the time interval between consent being given for children to participate in the trial and data collection on their IQ at 24 months. Data for most children were collected in a timely manner, with IQ assessments taking place for about 95 per cent of children within the six-months window around the two-year anniversary of their entry to the trial. Delays in data collection, where they have occurred, have been concentrated among the control group. This is explained by greater difficulties in scheduling data collection for this group, compared with the intervention group who are attending the EYEP centre.

Method of data collection

Data collection and analysis in this project have been non-blind. With it being easiest to collect data for the intervention group at the EYEP centre, and with some data items being related to assignment status, it would have been impossible to undertake blind data collection for this trial. Similarly, continuous monitoring of the numbers of children in the intervention and control groups remaining in the trial meant it was not possible to undertake the empirical analysis in a genuinely blind manner.

Figure 2: Length of time between time of consent to participate in trial and IQ assessment at 24 months data collection



Note: 24m means 731–759 days.

7. Details of participation in EYEP by the intervention group and in early years care and education services by the control group in the first 24 months

The estimated impact of EYEP is based on comparisons of outcomes between the intervention and control groups after 24 months in the trial. During that time the main difference between the groups is the early years care and education services they receive. First, the groups may differ in their intensity of usage of early years care and education services. Second, there are differences between the design and attributes of EYEP, which is attended by the intervention group, and the services received by children in the control group. Hence, the estimated impact of EYEP will depend on the quality of EYEP and engagement of intervention group children with the program, compared to the amount and quality of early years care and education received by the control group.

An important corollary is that estimates of the impact of an early years program can only be interpreted in the context of the amount and quality of services received by the control group. Recent analysis of the impact of demonstration early years programs in the United States has reinforced this point (see Elango et al., 2015, p. 8; Almond et al., 2018, pp. 1430–31). For example, one likely explanation for the progressive decreases in the estimated impacts of early years programs targeted at disadvantaged children since the 1960s is the increase over time in the amount and quality of early years care and education available to children from disadvantaged backgrounds in the control groups in those trials.

Attendance patterns at EYEP

Attendance at EYEP by the intervention group in the first 24 months in the trial is described in Figures 3a and 3b. These attendance data come from EYEP administrative records. Figure 3a shows the distribution of total days

Figure 3a: Days attended at EYEP after 24 months – All children assigned to EYEP

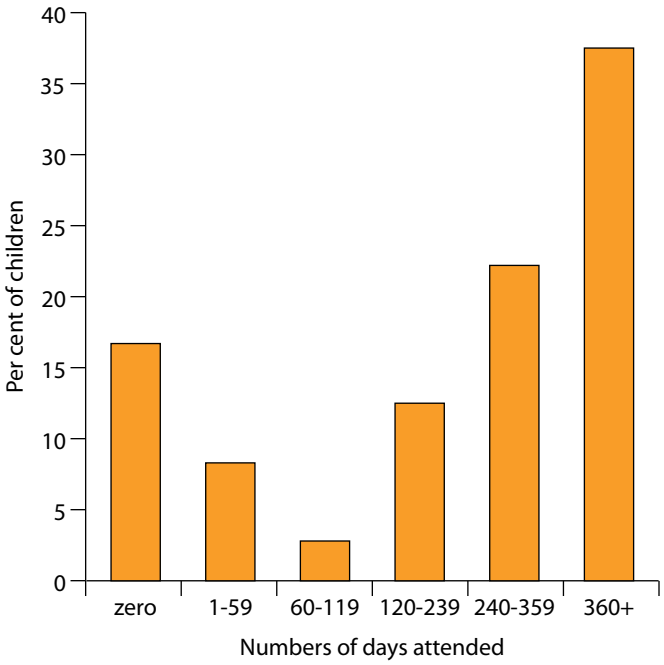


Figure 3b: Percentage of available days at EYEP attended after 24 months – Children assigned to EYEP who attended for more than 60 days

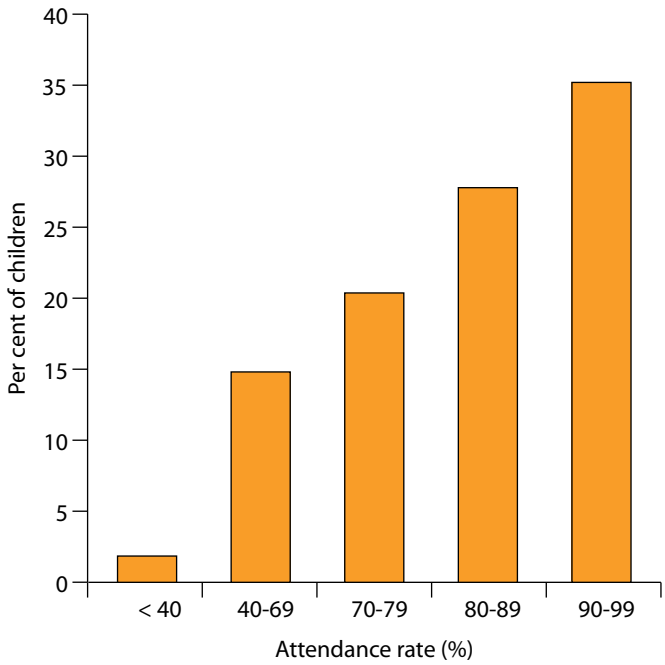
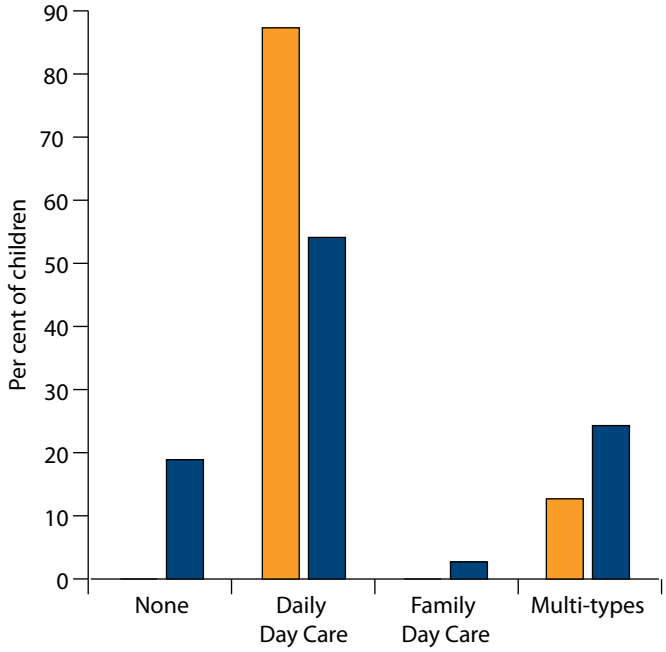
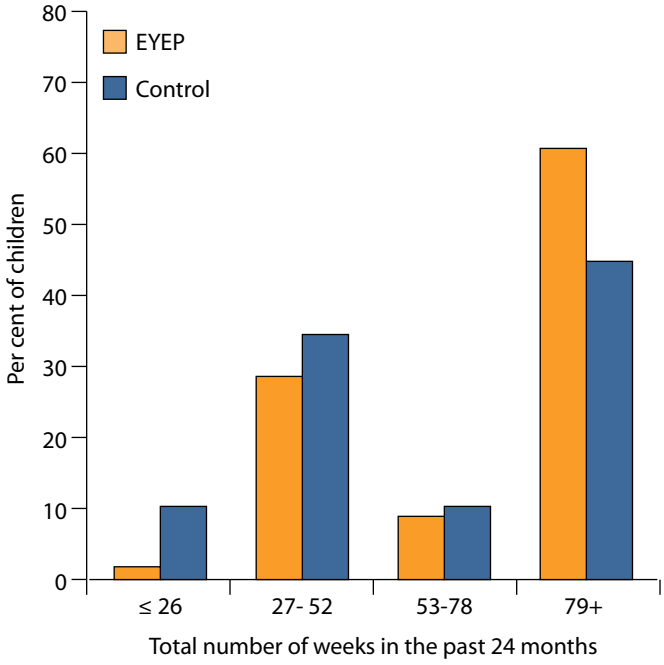


Figure 4a: Type of early years care and education attended by children in the EYEP and control groups in the first 24 months of EYEP trial



Note: Sample is children who provided data on IQ at time of entry to the trial as well as data on annual child care usage at both twelve months and 24 months.

Figure 4b: Total number of weeks of formal early years care and education in the first 24 months of EYEP trial



Note: See Figure 4a.

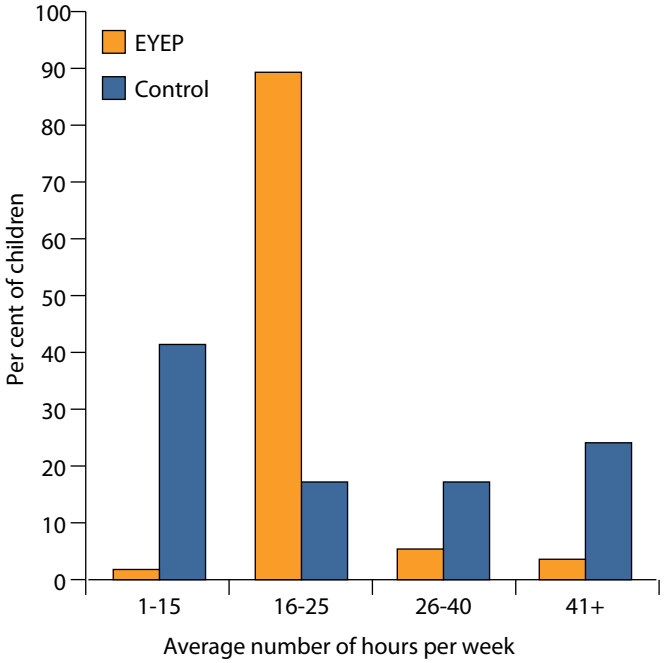
attended over the first 24 months for the full sample of children assigned to EYEP. The child at the median of the distribution attended EYEP for 313.5 days, a bit over three-fifths of the available (approximately) 480 days. The distribution of days attended is quite dispersed – with one-quarter of children attending for less than 120 days and one-third for more than 360 days. Figure 3b shows the attendance rate (proportion of available days attended) over the first 24 months for children who attended EYEP for more than 60 days in that time. This is the sample of children enrolled in EYEP for whom estimates of impact of the EYEP will be derived. Most children in this sample had relatively high rates of attendance at EYEP over the first 24 months. The attendance rate for the child at the median of the distribution was 83.4 per cent. All but fifteen per cent of children attended for at least 70 per cent of available days, and 35 per cent had an attendance rate of at least 90 per cent. Compared to attendance rates at twelve months, this represents an increase of about ten percentage points in the proportion of children attending for more than 90 per cent of available days, and a slight increase of about five percentage points in the proportion attending for less than 70 per cent of days.

Comparison of early years care and education services received by the intervention and control groups

Descriptive information on the type and amount of formal early years care and education used by children in the intervention and control groups over the first 24 months in the EYEP trial is shown in Figures 4a to 4d. Information on children's participation in early years care and education outside EYEP for the previous twelve months was obtained from primary caregivers at the annual data collection. For the control group this data source is used as the measure of their total usage of early years care and education. For the intervention group total usage is defined to be equal to attendance at EYEP plus usage of other services.

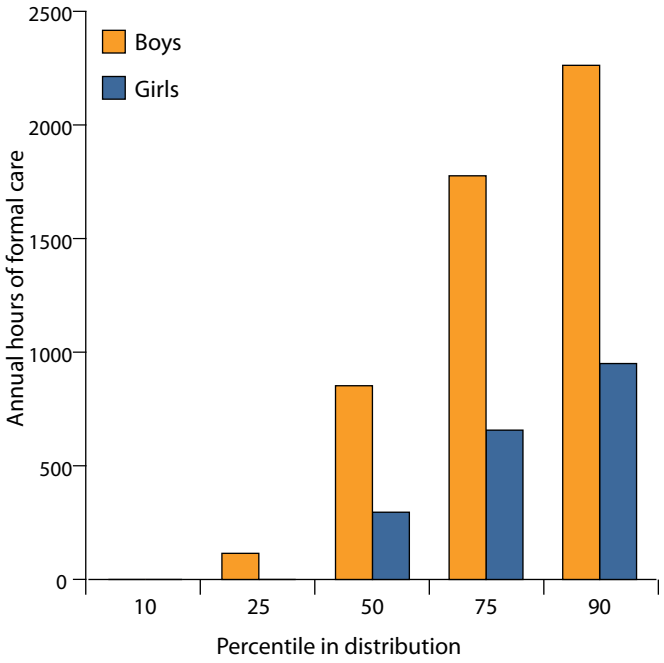
The information presented in Figures 4a to 4d is based on the sample of children for whom data was collected on usage of formal care both at twelve months and 24 months in the trial. Hence, the sample of children enrolled in EYEP in these figures differs from the sample used to report on attendance patterns at EYEP. As well, in interpreting information on usage of early years care and education services by the control group, it is important to keep in mind that eligibility for the EYEP trial required a child to 'be currently engaged with family services or child protection services and have early education as part of

Figure 4c: Usual weekly hours in formal care and education in the first 24 months of EYEP trial



Note: See Figure 4a.

Figure 4d: Distribution of annual hours in formal care and education in the first 24 months of EYEP trial – Control group by gender



Note: See Figure 4a.

their care plan'. This is likely to cause higher levels of usage of early education services in the control group than would otherwise be the case.

Figure 4a shows the types of early years care and education services received by the intervention and control groups in the first 24 months of the trial. Most children in the intervention group had 'Day Care Centre' as their only type of care, with a small proportion using multiple types of services. The category of 'Day Care Centre' includes centre-based child care and kindergarten. In the control group about one-half of children had 'Day Care Centre' as their care type, with the remainder being evenly distributed between not using early years care and education services and using multiple types of services.

Figures 4b and 4c compare the extent of use of early years care and education services between children in the intervention and control groups – for those children who used some type of services. Annual hours of services received by children enrolled in EYEP were relatively concentrated, with a majority of children receiving services for greater than 18 months and for 16 to 25 hours per week. By contrast, the distribution of annual hours of services for children in the control group was more dispersed, with concentrations of children who used services for small and large numbers of hours. The dispersion is due to two features of usage of early years care and education services by the control group – on the one hand, a relatively large proportion who spent few weeks in early years care and education services; but on the other hand, those children who attended for greater numbers of weeks being more likely to spend above 25 hours per week using the services.

chart. Per cent of childr
average number of hours

For the control group there also appear to be quite large differences by gender in the usage of early years care and education services. Figure 4d shows the annual hours of usage of services for children in that group at different percentile points in the distribution of annual hours, separately for boys and girls. It is evident that boys had much higher usage of services in the first 24 months of the trial. For example, the median of the distribution of annual hours for boys is 993.3 hours compared with 483.1 hours for girls.

Summary

Children enrolled in EYEP had received a greater amount of early years care and education services than children in the control group after 24 months of the trial – on average 20.4 hours per week compared with 15.7 hours per week. Hence, any impact of EYEP on outcomes for children after 24 months of the trial may partly derive from this difference, as well as from differences between the quality and attributes of EYEP and the services received by the control group.

8. Empirical methodology

Objective

The main goal of the empirical analysis is to estimate the impact of EYEP on outcomes for children and their primary caregivers after 24 months of enrolment in the program. The estimated impact is intended to show the difference in an outcome between children or primary caregivers in the intervention control groups that can be attributed solely to enrolment in EYEP.

As an illustration, suppose the impact of EYEP on IQ for children is estimated to be plus five points. This means that after 24 months, the intervention group is assessed, on average, to score five points higher for their IQ, compared to if they had not been enrolled in EYEP and instead received usual early years care and education.

en versus
s per week.

What impact of EYEP is estimated?

We investigate three aspects of the impact of EYEP at 24 months. First, an average impact of EYEP across all children (or primary caregivers) is estimated for each outcome. Second, the average impact of EYEP on each outcome is estimated separately for boys and girls. For these two aspects the impact of EYEP on any outcome is estimated using the sample of all children or primary caregivers for whom data on that outcome were available from the data collection at 24 months. Third, the evolution of the impact of EYEP is examined – comparing between impacts at twelve months and 24 months. To do this it is necessary to restrict attention for each outcome to a sample of children and primary caregivers for whom information on that outcome was available in both time periods. (For analysis of the impact of EYEP on IQ and language skills it is also necessary to have data on those outcomes at the time of entry to the trial).

Throughout, the impact of EYEP is estimated by comparing outcomes for children in the intervention group who attended the program for at least 60 days in the first 24 months with outcomes for children in the control group. Hence, we exclude children for whom consent was given to participate in the trial, and who were assigned to EYEP, but who never attended the program; as well as those children who had attended the program for less than 60 days after 24 months. The threshold of 60 days is the same as in the report on the impact of EYEP at twelve months (Tseng et al., 2018). Making this restriction ensures that children have spent the minimum amount of time attending EYEP needed for the program to have had an impact on them. It is also important to note that setting this threshold causes only three children or fewer who attended EYEP to be excluded from the analysis, depending on the outcome.

Identifying the causal impact of EYEP

Program evaluation is intended to provide an estimate of the impact of a program on an outcome that can be interpreted as causal. That is, the estimate should reflect only that part of the difference in the outcome between the intervention and control groups that is due to the program; and it should exclude, for example, any difference in the outcome due to differences in the characteristics of individuals in those groups.

A major potential advantage of a RCT is that it allows the impact of a program to be evaluated simply by comparing the average values of an outcome between the intervention and control groups. Randomisation implies that trial participants assigned to either group have the same characteristics, the only difference being that the intervention group has participated in the program being studied. Hence, any significant difference in the average values of an outcome between the groups can be attributed to a causal effect of the program. This property only holds, however, where random assignment results in balance between the characteristics of the intervention and control groups at their time of entry to the trial, and where attrition from the trial since entry has not subsequently created imbalance in these characteristics.

In an earlier report we assessed the balance between the characteristics of children and their primary caregivers in the intervention and control groups at their times of entry to the trial (Tseng et al., 2018). Random assignment in the EYEP trial was found to have been implemented

successfully, with balance achieved for almost all characteristics on which data were collected at the time the primary caregiver consented to participate in the trial. Significant differences, however, were found to exist between the intervention and control groups for some key characteristics, for which data were collected at the subsequent stage when a child in the intervention group commenced participation in EYEP (or would have commenced for children from the control group). In particular, the Bayley Scales outcome measures for children's cognitive development, motor skills, and social-emotional development were found to be unbalanced. In the earlier report we showed that the main reason the Bayley Scales measures are significantly higher for the control group than the intervention group is the sample attrition that occurred between the time of consent to participate in the trial and the subsequent stage of data collection when Bayley Scales assessments were undertaken (Tseng et al, 2018, pp. 17–18).

For this report we have done extra analysis of sample attrition through to 24 months. Several further variables have now been found to be unbalanced in the remaining sample – first, child and family risk factors relating to alcohol/substance abuse, disability or complex medical needs, and family violence; and second, whether a child is from a family from which multiple children entered the EYEP trial at the time of referral.

Imbalance between characteristics of the intervention and control groups needs to be taken account of by the method used to estimate the impact of EYEP. Where imbalance exists, differences in outcomes between the groups can reflect either the impact of EYEP or differences in their characteristics. Identifying the casual impact of EYEP therefore requires a method that removes the effect that imbalance in characteristics would otherwise have on the estimated impact of EYEP.

To illustrate, suppose that in the sample of children for whom data have been collected at 24 months, those enrolled in EYEP are less likely than children in the control group to live with a primary caregiver with the risk factor of alcohol or substance abuse. A finding that being enrolled in EYEP is associated with a score of plus five on the IQ measure could then reflect either the impact of EYEP or the fact that children enrolled in EYEP have better home environments. Hence, it is necessary to use a method to estimate the impact on IQ that can control for the difference between children in the incidence of alcohol or substance abuse in their families.

Empirical methods applied in this study

Estimation of the impact of EYEP is done using two main methods – a regression method and a propensity score matching method. These methods have the advantages of being relatively transparent and robust with small samples (Huber et al., 2013). Both estimate the impact of EYEP while controlling for differences in the characteristics of the intervention and control groups (see Tseng et al., 2018, Appendix 6.2 for more details). Hence, using these methods implies that the estimated impact of EYEP on an outcome should reflect only the effect of the program.

What characteristics of children and their families need to be controlled for in order that the estimated impact of EYEP reflects only attendance at EYEP? Statistical theory provides some guidance – directing that it is necessary to control for characteristics of children and their families that are unbalanced between the intervention and control groups, or that might affect the outcome being examined (Stuart, 2010). Applying these criteria, however, is a matter of judgment.

In our report on the impact of EYEP at twelve months we specified a preferred model with a set of characteristics we believed should be controlled for (Tseng et al., 2018, Appendix 6.2). For example, drawing on recent research which establishes the importance of controlling for baseline outcomes, our model included variables for measures of children's IQ and language development at the time of entry to the trial (Griffen and Todd, 2017). With the extra attrition in the sample that occurred prior to the 24 months data collection having caused imbalance in some further characteristics of children and their families, we need to revisit the choice of characteristics to control for.

Our approach in this report is to present estimates of the impact of EYEP from two alternative model specifications that control for different sets of characteristics. First, the impact of EYEP on each outcome is estimated using the preferred model specification at twelve months. This is done for the sake of transparency, and because testing generally indicates that the model specification used at twelve months is also the preferred model at 24 months, even though it does not control for the extra imbalanced

characteristics. Second, estimates of the impact of EYEP on each outcome are presented for a model specification including the same set of characteristics as in preferred model specification at twelve months plus the extra characteristics now found to be unbalanced: child and family risk factors relating to alcohol/substance abuse, disability or complex medical needs, family violence, and whether a child is from a family from which multiple children entered the EYEP trial at the time of referral. Full details of the model specifications are shown in Appendix 3.

Impacts of EYEP for all children at 24 months, and the evolution of the impact over time, are estimated using regression analysis and a matching method. Findings from the regression analysis are presented in the main body of the report. Estimates of the impact of EYEP using the matching method and regression on a matched sample are presented in Appendix 5. Regression on a matched sample applies the specified set of characteristics to match EYEP participants to the control group, and then uses the same set of characteristics in a weighted regression (with weights derived from the propensity score matching). Results derived using these alternative methods are treated as a robustness check. For all outcomes, estimates of the impact of EYEP using the alternative methods are quite similar.

Separate impacts of EYEP for boys and girls are estimated using the regression method only. This is done with the same model specifications as for all children. The impact of gender is captured by including an extra variable – the interaction of gender and being enrolled in EYEP. It is not feasible to apply the matching method to derive separate estimates of the impact of EYEP for boys and girls due to small sample sizes once the sample is disaggregated by gender.

Statistical significance

The statistical significance level associated with the estimated impact of EYEP on each outcome is reported using p-values. The p-value is a way to assess how likely it is (or how confident we should be) that there is a difference in an outcome between the intervention and control groups. We follow the literature in interpreting statistical significance using what is known as a one-tailed test with 5 or 10 per cent level of significance (for example, Karoly, 2005; Elango et al., 2015; Doyle et al., 2017). The statistical significance of estimates of the impact of EYEP from the regression method is assessed with standard errors calculated using the permutation method; and standard errors for the matching method are calculated using a bootstrap method. Further details on statistical significance are provided in Appendix 4.

9. Impacts of EYEP after 24 months

Main results

Estimates of the impact of EYEP after 24 months using the regression method are presented in Table 5 for: (i) children's development (rows (1) to (4)); and (ii) outcomes for their primary caregivers and home environment (rows (5) to (8)). Estimates from both specifications of the regression model are shown. Results from robustness analysis – using the alternative matching estimation methods and (where possible) alternative specifications of the outcome variable – are reported in Appendix 5 (Tables 5.1 and 5.2).

Overall, the findings suggest that at 24 months the impact of attending EYEP on children and their families is broad and powerful. Large and significant impacts of EYEP on children's IQ, protective factors related to resilience and social-emotional development are found. There is also some evidence of a favourable impact from EYEP on children's language skills and the psychological distress of their primary caregivers.

IQ. The estimated impact of EYEP on children's IQ is positive and statistically significant. The alternative model specifications find estimated impacts of 5 to 7 points. A one-tail test of the impact of EYEP on IQ is significant at either the 5 per cent or 10 per cent levels, depending on the model specification. The size and statistical significance of the estimated impact is robust to using the matching or regression on matched sample methods, to defining the outcome as the change in IQ from the time of entry to the trial to 24 months, and to including a control for whether the Bayley Scales or WPPSI assessment was used.

Table 5: Impact of enrolment in EYEP for 24 months – Regression estimates - Children who attended for at least 60 days in the first 24 months

	Outcome		Model 1		Model 2		
		EYEP mean (24 months)	EYEP impact	1-tail p-value	EYEP impact	1-tail p-value	Number of observations (EYEP/ Control)
	Children's development						
(1)	IQ	99.56	5.010	0.068	7.057	0.017	50/43
(2)	Language	96.16	2.865	0.225	5.153	0.072	50/43
(3)	Protective factors	44.93	3.289	0.047	3.340	0.059	46/41
(4)	Social and emotional (Per cent in clinical range; Below norm for bottom 10% of population)	12.8	-31.6	0.001	-29.2	0.003	47/39
	Primary caregiver						
(5)	Psychological distress	13.91	-1.651	0.098	-1.776	0.098	46/45
(6)	Parenting daily hassles - Frequency	44.78	-1.011	0.324	-2.111	0.171	42/41
(7)	Parenting daily hassles - Intensity	43.55	-3.028	0.200	-2.976	0.228	43/40
(8)	**Home environment**	69.18	-0.549	0.440	-0.452	0.454	30/32

Notes:
1] Model 1 is the basic method from the twelve months report. Model 2 is the basic method with extra covariates to control for variables that are unbalanced at 24 months. See Appendix 3 for details.
2] The EYEP impacts in rows (1)-(3) and (5)-(8) are the estimated impacts of attending EYEP from an OLS regression. The EYEP impact in row (4) is the marginal impact on the per cent of children below the 10% threshold on the social-emotional measure estimated from a probit model.

On the IQ scale, an increase of 5 to 7 points is relatively large, representing about one-third to one-half of a standard deviation. By comparison, recent reviews of early years demonstration programs in the United States generally find average impacts on IQ of about one-quarter of a standard deviation. For example, Karoly et al. (2005, p. 67) review estimates of impacts on IQ for children near to or soon after commencing in primary school from 20 studies and find an average estimated impact of 0.28 of a standard deviation. Other reviews of programs from the United States find average impacts on IQ of 0.23 (Camilli et al., 2010) and 0.21 (Duncan and Magnuson, 2013).

Some caution does need to be exercised in making comparisons between the estimated impacts of EYEP and average impacts of these other programs. First, the estimate for EYEP is in-program whereas for programs included in the reviews cited above estimates are often from after the program has concluded – and the phenomenon of fade-out in the impact on IQ from early years programs is well-established (Hojman, 2015; Elango et al., 2015, pp. 31–32). Second, estimated impacts of other programs on IQ do tend to be larger for more intensive programs such as EYEP. Against this, however, it is important to note that estimated impacts of early years programs have tended to be smaller for programs implemented in more recent years (Elango et al., 2015, p. 32; Duncan and Magnuson, 2013, p. 123).

Language. The estimated impact of EYEP on children's language skills is positive but has limited statistical significance. The alternative model specifications find an estimated impact of 3 to 5 points, about one-quarter of a standard deviation. Similar results on the size of impact of EYEP are found from the robustness analysis using alternative estimation methods, the alternative outcome definition, and controlling for the type of language assessment. At this stage, however, there is not as strong evidence of statistical significance as for the other measures of child development. Across the eight estimates of the impact of EYEP on language skills using the alternative estimation methods and specifications, a one-tail test achieves significance at the 10 per cent level only twice.

Within-child protective factors related to resilience. The estimated impact of EYEP on children's protective factors is positive and statistically significant. The estimated impact of EYEP is about 3.5 points which is equal to one-third of a standard deviation. It is significant with a one-tailed test at about the 5 per cent level in both model specifications. The size and statistical significance of the estimated impact of EYEP is robust to using matching or regression on matched sample.

Social-emotional development. The estimated impact of EYEP on children's social-emotional development is positive and statistically significant. The proportion of children enrolled in EYEP who are classified in the clinical range for development is lower by 30 percentage points than the control group, which is a substantial impact. Interestingly, the impact appears to derive both from a decrease in the proportion of children enrolled in EYEP who are in the clinical range compared to at twelve months, and an increase in the proportion of children in the clinical range for the control group. In both model specifications the estimated impact is significant with a one-tailed test at the 1 per cent level. The size and statistical significance of the estimated impact of EYEP is robust to using the matching estimation method and to controlling for the type of assessment of social-emotional development used. A further robustness check is to use bounds analysis to estimate upper and lower bounds for the impact of EYEP (see Appendix 4 for more details of this method). The bounds estimated for the decrease in the proportion of children in the clinical range for development are from 7.7 per cent to 50.6 per cent, with the upper but not lower bound being statistically significant.

Psychological distress of primary caregivers (K6). EYEP is estimated to reduce the psychological distress of primary caregivers, an effect which is marginally statistically significant. The size of impact estimated using the regression method is a decrease of about 1.5 points on a zero to 30 points scale. In both model specifications the estimated impact is significant with a one-tailed test at about the 10 per cent level. The size of estimated impact is robust to using the matching estimation method, but the estimated impact is not close to statistically significant at the 10 per cent level.

Parenting daily hassles. The impact of EYEP is estimated as small decreases in the frequency (1 point on zero to 80 points scale) and in the intensity (3 points on zero to 100 points scale) of parenting daily hassles. Neither estimated impact is close to being significant with a one-tailed test at the 10 per cent level for either regression

model specification or for the alternative matching estimation method.

HOME. The estimated impact of EYEP on the home environment is small and not statistically significant. Using the regression method the impact is estimated to be a reduction in the quality of the home environment by 0.5 points on a zero to 100 points scale. This estimated impact is not at all close to being statistically significant with a one-tailed test.

Impacts by gender

Table 6 reports estimates of the impact of EYEP on each outcome separately for boys and girls. Estimates are reported for both specifications of the regression model.

Quite distinct patterns of impact are found for boys and girls.

For children's IQ and language skills the estimated impact of EYEP is larger and more statistically significant for boys than girls. However, the size of the gender gap does depend on which regression model specification is used; and the difference in estimated impact between boys and girls is not as large as was estimated at twelve months.

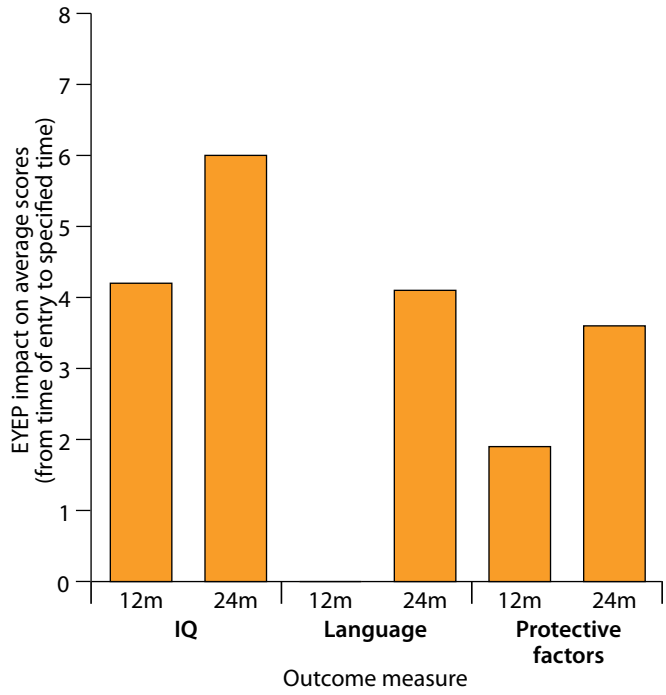
Non-cognitive skills show the biggest difference between boys and girls. For protective factors related to resilience a large (about one-half standard deviation) and highly significant impact is found for boys, compared to a zero impact for girls. The estimated impact of EYEP on social-

Table 6: Impact of enrolment in EYEP for 24 months by gender – Regression estimates – Children who attended for at least 60 days in the first 24 months

Outcome			Model 1		Model 2		
		EYEP mean (24 months)	EYEP impact	1-tail p-value	EYEP impact	1-tail p-value	No. of observations (EYEP/Control)
Child development							
IQ	Boys	95.81	6.389	0.056	7.274	0.039	26/25
	Girls	103.63	3.106	0.226	6.699	0.053	24/18
Language	Boys	92.69	4.666	0.146	5.702	0.084	26/25
	Girls	99.92	0.378	0.470	4.251	0.190	24/18
Protective factors	Boys	45.33	5.465	0.013	5.369	0.015	24/24
	Girls	44.50	0.150	0.478	-0.002	0.500	22/17
Social and emotional	Boys	16.0	-13.2	0.171	-10.4	0.241	25/23
	Girls	9.0	-59.3	<0.001	-52.7	<0.001	22/16
Primary caregiver							
Psychological distress	Boys	14.68	-0.977	0.262	-1.033	0.265	25/27
	Girls	13.00	-2.565	0.093	-2.973	0.076	21/18
Parenting daily hassles - Frequency	Boys	47.27	2.022	0.233	1.122	0.344	22/24
	Girls	42.45	-3.196	0.252	-3.146	0.261	20/17
Home environment							
	Boys	70.19	3.141	0.287	3.310	0.296	13/21
	Girls	68.40	-5.864	0.104	-6.510	0.106	17/11

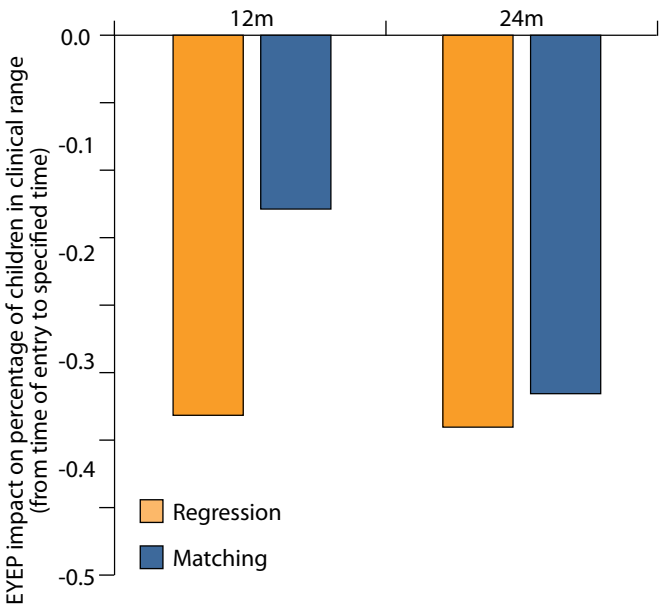
Note: See Table 5.

Figure 5a: Impact of enrolment in EYEP after twelve months and 24 months – Regression method – Selected outcome measures for children – Impact on average scores



Note: 1] Sample is children who attended for at last 60 days and for whom data were collected on an outcome at twelve months and 24 months; and 2] Regression method is applied using the model specification 2.

Figure 5b: Impact of enrolment in EYEP after twelve months and 24 months – Social-emotional development – Impact on percentage of children in clinical range



Note: See Figure 5a. Matching method is applied using the model specification 2.

emotional development exhibits the opposite pattern. For girls there is a substantial and highly statistically significant estimated impact, with EYEP decreasing the proportion in the clinical range by 50 percentage points, whereas for boys the estimated impact is a decrease of only 10 percentage points which is not statistically significant.

Impacts on outcomes for primary caregivers at 24 months are confined to families with girls enrolled in EYEP. For the primary caregivers of girls there is a decrease in psychological distress (K6) of about 3 points (zero to 30 points scale), and in the frequency of parenting hassles of 5 to 7 points (zero to 80 points scale). Both impacts are statistically significant at the 10 per cent level. No significant impact on intensity of parenting hassles for primary caregivers of girls is found. For primary caregivers of boys there is no evidence of a significant impact on any of these outcomes related to psychological distress or parenting hassles of primary caregivers. No significant impact of EYEP on home environment is found for either boys or girls.

Evolution of impacts over 24 months

Figures 5a to 5c present estimates of the impact of EYEP at twelve and 24 months using the regression method (model specification 2). Detailed information on estimates of the impacts of EYEP at twelve and 24 months using the regression and matching methods are presented in Appendix 5 (Table 5.3).

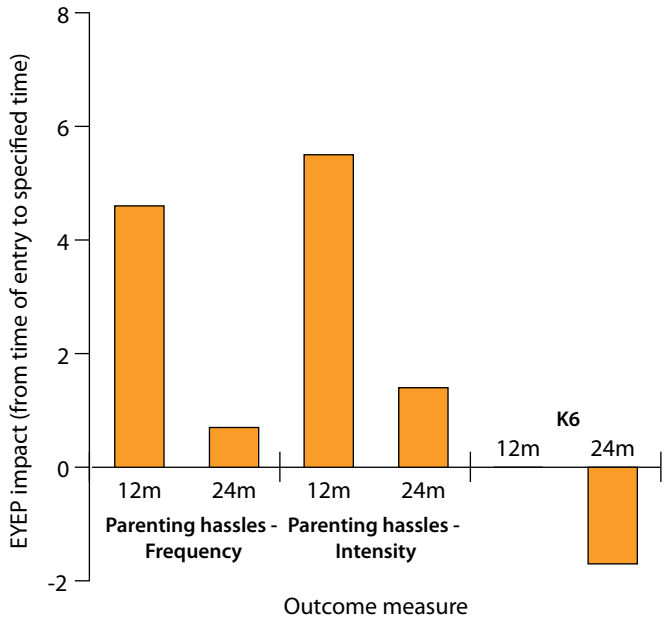
For each outcome these estimates are for the sample of children and primary caregivers for whom data on that outcome were collected at both twelve months and 24 months (and for IQ and language skills information is also required at the time of entry to the trial). Requiring children to have data available on an outcome at both twelve and 24 months means that the sample size is reduced compared to the analysis of impacts at 24 months presented thus far in this report. With a smaller sample size the levels of statistical significance for the estimated impacts of EYEP are generally lower.

The estimated impact of EYEP on children's IQ remains at about the same level between twelve months and

24 months, being 4 to 6 points at both times. This pattern matches with findings using the full samples of children for whom data on IQ are available at twelve and 24 months. Hence it seems that thus far the impact of EYEP on IQ has been concentrated in the first twelve months of attending the program. This result is consistent with evidence from previous trials of early years demonstration programs. For example, Hojman (2015) examines six early childhood programs in the United States and concludes that gains in IQ experienced by the intervention group occurred rapidly in the first few months of the program and were followed by small or zero gains in subsequent years.

Other outcomes for children show a more pronounced impact of EYEP after the second year of being enrolled in the program than after the first year. The estimated impact on children's language skills increases from zero after twelve months to about 3 to 4 points after 24 months. For protective factors related to resilience the estimated impact size after 24 months is two to three times as large as after twelve months. These patterns are found using both estimation methods, and are consistent with findings when impacts are estimated for the full samples of children for whom data on these outcomes are available at twelve and 24 months.

Figure 5c: Impact of enrolment in EYEP after twelve months and 24 months – Regression method – Selected outcome measures for primary caregivers



Note: See Figure 5a

Distinguishing the exact timing of the impact of EYEP on social-emotional development is more difficult. The estimated impact size and significance level are higher after 24 months than twelve months, regardless of the estimation method or model specification. Whether it was during the first or second year when EYEP had the largest impact on social-emotional development, however, cannot be determined. The most that can be said is that some positive impact on social-emotional development occurred in each year.

For outcomes for primary caregivers, the positive impacts of EYEP have been concentrated in the second year of their child's enrolment in the program. The estimated impact on the K6 measure of psychological distress was zero after twelve months and a reduction of 1.5 points after 24 months. This pattern is robust to the estimation method and is consistent with findings when the full samples of primary caregivers for whom data are available at both times are used. For parenting daily hassles, after twelve months EYEP was associated with increases in the frequency and intensity of hassles, but that effect was reversed by 24 months. This is different from results using the full samples at twelve and 24 months where impacts of EYEP were found to be respectively zero and slight negative impacts. For both samples, however, the pattern of impact over time is the same – that is, any decrease in parenting hassles has happened in the second year of children's enrolment in EYEP.

Summary and interpretation

By 24 months EYEP has been found to have a variety of important impacts on children's cognitive and non-cognitive development:

1] The impact on IQ has been sufficiently large that the objective of EYEP to make participants developmentally equal to their peers has been achieved. At 24 months children attending EYEP had an average IQ score of 99.6. Statistical significance tests do not reject this as being equal to the population average score of 100. The estimated impact of EYEP on IQ is comparable in size to impacts from early years demonstration programs implemented in the United States. Some weaker evidence of an impact from EYEP on language skills is also found. Slightly larger and more significant impacts from EYEP on cognitive development are estimated for boys than girls. The impact of EYEP on IQ is evident after twelve months

and then stable to 24 months, whereas the impact on language was concentrated in the second year of the program.

2] The impacts on protective factors related to resilience and social-emotional development seem large – for example, a decrease of 30 percentage points in the proportion of children in clinical range for social-emotional development. However, it is not possible to benchmark these impacts on non-cognitive development against estimated impacts from early years programs in the Untied States (partly due to problems in comparing between alternative measures of non-cognitive development and partly because a smaller set of studies have examined non-cognitive outcomes). A larger impact on social-emotional development is estimated for girls than boys, with the opposite pattern for protective factors related to resilience. EYEP appears to have benefitted children's social-emotional development during both years, whereas the impact on resilience was confined to the second year of the program.

At 24 months there is also some evidence of impacts from EYEP on the well-being of primary caregivers. A marginally significant impact in reducing psychological distress of primary caregivers is found. That impact occurs only for primary caregivers of girls. When attention is restricted to that group of caregivers, a significant impact from EYEP in lowering the frequency of parenting hassles is also found. The impact of EYEP on the psychological well-being of primary caregivers is concentrated in the second year of the program. No impact on home environment is detected at 24 months.

The findings at 24 months raise two interesting questions about how EYEP is affecting children and their primary caregivers. First, with regard to impacts by gender, why are the impacts on IQ, language skills, and resilience largest for boys and the impacts on social-emotional development and the psychological well-being of the primary caregiver largest for girls? Second, in relation to the timing of impacts on outcomes: Why is the impact on IQ concentrated in the first year? Why is the impact on social-emotional development spread across the 24 months? And, why do the impacts on other outcomes occur in the second year?

On the first question of impacts by gender, existing international studies of centre-based early years interventions for the most part find either no difference in impacts by gender or larger effects for boys than girls (Magnuson et al., 2016; Garcia et al., 2018, Appendix C; Elango et al., 2015, p. 33; Anderson, 2008). The primary explanation provided for why there is a larger impact for boys than girls is that boys' development is more sensitive to the quality and stressfulness of their home and care environments. Therefore, boys are likely to begin behind girls at the time they commence in high quality early years programs and will also benefit more from the program (for example Autor et al., 2016; Golding and Fitzgerald, 2017; Bertrand and Pan, 2013). However, there are also exceptions to this pattern – for example, larger impacts on social-emotional development for girls than boys are found in re-evaluations of the Perry Preschool and Abecedarian projects (Heckman et al., 2013; Garcia et al., 2018).

Our finding that relative impacts of EYEP by gender are mixed is therefore not entirely at odds with previous studies of early years programs in the United States – and evidence on the impact of EYEP on cognitive development for boys and girls seems largely consistent with those studies. Overall, however, it has to be said that the evidence on the relative size of impacts by gender is more mixed in our study than in that literature. This difference from previous studies may simply be due to features of the sample of participants in the EYEP trial. Alternatively, analysis of EYEP may be yielding new insights into impacts by gender due to the wide range of outcome variables being considered. Sorting out the sources of gender differences in the impact of EYEP is an important future task.

The timing of impacts of early years interventions on outcomes is not a topic that has been much addressed in previous studies of programs in the United States. To the extent that most of the impact of EYEP has become evident during children's second year of attendance at the program, it could be interpreted as showing the amount of time necessary to begin to have an impact on children from highly disadvantaged and stressful environments. However, as the impact of EYEP on IQ thus far has been concentrated in the first year of the program, as well as some positive impact on social-emotional development in that time, that does not seem a sufficient explanation for the timing of impacts. The spread of timing of impacts could reflect design details of EYEP, but might also be due to there being a natural progression of development for children. Seeking to understand more about the ordering of impacts is therefore also a valuable future goal.

References

Achenbach, T. and L. Rescorla (2000), *Manual for the ASEBA Preschool Forms and Profiles* (Burlington: VT, University of Vermont, Research Center for Children, Youth and Families).

Almond, D., J. Currie and V. Duque (2018), 'Childhood circumstances and adult outcomes: Act II', *Journal of Economic Literature*, 56(4): 1360-1446.

Anderson, M.L. (2008), 'Multiple inference and gender differences in the effects of early interventions: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects', *Journal of the American Statistical Association*, 103(484), 1481-95.

Australian Children's Education and Care Quality Authority (2011), National Quality Standard 2011.

Australian Institute of Health and Welfare, (2018), *Child Protection Australia 2016-17*, Child Welfare series no.66, Catalogue no. CWS 63 (Canberra: AIHW).

Autor, D., D. Figlio, K. Karbownik, J. Roth and M. Wasserman (2015), 'Family disadvantage and the gender gap in behavioural and educational outcomes', Institute for Policy Research, Northwestern University Working Paper Series WP15-16.

Bayley, N. (2006), *Bayley Scales of Infant and Toddler Development* (Third edition, San Antonio, TX: Harcourt Assessment, Inc.).

Bertrand, M. and J. Pan (2013), 'The trouble with boys: Social influences and the gender gap in disruptive behavior', *American Economic Journal: Applied Economics*, 5(1), 32-64.

Biddle, N., R. Seth-Purdie and H. Crawford (2017), 'Risk burden, participation in early childhood education and care, and child outcomes', *Australasian Journal of Early Childhood*, 42(1), 49-59.

Bierhaus, A., J. Wolf, M. Andrassy, N. Rohleder, P.M. Humpert, D. Petrov, D. Ferstl, M. von Eynatten, T. Wendt, G. Rudofsky, M. Joswig, M. Morcos, M. Schwaninger, B. McEwen, C. Kirschbaum and P.P. Naworth (2003), 'A mechanism converting psychosocial stress into mono-nuclear cell activation', *Proceedings of the National Academy of Sciences*, 100(4), 1920-5.

Briggs-Gowan, M. and A. Carter (2006), *BITSEA Brief Infant-Toddler Social and Emotional Assessment* (Bloomington, MN: Pearson).

Caldwell, B., and R. Bradley (2003), *Home Inventory Administration Manual – Comprehensive Edition* (Eau Claire: Wisconsin, Minuteman Press).

Camilli, G., S. Vargas, S. Ryan, and W.S. Barnett (2010), ' Meta-analysis of the effects of early education interventions on cognitive and social Development', *Teachers College Record*, 112(3), 579–620.

Campbell, F.A., G. Conti, J.J. Heckman, S.H. Moon, R. Pinto, E.P. Pungello and Y. Pan (2014), 'Early childhood investments substantially boost adult health', *Science*, 343, 1478-85.

Campbell, F.A. and C.T. Ramey (1994), 'Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low income families', *Child Development*, 65(2), 684-98.

Cannon, J., M. Kilburn, L. Karoly, T. Mattox, A. Muchow and M. Buenaventura (2017), *Investing Early: Taking Stock of Outcomes and Economic Returns from Early Childhood Programs* (Rand Corporation).

Carrell, S. and M. Hoekstra (2010), 'Externalities in the classroom: How children exposed to domestic violence affect everyone's kids', *American Economic Journal: Applied Economics*, 2(1), 211-28.

Carrell, S., M. Hoekstra and E. Kuka (2018), 'The long-run effects of disruptive peers', *American Economic Review*, 108, 3377-3415.

Caspi, A., R. Houts, D. Belsky, H. Harrington, S. Hogan, S. Ramrakha, R. Poulton, R. and T. Moffitt (2017), 'Childhood forecasting of a small segment of the population with large economic burden', *Nature Human Behaviour* 1(1), 0005.

Centre for Community Child Health (2011), *Review of the evidence base in relation to early childhood approaches to support children in highly disadvantaged communities for the Children's Ground Project*. Report prepared for the Commonwealth Department of Education, Employment and Workplace Relations (Melbourne: Murdoch Children's Research Institute).

Center on the Developing Child at Harvard University (2016a), *Applying the Science of Child Development in Child Welfare Systems*, accessed at: http://www.developingchild.harvard.edu

Center on the Developing Child at Harvard University (2016b), *From Best Practices to Breakthrough Impacts: A Science-Based Approach to Building a More Promising Future for Young Children and Families*, accessed at: http://www.developingchild.harvard.edu

Commonwealth Department of Education, Employment and Workplace Relations (2009), N*ational Early Years Learning Framework* (Canberra).

Craig, P., P. Dieppe, S. Macintyre, S. Mitchie, I. Nazareth, and M. Petticrew (2008), 'Medical Research Council Guidance. Developing and evaluating complex interventions: The new Medical Research Council guidance', *British Medical Journal*, 337(7676), 979-83.

Crnic, K.A. and M.T. Greenberg (1990), 'Minor parenting stresses with young children', *Child Development*, 61(5), 1628-37.

Cunha, F. and J. Heckman (2007), 'The technology of skill formation', *American Economic Review*, 97(2), 31-47.

Currie, J. and C. Spatz Widom (2010), 'Long-term consequences of child abuse and neglect on adult on well-being', *Child Maltreatment*, 15(2): 111-20.

Doyle, O., C. Harmon, J.J. Heckman, C. Logue and S.H. Moon (2017), 'Early skill formation and the efficiency of parental investment: A randomized controlled trial of home visiting', *Labour Economics*, 45, 40-58.

Duncan, G.J. and K. Magnuson (2013), 'Investing in preschool programs', *Journal of Economic Perspectives*, 27(2), 109-32.

Elango, S., J.L. Garcia, J.J. Heckman and A. Hojman (2015), 'Early childhood education', Human Capital and Economic Opportunity Global Working Group, Working Paper No.2015-017.

Evans, S.E., C. Davies and D. DiLillo (2008), 'Exposure to domestic violence: A meta-analysis of child and adolescent outcomes', *Aggression and Violent Behavior*, 13(2), 131-40.

Fergusson, D.M. and L.J. Horwood (2003), 'Resilience to childhood adversity: Results of a 21 year study', *Resilience and Vulnerability: Adaption in the Context of Childhood Adversities*, 130-55.

Ford, E., R. Anda, V. Edwards, G. Perry, G. Zhao, C. Li and J. Croft (2011), 'Adverse childhood experiences and smoking status in five states', *Preventive Medicine*, 53(3), 188-93.

Furukawa T.A.,R.C. Kessler R.C. and G. Andrews (2002), 'The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being', *Psychological Medicine*, 33(2), 357–62.

Garcia, J.L., J.J. Heckman and A.L. Ziff (2018), 'Gender differences in the benefits of an influential early childhood program', *European Economic Review*, 109, 9-22.

Golding, P. and H. Fitzgerald (2017), 'Psychology of boys at risk: Indicators from 0 to 5', *Infant Mental Health Journal*, 38(1), 5-14.

Good, P. (2005), *Permutation, Parametric, and Bootstrap Tests of Hypotheses* (Third edition) (Huntington Beach, California: Springer Science+ Business Media).

Griffen, A.S. and P.E. Todd (2017), 'Assessing the performance of nonexperimental estimators for evaluating Head Start', *Journal of Labor Economics*, 35(S1), S7-S63.

Heckman, J.J. (2008), 'Schools, skills and synapses', *Economic Inquiry*, 46(3), 289-324.

Heckman, J.J., R. Pinto and P. Savelyev (2013), 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes', *American Economic Review*, 103(6): 2052-86.

Heckman, J.J. and S. Mosso (2014), 'The economics of human development and social mobility', *Annual Review of Economics*, 6(1), 689-733.

Hojman, A. (2015), 'Evidence on the fade-out of IQ gains from early childhood interventions: A skill formation perspective', Mimeo, Center for the Economics of Human Development, University of Chicago.

Huber, M., M. Lechner and C. Wunsch (2013), 'The performance of estimators based on the propensity score', *Journal of Econometrics*, 175, 1-21.

Hughes, K., M.A. Bellis, K.A. Hardcastle, D. Sethi, A. Butchart, C. Mikton, L. Jones and M.P. Dunne (2017), 'The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis', *Lancet Public Health*, 2, e356-66.

Jordan, B. and A. Kennedy (2019), *Changing the Life Trajectories of Australia's Most Vulnerable Children – The Early Years Education (EYEP) model*; https://fbe.unimelb.edu.au/__data/assets/pdf_file/0008/3059297/EYERP-Report-3-web.pdf

Jordan, B., Y. Tseng, N. Coombs, A. Kennedy and J. Borland (2014), 'Improving lifetime trajectories for vulnerable young children and families living with significant stress and social disadvantage: The Early Years Education Program', *BMC Public Health*, 14(1), 965 (10 pages).

Karoly, L., M.R. Kilburn and J.S. Cannon (2005), *Early Childhood Interventions: Proven Results, Future Promise* (Santa Monica, California: Rand Corporation).

Kautz, T., J.J. Heckman, R. Diris, B. ter Weel and L. Borghans (2014), 'Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success', National Bureau of Economic Research, Working Paper no.20749.

Kessler, R.C., G. Andrews, L.J. Colpe, E. Hiripi, D.K. Mroczek, S.L. Normand, E.E. Walters and A.M. Zaslavsky (2002), 'Short screening scales to monitor population prevalences and trends in non-specific psychological distress', *Psychological Medicine*, 32, 959–76.

Knudsen, E.I, J.J. Heckman, J.L. Cameron and J.P. Shonkoff (2006), 'Economic, neurobiological and behavioral perspectives on building America's workforce', *Proceedings of the National Academy of Sciences*, 103(27), 10155-62.

Koenen, K.C., T.E. Moffitt, A. Caspi, A. Taylor and S. Purcell (2003), 'Domestic violence is associated with environmental suppression of IQ in young children', *Development and Psychopathology*, 15(2), 297-311.

LeBuffe, P. and J. Naglieri (2012), *Devereux Early Childhood Assessment for Preschoolers* (Lewisville, NC: Kaplan Early Learning Company).

Mackrain, G. M. LeBuffe and G. Powell (2007), *Devereux Early Childhood Assessment for Infants and Toddlers: User's Guide* (Lewisville, NC: Kaplan Early Learning Company).

Magnuson, K., R. Kelchen, G. Duncan, H. Schindler, H. Shager and H. Yoshikawa (2016), 'Do the effects of early childhood education programs differ by gender? A meta-analysis', *Early Childhood Research Quarterly*, 36, 521-36.

Manski, C. (2013), *Public Policy in an Uncertain World* (Harvard University Press).

Nicholson, J.M., N. Lucas, D. Berthelsen and M. Wake (2012), 'Socioeconomic inequality profiles in physical and developmental health from 0-7 years: Australian national study', *Journal of Epidemiology and Community Health*, 66(1), 81-7.

Perry, B. (2002), 'Childhood experience and the expression of genetic potential: What childhood neglect tells us about nature and nurture', *Brain and Mind*, 3(1), 79-100.

Productivity Commission (2014), *Childcare and Early Childhood Learning*, Inquiry Report No.73 (Canberra).

Raver, C.C., S.M. Jones, C. Li-Grining, F. Zhai, K. Bub and E. Pressler (2011), 'CSRP's impact on low-income pre-schoolers pre-academic skills: Self-regulation as a mediating mechanism', *Child Development*, 82(1), 362-78.

Rothman, E.F., E.M. Edwards, T. Heeren and R.W. Hingson (2008), 'Adverse childhood experiences predict earlier age of drinking onset: Results from a representative US sample of current or former drinkers', *Pediatrics*, 122(2), e298-e304.

Schweinhart, L.J., J. Montie, Z. Xiang, W.S. Barnett, C.R. Belfield and M. Nores (2005), *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40* (Ypsilanti, MI: High Scope Press).

Shonkoff, J.P. (2011), 'Protecting brains, not simply stimulating minds', *Science*, 333(6045), 982-83.

Shonkoff, J.P. (2012), 'Leveraging the biology of adversity to address the roots of disparities in health and development', *Proceedings of the National Academy of Sciences*, 109(Supplement 2), 17302-07.

Stuart, E.A. (2010), 'Matching methods for causal inference: A review and a look forward', *Statistical Science*, 25(1), 1-21.

St. Pierre, R.G., A.E. Riciutti and T.A. Rimdzius (2005), 'Effects of a family literacy program on low-literate children and their parents: Findings from an evaluation of the Even Start Literacy Program', *Developmental Psychology*, 41(6), 953-70.

Tapper, A. and J. Phillimore (2012), 'Prevention-based approaches to social policy: The case of early childhood development', *Evidence Base*, 2012(2), 1-20.

Tough, P. (2016), *Helping Children Succeed: What works and why*, (Boston, MA: Houghton Mifflin Harcourt).

Tseng, Y., B. Jordan, J. Borland, T. Clancy, N. Coombs, K. Cotter, A. Hill and A. Kennedy (2017), *Changing the Life Trajectories of Australia's Most Vulnerable Children – Report no.1: Participants in the Trial of the Early Years Education Program*; accessed at: https://fbe.unimelb.edu.au/__data/assets/pdf_file/0008/2548808/background-report-final-web.pdf

Tseng, Y., B. Jordan, J. Borland, T. Clancy, N. Coombs, K. Cotter, A. Hill and A. Kennedy (2018), *Changing the Life Trajectories of Australia's Most Vulnerable Children – Report no.2: The first twelve months in the Early Years Education Program: An initial assessment of the impact on children and their primary caregivers*; accessed at: https://fbe.unimelb.edu.au/__data/assets/pdf_file/0005/2781248/EYERP-Report-2-web.pdf

Turnbull, A.P., V. Turbville and H.R. Turnbull (2000), 'Evolution of family-professional relationships: Collective empowerment for the early 21st century'. In J.P. Shonkoff and S.J. Meisels (eds.) *Handbook of Early Childhood Interventions* (2nd edition) (Cambridge UK: Cambridge University Press).

United States Department of Health and Human Services (2010), *Head Start Impact Study Final Report* (Washington DC).

Wechsler, D. (2002), *Weschler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI III)* (San Antonio Texas: NCS Pearson).

Appendix 1

Staff involved in delivery of the Early Years Education Program and the research trial

Children's Protection Society/KidsFirst Presidents

Alice Hill	2005–2008
Tim Mulvany	2008–2011
Alice Hill	2011–2013
Jane Munro	2013–2016
Bernard Murphy	2016–2018
Sandy Forbes	2018–present

Children's Protection Society/KidsFirst CEOs

Bernadette Burchell	2006–2011
Dave Glazebrook	2011–2012
Aileen Ashford	2013–present

Assistant to Research/Governance Committee

Margaret Farquharson

Management/team leaders/administration/ infant mental health

Janet Williams Smith	Shannan Mudie
Natalie Bou-Ghosn	Monica Robertson
Aisha Bal	Madeleine Saffigna
Liza Farquhar	Liz Dullard
Joanne Kitto	Dianne Camilleri
Diana Pellegrino	Debra Parker
Nichola Coombs	Cath McPhee

Educators

Sonia Shard	Marylin Ellis
Val Farmer	Pnita Holthouse
Jenny Voogt	Jennifer Lovrek
Barbara Lacey	Donna Kavanagh
Natalie Boardman/O'Dath	Helen Brand
Sandra Athanasoupolos	Chiara Perri
Nerissa Linklater	Lisa McKibbin
Erin Maree Sharp	Jane Cecelia William
Catherine Quirk	Sarah Meldrum
Jacquelyn Clark	Tina Howard
Jaymi-Lee Warren	Robyn Ball
Lisa Barbaro	Sohayla Asari

Cooks

Edwina Pleming	Marcela Ramos
Lea Bautista	Patrick Carmody
Anne Flack	Gabbie Mantini

Researchers

Nichola Coombs	Jane Sheehan
Megan Clark	Penny Hartmann
Andrew Bevitt	Steph Brophy
Jonathan Reyes	Xuan Vu
Mael Guillou	Leng Lee
Tamera Clancy	Lauren McCabe
Kerry Ware	

Appendix 2

Victorian Department of Human Services 2007 Best Interest Case Practice Model - List of risk factors to healthy child development

Child and family risk factors

- family violence, current or past
- mental health issue or disorder, current or past (including self-harm or suicide attempts)
- alcohol/substance abuse, current or past, addictive behaviours
- disability or complex medical needs eg. intellectual or physical disability, acquired brain injury
- newborn, prematurity, low birth weight, chemically dependent, foetal alcohol syndrome, feeding/sleeping/settling difficulties, prolonged and frequent crying
- unsafe sleeping practices for infants eg. side or tummy sleeping, ill-fitting mattress, cot cluttered with pillows, bedding or soft toys which can cover an infant's face, co-sleeping with sibling or parent who is on medication, drugs/alcohol or smokes, using other unsafe sleeping place such as a couch or exposure to cigarette smoke
- disorganised or insecure attachment relationship (child does not seek comfort or affection from caregivers when in need)
- developmental delay
- history of neglect or abuse, state care, child death or placement of child or siblings
- separations from parents or caregivers
- parent, partner, close relative or sibling with a history of assault, prostitution or sexual offences
- experience of intergenerational abuse/trauma
- compounded or unresolved experiences of loss and grief
- chaotic household/lifestyle/problem gambling
- poverty, financial hardship, unemployment
- social isolation (family, extended family, community and cultural isolation)
- inadequate housing/transience/homelessness
- lack of stimulation and learning opportunities, disengagement from school, truancing
- inattention to developmental health needs/ poor diet
- disadvantaged community
- racism
- recent refugee experience

Parent risk factors

- parent/carer under 20 years or under 20 years at birth of first child
- lack of willingness or ability to prioritise child's needs above own
- rejection or scapegoating of child
- harsh, inconsistent discipline, neglect or abuse
- inadequate supervision of child or emotional enmeshment
- single parenting/multiple partners
- inadequate antenatal care or alcohol/substance abuse during pregnancy

Wider factors that influence positive outcomes

- sense of belonging to home, family, community and a strong cultural identity
- pro-social peer group

Appendix 3

Details of covariates in statistical models

The full details of the sets of covariates included for each outcome and each specification is below:

Covariates by outcome and specification (DV= dummy variable):		
Outcome	**Model 1**	**Model 2**
IQ; Language	Gender; Age at 24 months IQ assessment; Duration between IQ assessments at entry to trial and 24 months; DV for carer age 25–34; DV for carer age 35+; DV for whether carer has post-school qualification; DV for K6 category Medium; DV for K6 category High; DV for whether primary caregiver is immigrant; DV for whether language other than English is main language spoken at home; DV for whether both parents present at consent meeting; IQ score at time of entry to trial; Language score at time of entry to trial.	Model 1 plus DV for multiple children in family in trial from time of referral; DV for whether alcohol/ substance abuse was a risk factor at referral; DV for whether mental health issues was a risk factor at referral; DV for whether family violence (current or past) was a risk factor at referral.
Protective factors; Social-emotional development	Same as for IQ/Language – Except that replace age at 24 months IQ assessment and the duration between IQ assessments at time of entry to trial and at 24 months with age at 24 months assessments for protective factors and social and emotional development.	Model 1 plus DV for multiple children in family in trial from time of referral; DV for whether alcohol/ substance abuse was a risk factor at referral; DV for whether mental health issues was a risk factor at referral; DV for whether family violence (current or past) was a risk factor at referral.
Parenting daily hassles	Same as for IQ/Language – Except that replace age at 24 months IQ assessment and the duration between IQ assessments at time of entry and at 24 months with age at 24 months assessments for parenting daily hassles.	Model 1 plus DV for multiple children in family in trial from time of referral; DV for whether alcohol/ substance abuse was a risk factor at referral; DV for whether mental health issues was a risk factor at referral; DV for whether family violence (current or past) was a risk factor at referral.
Parent psychological distress	Same as for IQ/Language – Except that replace age at 24 months IQ assessment and the duration between IQ assessments at time of entry and at 24 months with age at 24 months for K6 assessment.	Model 1 plus DV for multiple children in family in trial from time of referral; DV for whether alcohol/ substance abuse was a risk factor at referral; DV for whether mental health issues was a risk factor at referral; DV for whether family violence (current or past) was a risk factor at referral.
Home environment	Same as for IQ/Language – Except that replace age at 24 months IQ assessment and the duration between IQ assessments at time of entry and at 24 months with age at 24 months for home environment.	Model 1 plus DV for multiple children in family in trial from time of referral; DV for whether alcohol/ substance abuse was a risk factor at referral; DV for whether mental health issues was a risk factor at referral; DV for whether family violence (current or past) was a risk factor at referral.

Appendix 4

Statistical significance

1] p-value

The p-value can range from zero to one; and lower p-values make it more likely that there is a significant difference in an outcome between the intervention and control groups. As an example, if the p-value for the estimated difference in an outcome between the groups equals 0.10, this means that there is only a 10 per cent chance that there is not a difference in that outcome between the intervention and control groups

2] Type of test

We report what are known as one-tailed and two-tailed tests of significance. A one-tailed test is appropriate if it is considered that the only possible effect of participation in EYEP could have been to cause a zero or positive impact on outcomes. A two-tailed test is relevant if it is also believed that participation in EYEP could have caused a negative impact on outcomes. It is standard in analysis of early years programs to put most weight on one-tail tests. This is supported by the fact that where significant impacts have been estimated for early years programs in previous research, those impacts overwhelmingly have been positive (see Cannon et al., 2017, p. 64). Generally, a 5 or 10 per cent level of significance is used as the threshold for concluding that it is possible to reject that the impact size is significantly different from zero. Hence, we regard an estimate of the impact of EYEP as significant when the p-value is equal to or less than 0.05 (five per cent level) or 0.1 (ten per cent level).

3] Permutation test

Applying the same block randomisation approach and sibling assignment rule as for the original assignment of trial participants to the intervention group and the control group, the sample of participants in the EYEP trial is reassigned 10,000 times between those groups. For each reassignment, the difference in the group mean of each variable can be calculated. From the process of repeated reassignment, a distribution of differences in means for each variable is generated. That distribution is then applied to undertake a two-tailed test for whether there is a significant difference in the mean values of the variable between the intervention group and the control group. For a general reference, see Good (2005).

4] Bounds analysis

To estimate the upper bound of the impact of EYEP on children's social-emotional development it is assumed that for children from whom data was collected at time of entry to the trial but not at 24 months: (i) A child enrolled in EYEP would not be in the clinical range at 24 months; and (ii) A child in the control group would be in the clinical range at 24 months. To estimate the lower bound of the impact of EYEP on children's social and emotional development, the opposite assumptions are made. See Manski (2013, pp. 57–58) for a discussion of the method of bounding impact sizes.

Appendix 5

Extra results on impacts of EYEP after 24 months

Table 5.1: Impact on IQ and Language of enrolment in EYEP for 24 months – Sensitivity analysis

	Outcome	Method	EYEP mean (24 months)	EYEP impact	1-tail p-value	2-tail p-value	Observations (EYEP/Control)
(1)	IQ	Matching – Model 1	99.560	5.952	0.087	0.174	50/43
(2)	IQ	Matching – Model 2	99.560	5.351	0.149	0.298	50/43
(3)	IQ	Regression on matched sample – Model 1	99.560	5.253	0.072	0.144	50/43
(4)	IQ	Regression on matched sample – Model 2	99.560	6.285	0.053	0.106	50/43
(5)	Change in IQ from entry to 24 months	Matching – Model 1	7.660	6.440	0.064	0.128	50/43
(6)	Change in IQ from entry to 24 months	Matching – Model 2	7.660	8.002	0.085	0.170	50/32
(7)	IQ	Regression – Model 1 – With control for type of test	99.560	5.838	0.034	0.063	50/43
(8)	IQ	Regression – Model 2 – With control for type of test	99.560	7.648	0.009	0.017	50/43
(9)	Language	Matching – Model 1	96.160	3.723	0.195	0.385	50/43
(10)	Language	Matching – Model 2	96.160	4.350	0.191	0.382	50/43
(11)	Language	Regression on matched sample – Model 1	96.160	3.537	0.161	0.322	50/43
(12)	Language	Regression on matched sample – Model 2	96.160	5.790	0.052	0.104	50/43
(13)	Change in Language from entry to 24 months	Matching – Model 1	7.580	3.936	0.147	0.294	50/43
(14)	Change in Language from entry to 24 months	Matching – Model 2	7.580	6.584	0.101	0.202	50/43
(15)	Language	Regression – Model 1 – With control for type of test	96.16	3.448	0.177	0.343	50/43
(16)	Language	Regression – Model 2 – With control for type of test	96.16	5.501	0.058	0.112	50/43

Table 5.2: Impact on Protective factors, Social and emotional development and K6 of primary caregivers of enrolment in EYEP for 24 months – Sensitivity analysis

	Outcome	Method	EYEP mean (24 months)	EYEP impact	1-tail p-value	2-tail p-value	Observations (EYEP/Control)
(1)	Protective factors	Matching – Model 1	44.94	3.131	0.067	0.134	46/41
(2)	Protective factors	Matching – Model 2	44.94	4.097	0.075	0.150	46/41
(3)	Protective factors	Regression on matched sample – Model 1	44.94	3.398	0.044	0.088	46/41
(4)	Protective factors	Regression on matched sample – Model 2	44.94	4.277	0.052	0.104	46/41
(5)	Social-emotional	Matching – Model 1	12.8	-34.7	0.002	0.004	47/39
(6)	Social-emotional	Matching – Model 2	12.8	-33.6	0.006	0.012	47/39
(7)	Social-emotional	Regression on matched sample – Model 1	12.8	-33.6	0.004	0.008	47/39
(8)	Social-emotional	Regression on matched sample – Model 2	12.8	-29.8	0.008	0.016	47/39
(9)	Social-emotional	Regression – Model 1 – With control for type of test	12.8	-31.7	0.002	0.003	47/39
(10)	Social-emotional	Regression – Model 2 – With control for type of test	12.8	-29.5	0.004	0.006	47/39
(11)	Social-emotional	Bounds approach – Model 1 - Lower bound	24.1	-7.7	0.242	0.484	54/55
(12)	Social-emotional	Bounds approach – Model 1 - Upper bound	11.1	-50.6	<0.001	<0.001	54/55
(13)	K6	Matching – Model 1	13.91	-1.210	0.204	0.408	46/45
(14)	K6	Matching – Model 2	13.91	-1.716	0.244	0.488	46/45
(15)	K6	Regression on matched sample – Model 1	13.91	-1.381	0.202	0.404	46/45
(16)	K6	Regression on matched sample – Model 2	13.91	-1.713	0.260	0.519	46/45

Table 5.3: Impact of enrolment in EYEP at 12 and 24 months – Regression and matching estimates - Children who attended for at least 60 days

Outcome		Regression		Matching		
		EYEP impact	1-tail p-value	EYEP impact	1-tail p-value	No. of observations (EYEP/Control)
IQ – Model 1	12 months	3.059	0.131	5.009	0.098	50/36
	24 months	3.986	0.127	3.975	0.204	50/36
IQ – Model 2	12 months	4.194	0.068	5.694	0.121	50/36
	24 months	6.045	0.035	4.138	0.242	50/36
Language – Model 1	12 months	-1.444	0.306	-1.029	0.326	50/36
	24 months	1.751	0.327	0.903	0.451	50/36
Language – Model 2	12 months	-0.002	0.489	0.697	0.436	50/36
	24 months	4.138	0.129	3.063	0.304	50/36
Protective factors – Model 1	12 months	1.093	0.324	-0.702	0.453	43/32
	24 months	3.442	0.058	3.560	0.055	43/32
Protective factors – Model 2	12 months	1.892	0.240	1.184	0.373	43/32
	24 months	3.590	0.067	4.549	0.051	43/32
Social and emotional – Model 1	12 months	-20.0	0.084	-13.9	0.165	42/32
	24 months	-29.0	0.006	-32.9	0.006	42/32
Social and emotional – Model 2	12 months	-35.2	0.011	-16.1	0.155	42/32
	24 months	-36.3	0.003	-33.2	0.016	42/32
Psychological distress – Model 1	12 months	-0.533	0.336	0.108	0.241	42/36
	24 months	-1.429	0.150	-0.098	0.255	42/36
Psychological distress – Model 2	12 months	0.009	0.497	0.172	0.156	42/36
	24 months	-1.692	0.126	-0.209	0.132	42/36
Parenting daily hassles – Frequency – Model 1	12 months	3.697	0.065	3.083	0.053	39/31
	24 months	1.477	0.270	0.449	0.492	39/31
Parenting daily hassles – Frequency – Model 2	12 months	4.609	0.044	3.952	0.092	39/31
	24 months	0.707	0.393	0.179	0.351	39/31
Parenting daily hassles – Intensity – Model 1	12 months	4.811	0.089	3.939	0.099	38/31
	24 months	0.419	0.457	-1.408	0.260	38/31
Parenting daily hassles – Intensity – Model 2	12 months	5.494	0.078	3.250	0.210	38/31
	24 months	1.389	0.370	-2.533	0.155	38/31
HOME – Model 1	12 months	0.894	0.409			25/26
	24 months	0.482	0.465			25/26
HOME – Model 2	12 months	2.694	0.261			25/26
	24 months	1.642	0.388			25/26

Changing the Life Trajectories of Australia's Most Vulnerable Children

Report no. 1	June 2017	Participants in the Trial of the Early Years Education Program
Report no. 2	March 2018	The first twelve months in the Early Years Education Program: An initial assessment of the impact on children and their primary caregivers
Report no. 3	May 2019	The Early Years Education Program (EYEP) Model
Report no. 4	May 2019	24 months in the Early Years Education Program: Assessment of the impact on children and their primary caregivers

Changing the Life Trajectories of Australia's Most Vulnerable Children

Report No. 4
24 months in the Early Years Education Program: Assessment of the impact on children and their primary caregivers