Changing the Life Trajectories of Australia's Most Vulnerable Children

Report No. 5

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

October 2022

THE UNIVERSITY OF MELBOURNE

MELBOURNE INSTITUTE
APPLIED ECONOMIC & SOCIAL RESEARCH

KIDS FIRST

36 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
Jeff Borland	Department of Economics, University of Melbourne
Megan Clark	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	Antipodean Family Foundation
Anne Kennedy	Melbourne Graduate School of Education, University of Melbourne
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

We acknowledge the role of the Board of Kids First Australia (previously the Children's Protection Society) in initiating the trial of the Early Years Education Program; the contribution of Kids First Australia, the Board and its staff in implementing the program; and the contribution by Kids First Australia in funding the trial. We owe a huge debt 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. We have appreciated the opportunity to work in partnership with, and the support we have received from, the Board of Kids First and its senior management. A complete list of staff from the University of Melbourne and Kids First/Children's Protection Society who have been involved with implementation 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 Kids First, 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 Department of Health and Department of Families, Fairness and Housing), 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, Vic Health, the Antipodean Family Foundation, the Murphy-McNicol Family, the Crawford Foundation, the Brougham Family Trust, the William Buckland Foundation, VivCourt Trading, the Faculty of Business and Economics at the University of Melbourne and Australian Research Council Linkage Grant LP140100897. This study uses part of questionnaires and unit-record data 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 of the Early Years Education Program (EYEP) on vulnerable and at-risk children and their primary caregivers after 36 months in the program. Australian children who live with multiple risk factors in their early years have compromised cognitive and non-cognitive skills; as demonstrated for example in our initial report (Tseng et al., 2017). In this report we document how participation for 36 months in EYEP can effectively reverse harms and restore the learning trajectories of Australia's most vulnerable children. Children's cognitive and non-cognitive development is substantially improved by participating in EYEP, continuing the broad positive impact of the program found after 24 months (Tseng et al., 2019).

Large impacts of EYEP are found on both outcomes relating to children's cognitive development, IQ and language skills. A significant impact on children's IQ existed after 12 months, which persisted throughout the trial. A significant impact on language development has mainly become evident at 36 months. Children's IQ and language development are sufficiently improved by EYEP that the program's objective to make participants developmentally equal to their peers is achieved.

A large impact of EYEP is found on children's social and emotional development, one of the two outcomes relating to non-cognitive development. This impact emerged at 24 months and is again found at 36 months. The impact on the other non-cognitive outcome, children's resilience, is less significant at 36 months than at 24 months. Taken across the whole trial we conclude that EYEP has a small positive average impact on resilience but with a relatively high degree of variability across children.

Evidence is not found, however, for an impact of EYEP on the psychological well-being of primary caregivers, or the level of stress they experience.

EYEP

EYEP is an early years education and care 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. EYEP is intended for implementation in purposefully designed and well-resourced small childcare centres (up to 42 children). A highly experienced and qualified multidisciplinary staff team collaborates with families and other professionals to provide very young children with a program designed to repair harms, reduce stress and to promote learning. Children who participate in EYEP are offered three years of education and care (50 weeks per year, five hours per day each weekday) at no cost to their families.

The ultimate objective of EYEP is to ensure that at-risk and vulnerable children realise their full potential and transition into school as confident, capable learners, developmentally and educationally equal to their peers.

The EYEP trial

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 who received 'usual care' (a mix of parental and guardian care as well as education and care sourced by caregivers from other childcare centres or kindergartens). The impact of EYEP on children and their primary caregivers is estimated by comparing outcomes for the intervention and control 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 referral to the trial, assessed as having two or more risk factors as defined in the Victorian 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.

Participants in the EYEP trial were highly disadvantaged on multiple dimensions when they entered the trial. Many aspects of their development - IQ, language skills, motor skills and adaptive behaviour – were compromised. Their primary caregivers had low levels of labour force engagement and family income, were likely to have recently experienced stressful events and had high levels of psychological distress.

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. Outcomes relating to children's IQ and language skills and their resilience and social-emotional development are therefore examined. Possible impacts of EYEP on primary caregivers are also evaluated using outcomes relating to perceptions of the level of stress they are experiencing.

Sample size

A total of 145 children were recruited to the EYEP trial. These children come from 99 families. There are 72 children 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 36 months, on which the analysis in this report primarily relies, obtained responses from 95 children and their primary caregivers. In addition, it was not possible to collect a complete set of data on outcomes for all children and primary caregivers. Hence, the analysis of the impact of EYEP on outcomes at 36 months is based on samples that usually consist of 75 to 85 children or primary caregivers, depending on the outcome.

Estimation method

The impact of EYEP is estimated using regression adjusted matching: Children in the intervention group are matched to children in the control group with similar characteristics; and a regression model is used to estimate the impact of EYEP for this matched sample.

Initial random assignment of participants in the EYEP trial did achieve balance in characteristics of the intervention and control groups. However, due to subsequent attrition of participants from the trial, some characteristics likely to affect children's outcomes at 36 months have become unbalanced in the sample available at that time. Hence it is necessary to use an estimation method that allows these differences in characteristics between intervention and control groups to be controlled for.

The impact of EYEP is estimated for children who attended the program for at least 60 days over 36 months (out of a possible total of approximately 720 days). We choose to restrict 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.

Main findings

Large and statistically significant impacts of EYEP are found on children's IQ and language development. The estimated impacts are about one-half of a standard deviation. For IQ this is twice as large as the average impact estimated from early years demonstration programs in the United States. The impact of EYEP on children's IQ was mainly achieved in the first 12 months of the program, with some evidence of further improvement in the next 24 months. The impact on language grew throughout the trial and becomes significant at 36 months. Much larger and more significant impacts on IQ and language development are found for children with low IQ and language scores at the time of entry to the trial. This suggests that the impact of EYEP primarily has been compensatory.

The estimated impact on children's protective factors related to resilience is relatively small and not statistically significant. However, throughout the trial a small positive impact on resilience has been sustained, which was highly significant at 24 months. Hence, we conclude there is likely to be a small average impact of EYEP on resilience, but with relatively high variability between children.

Outcome measures for the impact of EYEP on children's social-emotional development generally show large and statistically significant impacts of the program. Overall, it seems reasonable to conclude that EYEP had improved children's social-emotional development by 24 months and that this impact is sustained at 36 months.

The impacts of EYEP on children's development appear to be explained by the attributes and quality of EYEP compared to the formal early years education and care received by the control group; and by children in the intervention group having received a larger amount of formal early years education and care over the first 24 months of the trial.

1. Introduction

This report presents findings on the impact on children and their primary caregivers after 36 months of enrolment in the Early Years Education Program (EYEP). It follows earlier reports on evaluations of the impact of EYEP after 12 and 24 months (Tseng et al., 2018, 2019).

EYEP is a centre-based early years education and care program targeted at the needs of infants and young children who are exposed to significant family stress and social disadvantage, including being at heightened risk of, or having experienced, abuse and neglect. 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 Kids First, previously the Children's Protection Society (CPS), an independent not-for-profit child and family services organisation which was founded in 1896. CPS, as a large, well-established community services organisation including a family services team with expertise in child protection, was well placed for recruiting children and families along with other Child First and welfare agencies and Maternal and Child Health professionals. The program was designed and implemented by CPS in collaboration with Associate Professor Brigid Jordan and Dr Anne Kennedy.

The EYEP trial was undertaken by a consortium of researchers with support from their institutions and in partnership with Kids First. Funding for the research trial has come from Kids First, 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 the main findings on the impact of EYEP after 36 months. Our interpretation of the main findings is in Section 10.

2. Background to the EYEP trial

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, physiology and social and emotional development. 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 relations with adults that are 'reciprocal and dynamic' and avoiding excessive stress are regarded as essential to healthy development for young children (Center on the Developing Child at Harvard University, 2016b, pp.7-8). By contrast, prolonged exposure to physical, emotional and/or sexual abuse and traumatic experiences early in life are known 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 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 be unable to ever make up the lost ground, or may lack the necessary building blocks and foundation for subsequent learning (Cunha and Heckman, 2007; Heckman, 2008; Heckman and Mosso, 2014). 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).

Dealing with the problem of skill development for children who are exposed to significant family stress is widely agreed to require a new model for education and care compared to mainstream early childhood services. That new model must have as a core element to undo the negative impact of stress on brain development - as Shonkoff (2011, p.982) has put it: 'linking high-quality pedagogy to interventions that prevent, reduce, or mitigate the disruptive effects of toxic stress on the developing brain' – as well as redressing learning deficiencies.

Having a model that can address 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 2019-20 there were 49,646 children aged 4 years or below who were receiving child protection services (Australian Institute of Health and Welfare, 2021, table S2.3). Second, at-risk children in Australia currently seem the group least likely to be able to access early years education and care (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 and 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 influences motivated CPS to create and trial the Early Years Education 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 and HREC 2021-20679-16651-4). At the time it commenced, the EYEP trial was the first RCT of a centre-based early years education and care intervention in Australia (Tapper and Phillimore, 2012).

3. The Early Years Education Program

Overview

EYEP is an innovative Australian multi-disciplinary centre-based early years education and care program for infants and young children. 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. It is a targeted program that responds to the acute needs of children living with significant adversity in the same way that paediatric intensive care in hospitals is designed for children with acute medical problems. The program operates under the Australian National Quality Framework (NQF), the National Quality Standard (NQS) and the *National Early Years Learning Framework*.

EYEP has a dual focus: first, reducing stress and repairing the harms of curtailed cognitive and emotional development; and second, promoting learning within an enhanced early education and care model. The ultimate objective of EYEP is to ensure that at-risk and vulnerable children can enter their first year of formal schooling as confident, capable learners, developmentally and educationally equal to their peers.

EYEP 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).

A full description of EYEP and the underpinnings of its design is presented in Jordan and Kennedy (2019).

Education and care elements of EYEP

The foundation of EYEP is a holistic model of education and care 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 as well as on theories and best practice of early learning and pedagogy.

The EYEP model is designed around a set of core integrated elements which work in synergy to allow children to build their cognitive and emotional capacity for learning. 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 principle of high expectations for every child as a learner underpins the program.

The EYEP program is designed to be implemented in a purposefully designed and well-resourced small childcare centre, for between 36-42 children. Children who participate in EYEP are offered three years of education and care (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 teachers and educators, and an infant mental health clinician consultant and family support practitioner as core members of the staff team. A rigorous relationship-based curriculum informed by trauma and attachment theories and early learning theories is implemented. Structural and process quality elements in the model are well above Australian National Quality Standard requirements including better ratios of adults to children, higher qualification levels for all staff, small group sizes and the enactment of relational 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 objective of care is to redress harms and reduce stress in children so that they can avail themselves of the learning opportunities provided. A primary educator model is adopted. The purpose of the primary educator 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. Close attention is paid to transitions and a strong sense of belonging for every child and family is promoted through culturally responsive practices, non-judgemental approaches and respect for every family and child. The small size of the program supports each family's sense of belonging and the staff team's capacity to actively respond to individual families and children.

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). Teachers and educators plan a curriculum using a balance of child led, play-based approaches, and teacher-led intentional teaching to support each child's learning and development across outcomes in the *Early Years Learning Framework*. The infant mental health component means that emotional and behavioural issues, for example trauma responses, are taken into account when designing educational strategies for individual children. There is a focus on individual and small group learning experiences rather than whole group activities. The individualised and differentiated curriculum provision supports meaningful engagement in learning for every child. Teachers and educators engage in goal setting with families every twelve weeks and individual learning plans are developed accordingly. Play-based learning approaches are used and adapted to promote, support and extend children's identified interests. Teachers and educators understand the threshold for learning capacity of play-based approaches.

A multi-disciplinary model

An innovative feature of EYEP is a multi-disciplinary model with an infant mental health clinician consultant and a family support practitioner in-house as integral team members, and an early childhood curriculum consultant. 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.

Staffing

The EYEP model requires that only full-time teachers and educators are employed. This is intended to allow them to develop strong and consistent relationships with children for whom they are responsible. The EYEP model uses a range of strategies to support staff retention, wellbeing and engagement. Newly recruited teachers and educators receive introductory professional development based on attachment theory and key infant mental health concepts. Each teacher and educator receives weekly, formal, and individual reflective supervision from a member of the EYEP leadership team and quality individual and group based professional development and learning. Monthly staff meetings occur for all staff with a multi-disciplinary focus on reflection, professional learning, problem solving and decision making.

A critical part of EYEP is that teachers and educators have ten hours a week out of the room for planning, documenting and assessing children's learning, discussions with members of the leadership team, 12 weekly goal setting and review meetings with families, professional networking and other tasks related to supporting families and children. The infant mental health consultant provides group consultation for the educators in each of the children's rooms every fortnight.

Ensuring model fidelity

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 EYEP research team, 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.

Connection with parents

EYEP is a child-focused intervention, not intended to directly affect parenting behaviours. Development of sustained, meaningful partnerships with parents is however a core principle of the program. The model involves active engagement 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.

The orientation and attendance plan for a child enables the parent or primary caregiver to gradually build a trusting relationship with their child's teacher or educator. Goal setting meetings between parents and family support/child protection workers with the child's teacher or educator take place every twelve weeks. Individualised learning plans are developed and reviewed at these meetings.

4. The EYEP trial and participants

Eligibility criteria

Eligibility criteria 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 at the time of referral to the trial 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. Referrals of potential EYEP participants were made by caseworkers from clients of child and family services including (but not exclusively from) Child FIRST and Child Protection within the Victorian Department of Health and Human Services.

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 risk factors is included as Appendix 2.

Characteristics of EYEP trial participants

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.

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, most children had many more than the minimum number of two risk factors. About 30 percent of children had two or three risk factors, 35 percent had four or five risk factors, and 35 percent 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, the children exhibited compromised development in the areas of IQ, language skills, motor skills, and adaptive behaviour. They also 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). Table 2 (Panel A) presents summary information on children in the EYEP trial: on their development at the time of entry to the trial and on birth weights.

Third, the children were living in difficult family circumstances at their time of entry to the trial. Their primary caregivers were more likely to be young parents, to have fewer financial resources and to not be participating in the labour force. The number of stressful life events

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

	Number	Percent
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.

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

being experienced by the caregivers was extraordinarily high, and many had severe levels of psychological distress. Summary information on primary caregivers of children in the EYEP trial is presented 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 who are in the intervention group and 73 in the control group, and respectively 50 and 49 families in these groups.

Assignment was family-based, to ensure siblings were assigned to the same group. In addition, when a new sibling of a participant was born during the recruitment phase of the trial, and eligibility conditions for participation were met, the primary caregiver was invited to consent for the new infant to participate in the trial in the same group to which their older sibling had been assigned.

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

The control group received 'usual care', a mix of parental and guardian care as well as education and care 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 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 to provide families with a brief summary report of their child's development after their annual research assessment. This included any recommendation for the family to seek specialist assistance which the researchers believed was necessary. It is also possible that the data collection process, which asked primary caregivers about their parenting practices, exposed them to new positive behaviours that they could adopt. Hence, outcomes for the control group may be better than if they had not been involved in the trial.

Data collection

Data were 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 on outcome measures at 36 months were: (i) Collected directly by researchers (children's cognitive development); or (ii) From parent responses to self-report assessment instruments administered orally or in written form by the researchers (children's non-cognitive development and well-being of primary caregivers). Five researchers were involved in data collection, with about 80 per cent of the IQ data (for which we have information on assessor identity) being collected by two researchers.

5. Overview of outcome variables

The purpose of the EYEP trial is to test whether the program achieves the goals of building children's cognitive and non-cognitive development (Jordan et al., 2014, p.3). Hence, we investigate outcomes relating to these aspects of children's development (Craig et al., 2008).

Possible impacts of EYEP on primary caregivers are also evaluated using outcomes relating to perceptions of the level of stress they are experiencing and psychological well-being.

The outcomes and associated measures that we examine are listed 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' positive findings.

The set of outcomes and measures reported on at 36 months is the same as in previous reports, with two exceptions. First, we report an expanded set of estimates of the impact of EYEP on social-emotional development. In evaluating social-emotional development at 12 and 24 months, the mix of ages of children meant that two different measures needed to be used. Integrating those measures required us to examine the impact of EYEP on the percentage of children with social-emotional development below population norm thresholds (for example, bottom 10 per cent). At 36 months, however, almost all children have had data collected on their social-emotional development using the same measure. Hence, we are now able to report on the impact of EYEP on that measure – that is, on the average impact of EYEP on a measure which ranges over a continuous scale. Second, we have not been able to estimate the impact on the HOME measure at 36 months due to insufficient sample size.

In what follows, we provide a brief overview of the main measures and how they are used to estimate the impact of EYEP at 36 months. More details on the measures are provided in Tseng et al. (2018, Appendix 3).

Table 3: Outcomes and measures of the impact of EYEP

	Outcome	Measure	Impact estimated on
1	Child development - IQ	Bayley Scales of Infant and Toddler Development III (BSID); Wechsler Preschool and Primary Scale of Intelligence (WPPSI)	Points [Scale with population mean = 100; SD = 15]
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	Points [Scale with population mean = 100; SD = 15]
3	Child development – Protective factors related to resilience (initiative, self-regulation, attachment/relationships, behavioural concerns)	Devereux Early Childhood Assessment Program (DECA)	T score [Scale with population mean = 50; SD = 10]
4	Child social and emotional development	Brief Infant Toddler Social Emotional Assessment (BITSEA); Child Behaviour Checklist (CBCL)	a] Per cent of children with development in bottom 10/25 per cent of norm population; b] CBCL – T-score [Scale from 28 to 100]
5a	Parent psychological distress	K6	Points [Scale from 6 to 30]
5b	Parenting stress	The Parenting Daily Hassles Scale	a] Frequency [Scale from 0 to 80] b] Intensity [Scale from 0 to 100]

- **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).

The Bayley Scales and WPPSI 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 36 months, the scores from each test are treated as being directly comparable.

- **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 1 month to less than 18 months, DECA-T is used for toddlers from ages 18 to less than 36 months, DECA-P2 is used for children aged 3 to 5 years and DESSA for children in school. 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 and 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. 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 5 years. Raw scores on the CBCL are converted to a T-score with mean of 50 and SD of 10. A percentile ranking based on age-based population norms can also be assigned to each score (although scores below the 50th percentile are aggregated). Separate sub-scores are calculated relating to internalizing behaviour (problems mainly within the self) and externalizing behaviour (problems that mainly involve conflict with other people and their expectations for the child) (Achenbach and Rescorla, 2000, p.13).

As in previous reports, the BITSEA and CBCL measures are integrated to obtain a consistent measure of problems with children's social and emotional development. This is done by classifying children based on whether their score for social and emotional development was below specified percentile levels in the population norm distribution for each measure; and then calculating the percentage of children whose score was below that percentile level. The percentages of children with scores below the 10th and 25th percentiles are again used for estimating the impact of EYEP. As well, the percentage

of children with scores below the 17th percentile is used as an outcome. Children below the 10th percentile are classified as being in the clinical range for the CBCL measure. Children below the 17th percentile in the CBCL are classified as borderline for the clinical range; and children below the 25th percentile in BITSEA are classified as being in the Possible Problem range.

To further consider the effect of EYEP on social and emotional development at 36 months, an estimated impact on children's CBCL score is also reported. This approach is possible because all but one child were assessed with the CBCL measure at 36 months. The advantage of using the CBCL is that it provides an estimate of the average impact on a continuous scale. By contrast, where the outcome is the percentage of children below a specified percentile in the population norm distribution, the only information about a child's score that is used is whether the score is above or below that percentile level. Estimates of the impact of EYEP based on categorical outcomes are likely to exhibit greater variability and sensitivity to small changes in assessed scores compared to using a continuous scale.

- **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.

6. Description of data and analysis of randomisation and attrition

Background

Analysis of the impact of EYEP in this report is mainly based on data collected from children and their primary caregivers after 36 months in the trial. 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 children's development and demographic characteristics were collected prior to commencement in EYEP for 134 children. By the time of the data collection at 36 months, there had been further attrition, resulting in data being available for a maximum of 95 children and their primary caregivers. Details of the evolution of the sample sizes for the intervention and control groups are shown in Figure 1 (see next page).

For children and primary caregivers from whom data were collected at 36 months, there is also some non-response on specific items due to scheduling issues and time constraints. For example, some tests must be completed in a specified time period in order to be valid (defined relative to months since last assessment with that measure). 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).

Timing of data collection

Summary information on the distribution of timing of data collection for the sample of children whose outcomes are studied in this report is shown in Figure 2. 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 36 months.

Data for a majority of children were collected in a timely manner, with IQ assessments taking place for 51.5 percent of children within the six-months window around the three-year anniversary of their entry to the trial (60.4 per cent for the intervention group and 42.6 per cent for the control group). Data collection for most other children took place after the six months window. For children who had started school prior to the 36 months data collection it was decided to combine the 36 months and at-school assessments and for these to take place during the mid-year (July) school holidays, which resulted in longer times to data collection for those children. As well, delays were more likely to occur for children in the control group due to greater difficulties in scheduling data collection for this group, compared to the intervention group, most of whom were attending the EYEP centre.

Method of data collection

Data collection and analysis in this project have been non-blind. 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 36 months data collection

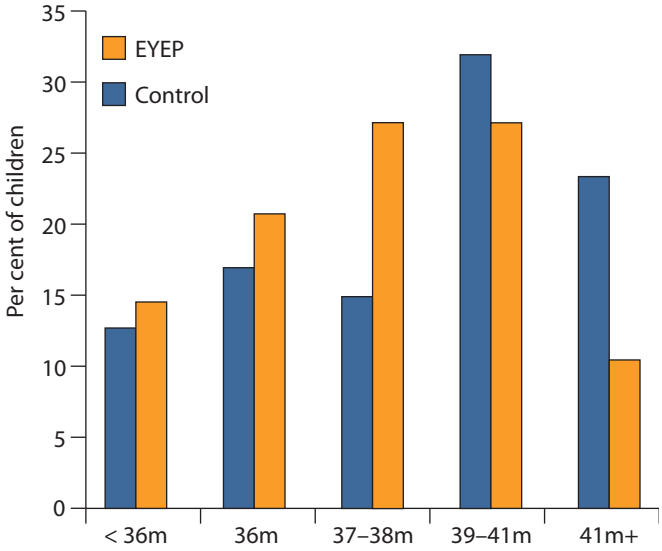
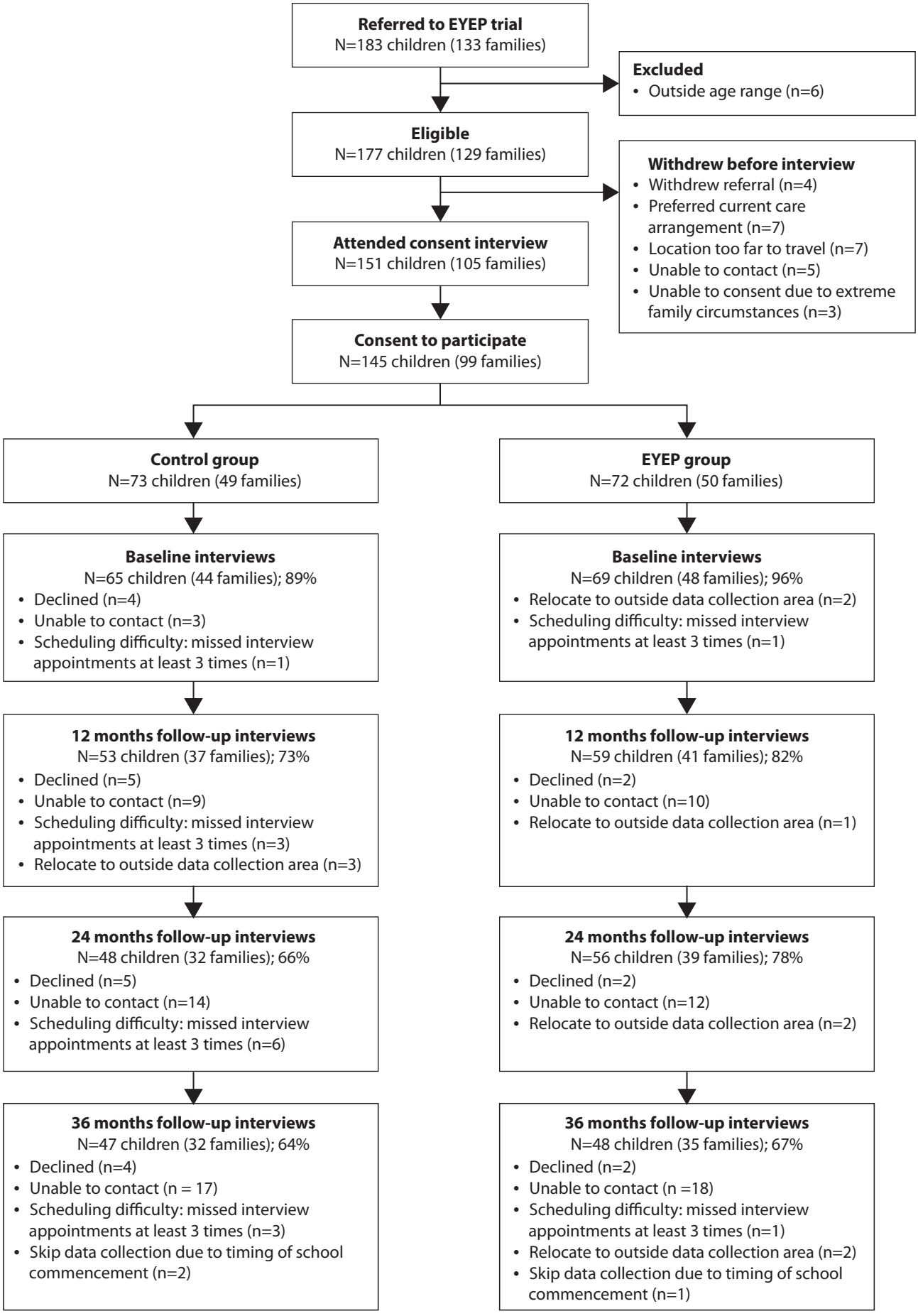


Figure 1: Flow chart of EYEP trial participation and attrition



7. The impact of EYEP – Background

Theory

The estimated impact of EYEP is based on comparisons of outcomes between the intervention and control groups after 36 months. The main difference between the groups during that time is the early years education and care services they receive.

There are several channels by which assignment to EYEP could directly affect children's development compared to being in the control group. First, EYEP affects the attributes and quality of early years education and care received by children. The program's objective of forming secure relationships between children and educators is intended to provide the children with a sense of safety and security to enable them to learn. EYEP also seeks to assist children develop the capacity to regulate their emotions and behaviour, and to mitigate the negative impacts of toxic stress (such as hypervigilance, inability to concentrate and preoccupation) that can impair learning. With these foundations, children attending EYEP can benefit from a high-quality child-focused education program that combines individualised programming with a rigorously developed pedagogy. Second, the quantity of early years education and care received by a child may be increased by being able to attend EYEP. Third, attending EYEP may improve children's nutrition via provision of lunch and snacks.

In addition, EYEP may also have indirect effects on children's development through impacts on parental engagement or by allowing reduced spending on childcare which enables increased spending on other goods or services that benefit children.

Attendance patterns at EYEP

Attendance patterns by the intervention group at EYEP at 36 months are described in Table 4. These attendance data come from EYEP administrative records. From the intervention groups recruited to the trial, 54 children attended for at least 60 days, six for less than 60 days, and 12 did not commence regular attendance. The average total days attended at EYEP by those who commenced attendance was 446 days. Variation in children's total days attended at EYEP is explained by two main factors: first, whether a child dropped out of the trial; and second, their attendance rate while enrolled in EYEP.

Table 4: Measures of attendance at EYEP over 36 months

		By days attended			
	Total	1-59	60-239	240-479	480+
Children still attending at 36 months IQ assessment (per cent)	53.3	0.0	0.0	37.5	78.8
Average total days attended (days)	446.0	24.8	184.2	400.2	584.5
Attendance rate (per cent)	77.4	43.3	71.9	72.4	86.9
Duration from consent date to 36 months IQ assessment	1170.3	1190.7	1140.8	1164.6	1173.8
Duration from consent date to last day attended prior to 36 months IQ assessment (days)	849.2	158.0	462.2	858.5	1028.9
Duration from consent date to first full-time attendance (days)	114.0	119.8	90.2	128.9	109.3
Number of Children	60	6	5	16	33

Notes:
1] Sample is children in the EYEP trial in the intervention group who commenced full-time attendance (12 children did not commence full-time attendance); 2] a] Still attending = Child attended centre for at least a day in the 2 weeks prior to time of 36 months IQ test; b] Attendance rate = Number of days attended/Number of available days. Number of available days = Total potential days for a child to attend EYEP from date of first full-time attendance to last day attended (not including weekends and holiday periods when centre closed); c] Duration from consent date to 36 months IQ assessment = Number of days (including weekends) from first day of full-time attendance to last day of attendance prior to 36 months IQ test; d] Full-time attendance = Attends for 3 hours or more on a day.

What impact of EYEP is estimated?

We investigate three aspects of the impact of EYEP at 36 months. First, an average impact of EYEP across all children (and primary caregivers) who attended the program for at least 60 days in the first 36 months is estimated for each outcome. Second, impacts of EYEP on each outcome are estimated for boys and girls. Third, impacts of EYEP are estimated for children with IQ and language scores equal to and below 90 points and above 90 points at the time of entry to the trial. The threshold of 90 points is chosen to approximately divide the sample into groups of children with IQ and language scores in the top and bottom halves of the distributions of those scores at the time of entry to the trial.

Estimating the impact of EYEP for children who attended for at least 60 days is the same as in previous reports (Tseng et al., 2018, 2019). Making this restriction ensures that children have spent the minimum amount of time attending EYEP needed for the program to have an impact on them. The restriction means we exclude children 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 in the 36 months after enrolment. Setting this threshold causes five children who attended EYEP and for whom data were collected at 36 months to be excluded from the analysis. We do present, however, findings from sensitivity analysis that includes the children who attended EYEP for less than 60 days.

8. Empirical methodology

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 those groups. In our case, this means that the estimated impact should identify that part of the difference in an outcome between children or primary caregivers in EYEP and those in the control group that can be attributed solely to participating in EYEP.

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 each group do not differ systematically in their characteristics; the only distinction is that the intervention group has participated in the program being studied. Hence, any 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 resulted 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 the characteristics.

Whether balance between the characteristics of children and their primary caregivers in the intervention and control groups had been achieved in the EYEP trial was assessed in our report on the impact of the program at 12 months (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 when the primary caregiver consented to participate in the trial. Significant differences, however, were found between the intervention and control groups for some characteristics, data on which were collected in the main stage of data collection at the time of entry to the trial (generally about three months after consent to participate in the trial). 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 why the Bayley Scales scores are significantly higher for the control group than the intervention group is due to selection effects. The selection effects are caused by not being able to collect data for all children on the Bayley Scales measures; for example, some children left the trial between consent to participate and the main stage of data collection at the time of entry to the trial (Tseng et al, 2018, pp.17-18).

Extra analysis of sample attrition was done in our report on the impact of EYEP at 24 months (Tseng et al., 2019) and has also been undertaken for this report. The analysis at 24 months found that several further variables had become unbalanced in the remaining sample or had significantly different attrition patterns – first, risk factors relating to alcohol/substance abuse by a family member, disability or complex medical needs of the child or a family member, and family violence; and second, whether a child was from a family from which multiple children entered the EYEP trial at the time of referral. The analysis for this report has found that the variables which were unbalanced at 24 months remain unbalanced at 36 months. But no extra variables were found to have become unbalanced between 24 months and 36 months in the sample we are able to study for this report.

The imbalance between characteristics of the intervention and control groups implies that to identify the casual impact of EYEP it is necessary to choose an estimation method that can correct for that imbalance. Otherwise, it would not be possible to assess whether differences in outcomes reflect the differences in the characteristics of the intervention and control groups or the effect of EYEP.

To illustrate, suppose hypothetically that in the sample of children from whom data have been collected at 36 months, those enrolled in EYEP are less likely than children in the control group to live with a primary caregiver who has 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 a positive impact from participating in EYEP or that children enrolled in EYEP are living in better home environments because their caregiver is less likely to have the risk factor of alcohol or substance abuse. 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 by applying regression adjusted matching. Regression adjusted matching applies specified covariates to match children in the intervention group to the control group, and then uses the same covariates in a weighted regression (with weights derived from the matching). Essentially, this method estimates the impact of EYEP while controlling for differences between the characteristics of the intervention and control groups that might have caused differences in their outcomes. Using these methods, estimates of the impact of EYEP should reflect only the causal impact of participation in the program.

The key question in application of regression adjusted matching is what characteristics of children and their families need to be controlled for in order that the estimated impact of EYEP reflects only the impact of participation in the program. Statistical theory dictates that it is necessary to control for characteristics of children and their families that are unbalanced between intervention and control groups; or that might affect the outcome being examined (Stuart, 2010).

Using this approach, in our report on the impact of EYEP at 12 months, we specified a set of characteristics that we regarded as the preferred model (Tseng et al., 2018, Appendix 6.2). This included controlling for variables which were unbalanced in that sample – primarily IQ and language development at the time of entry to the trial. Other variables controlled for in the analysis as 12 months were: Gender; Age at 12 months IQ assessment; Duration between IQ assessments at entry to trial and 12 months; a set of dummy variables for age of primary caregiver; a dummy variable for whether the primary caregiver had a post-school qualification; a set of dummy variables for the primary caregiver's level of psychological distress at the time of entry to the trial; a dummy variable for whether a language other than English was the main language spoken at home; a dummy variable for whether both parents were in the household at the time of the consent meeting; and a dummy variable for whether the primary caregiver is an immigrant.

For our report on the impact of EYEP at 24 months, where we found that several further variables had become unbalanced or had significantly different attrition patterns, the same theoretical approach required those characteristics to be controlled for. Our approach in the report on the impact of EYEP at 24 months was therefore to present estimates of the impact of EYEP from two alternative models. The first model was the preferred specification for 12 months. The second model included the same variables as the preferred specification for 12 months plus the extra characteristics: risk factors relating to alcohol/substance abuse, to disability or complex medical needs and to family violence; and whether a child is from a family from which multiple children entered the EYEP trial at the time of referral. Estimated impacts of EYEP at 24 months were found to be similar for the alternative model specifications.

In this report our main empirical approach is to use regression adjusted matching with the second model specification from 24 months (referred to henceforth as the main model). We choose that as the preferred model as it includes the full set of variables found to be unbalanced between the intervention and control groups at 36 months. However, the variables for whether the primary caregiver is an immigrant and for whether a family had multiple children in the trial at the time of referral are included only in the regression stage of analysis, due to problems with convergence in matching when they are also included at that stage.

A range of alternative empirical approaches are used as a robustness check: (i) regression adjusted matching with the preferred model specification from the 12 months

report (that is, the first model used in the 24 months report); (ii) regression adjusted matching with the main model with alternative samples (such as including children who attended EYEP for less than 60 days or excluding children whose initial IQ measure was collected when they were less than six months old); and (iii) using a regression estimation method with both model specifications from 12 months and 24 months. Appendix 3 provides full details of the main model specification in this report and the alternative specification used.

Separate impacts of EYEP after 36 months for boys and girls and for children by their IQ and language development at the time of entry to the trial are estimated using the regression method with the main model specification. It is not feasible to apply the regression adjusted matching method to derive estimates of the impact of EYEP for these separate groups due to small sample sizes once the sample is disaggregated.

Statistical significance

We follow the literature in interpreting statistical significance using what is known as one-tailed test (for example, Karoly, 2005; Elango et al., 2015; Doyle et al., 2017). The estimated impact of EYEP on an outcome is assessed to be statistically significant when the impact is significant at the 10 per cent level. 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. We refer to an estimated impact as being highly statistically significant when the impact is significant at the 5 per cent level. Our choice of a 10 per cent level to distinguish statistical significance is motivated by the sample size in our study and by other studies of small-scale trials having used the same threshold for significance. Statistical significance of estimates of the impact of EYEP using regression adjusted matching is assessed with standard errors calculated using a bootstrap method with bias-corrected p-values.

9. Results

Main results

Estimates of the impact of EYEP after 36 months using regression adjusted matching with the main model specification are presented in Table 5 for: (i) children's development (rows (1) to (4)); and (ii) outcomes for their primary caregivers (rows (5) to (7)). Results from robustness analysis – using the alternative model specification, samples and estimation method – are reported in Table 6. Other sensitivity analysis – mainly for children's IQ and language development - is reported in Table 7. The evolution of estimated impacts of EYEP throughout the trial (at 12, 24 and 36 months) is reported in Table 8. Note that the impacts in Table 8 are not for a balanced sample of children, they are simply the results reported at 12, 24 and 36 months for the full samples of children available for analysis at those times.

IQ The estimated impact of EYEP on children's IQ at 36 months is positive, statistically significant and robust. In the main model the estimated impact is 7.7 points. The size and statistical significance of the estimated impacts are robust to the alternative model specification, using a regression method, changes to the sample and including

Table 5: Impact of enrolment in EYEP for 36 months (Children in EYEP group who attended for at least 60 days)

	Outcomes	EYEP mean: at trial entry	EYEP mean: at 36months	EYEP impact	p-value: 1-tail	Number of observations (EYEP/Control)
	Children's outcomes					
(1)	IQ	92.44	99.63	7.65	0.076	43/42
(2)	Language	90.19	99.47	6.83	0.119	43/42
(3)	Resilience		44.75	2.11	0.361	40/36
(4a)	Social and emotional development (Per cent below bottom 10th percentile population norm)		26.1	-17.7	0.309	38/38
(4b)	Social and emotional development (Per cent below bottom 17th percentile population norm)		29.7	-27.0	0.077	38/38
(4c)	Social and emotional development (Per cent below bottom 25th percentile population norm)		44.7	-30.8	0.015	38/38
(4d)	Child behaviour problems		55.68	-6.19	0.036	37/38
(4e)	Child – Externalising problems		53.49	-5.46	0.072	37/38
(4f)	Child – Internalising problems		55.38	-5.03	0.078	37/38
	Parents' outcomes					
(5)	Parenting daily hassles - Frequency		45.64	-0.493	0.432	34/37
(6)	Parenting daily hassles – Intensity		46.20	1.067	0.402	34/37
(7)	Psychological distress (k6)		14.93	-0.061	0.483	40/40

Notes:
1] The EYEP impacts in rows (1)-(3), (4d) to (4f) and (5)-(7) are the estimated impacts of attending EYEP using a regression adjusted matching method. The EYEP impacts in rows (4a) to (4c) are the mean marginal impacts on the per cent of children below the respective threshold on the social-emotional measure estimated from a probit model on a matched sample; and
2] Bootstrap bias adjusted p-values based on 999 replications are reported.

a dummy variable for whether the Bayley Scales or WPPSI measure was used. These robustness checks give estimates of the impact of EYEP between 6.6 and 8.0 points, generally highly significant.

At 36 months children attending EYEP had an average IQ score of 99.6 points. A statistical significance test cannot reject this score as being equal to the population norm average of 100 (at the 80 per cent level of significance). By contrast, the average IQ score of the control group at 36 months can be rejected as equalling 100 (at the 1 per cent level of significance).

The estimated impact of EYEP on IQ at 36 months is large, representing one-half of a standard deviation on the scale for that outcome. By comparison, 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 twenty 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). Estimated impacts of early years 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 are generally smaller for programs implemented in more recent years (Elango et al., 2015, p.32; Duncan and Magnuson, 2013, p.123).

Looking across the whole trial, EYEP had an impact on IQ at 12 months that has been sustained and may even have increased over the next two years, with the effect being statistically significant throughout. Using the main model for this report, the estimated impact at 12 months was 5.7 points, at 24 months was 7.1 points and 7.7 points at 36 months. Using the model specification at 12 months, the EYEP impact on IQ at 12 months was 3.8 points, at 24 months was 5.0 points and at 36 months was 6.8 points.

The finding of a major impact on IQ at 12 months, that persists throughout the program, is consistent with evidence from previous trials of early years demonstration programs. What is notable about EYEP is the indication that the impact of the program has continued to increase after 12 months. By contrast, 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. Estimates of the impact on IQ from when children have commenced school – to be undertaken for a future report - will be informative about whether there is fade-out in the impact of EYEP on IQ once children have completed the program as has been found for the demonstration programs in the United States (Hojman, 2015; Elango et al., 2015, pp.31-32).

Language The estimated impact of EYEP on children's language skills at 36 months is positive with sufficient evidence to regard the impact as statistically significant. In the main model the estimated impact is 6.8 points and almost statistically significant. The robustness checks give estimates of the impact of EYEP on language between 6.3 and 7.2 points, always highly statistically significant.

At 36 months children attending EYEP had an average language score of 99.5 points. A statistical significance test cannot reject this score as being equal to the population norm average of 100 (at the 80 per cent level of significance). By contrast, the average language score of the control group at 36 months can be rejected as equalling 100 (at the 1 per cent level of significance).

The impact of EYEP on children's language skills appears to have steadily strengthened through their time in the trial – until at 36 months the impact is large and statistically significant. For example, the EYEP impact on language at 12 months was 0.6 points; and using same estimation method was 2.9 points at 24 months and 6.6 points at 36 months.

Within-child protective factors related to resilience

The estimated impact of EYEP on children's protective factors related to resilience at 36 months is positive but does not approach statistical significance. Using the main estimation method, the impact is 2.1 points. Estimated impacts from the robustness checks range from 2.1 to 2.6 points, with slightly higher levels of statistical significance.

The size of impact of EYEP on resilience has been relatively steady throughout children's time in the program, varying in narrow range from 0.2 to 0.33 of a standard deviation on the DECA measure. For example, the EYEP impact at

Table 6: Impact of enrolment in EYEP for 36 months – Robustness analysis

	Outcomes	(1) Main model	(2) Alternative model	(3) Full sample	(4) DV for children who attend school	(5) Regression – Main model	(6) Regression – Alternative model
	Children's outcomes						
(1)	IQ	7.646 (0.076)	6.825 (0.042)	7.939 (0.029)	7.650 (0.048)	8.226 (0.019)	7.395 (0.018)
(2)	Language	6.834 (0.119)	6.606 (0.068)	7.231 (0.049)	6.830 (0.100)	7.019 (0.028)	6.591 (0.027)
(3)	Resilience	2.113 (0.361)	2.684 (0.204)	2.210 (0.445)	2.107 (0.368)	2.636 (0.178)	2.815 (0.131)
(4a)	Social and emotional development (Per cent below bottom 10th percentile population norm)	-16.4 (0.329)	-4.9 (0.479)	-14.6 (0.344)	-20.8 (0.233)	-18.2 (0.095)	-8.9 (0.248)
(4b)	Social and emotional development (Per cent below bottom 17th percentile population norm)	-26.4 (0.077)	-14.5 (0.247)	-21.5 (0.139)	-27.6 (0.076)	-21.9 (0.083)	-14.4 (0.155)
(4c)	Social and emotional development (Per cent below bottom 25th percentile population norm)	-30.6 (0.015)	-23.7 (0.061)	-28.8 (0.027)	-30.5 (0.021)	-17.9 (0.112)	-16.7 (0.101)
(4d)	Child behaviour problems	-6.187 (0.036)	-4.811 (0.067)	-5.964 (0.099)	-6.190 (0.045)	-4.590 (0.096)	-3.623 (0.132)
(4e)	Child – Externalising problems	-5.461 (0.072)	-4.103 (0.128)	-5.006 (0.154)	-5.305 (0.097)	-3.719 (0.151)	-2.845 (0.191)
(4f)	Child – Internalising problems	-5.025 (0.078)	-4.341 (0.088)	-5.122 (0.126)	-5.160 (0.075)	-4.156 (0.090)	-3.315 (0.123)
	Parents' outcomes						
(5)	Parenting daily hassles – Frequency	-0.493 (0.432)	1.631 (0.258)	-0.635 (0.424)	-0.494 (0.440)	-0.044 (0.502)	1.560 (0.330)
(6)	Parenting daily hassles – Intensity	1.067 (0.402)	2.991 (0.197)	0.887 (0.390)	1.095 (0.399)	1.350 (0.398)	3.217 (0.251)
(7)	Psychological distress (K6)	-0.061 (0.483)	-0.048 (0.434)	0.251 (0.464)	-0.071 (0.497)	-0.199 (0.462)	-0.109 (0.477)

Note:
1] p-values are in parentheses;
2] a] Columns (1) to (4): (i) The EYEP impacts in rows (1)-(3), (4d) to (4f) and (5)–(7) are the estimated impacts of attending EYEP using a regression adjusted matching method. The EYEP impacts in rows (4a) to (4c) are the mean marginal impacts on the per cent of children below the respective threshold on the social-emotional measure estimated from a probit model on a matched sample. Column (1) reports estimates from the main model – see Table 6. Column (2) reports estimates from the alternative model specification used in the 12 months report. Column (3) reports estimates using the main model and expanding the sample to include children in the intervention group who attended EYEP for 60 days or less over the 36 months of the trial. Column (4) reports estimates using the main model and including a dummy variable for whether a child attended school during the 12 months preceding the 36 months data collection. b] Columns (5) and (6): (i) The EYEP impacts in rows (1)–(3), (4d) to (4f) and (5)–(7) are the estimated impacts of attending EYEP using Ordinary Least Squares. The EYEP impacts in rows (4a) to (4c) are the mean marginal impacts on the per cent of children below the respective threshold on the social-emotional measure estimated from a probit model. Column (5) reports estimates from the main model. Column (6) reports estimates from the alternative model specification used in the 12 months report;
3] Bootstrap bias adjusted p-values based on 999 replications are reported.

12 months was 2.0 points; and using the same estimation method was 3.3 points at 24 months and 2.7 points at 36 months. The estimated impacts at 12 and 36 months are not statistically significant, although they do come close to being at that level. At 24 months the estimated impact is highly statistically significant.

The consistency of impact throughout the EYEP trial, combined with the mixed evidence on statistical significance, appear to justify a conclusion that the program has had a small positive average impact on children's resilience, but with a relatively high degree of variability across children.

Social and emotional development In previous reports we have represented the estimated impact of EYEP on social and emotional development as its impact on the percentage of children below the 10th percentile of the norm population distribution of the BITSEA/CBCL measure. On that measure, the EYEP impact at 36 months is positive. Using the main model, the estimated impact of EYEP is to reduce the percentage of children falling in the bottom 10 per cent by 16.4 percentage points; and with the robustness checks the estimated impact ranges from 4.6 to 19.7 percentage points. However, none of these estimated impacts are close to being statistically significant. This is quite different from the estimated impacts at 12 months (lower by 12.2 percentage points and almost statistically significant); and at 24 months (lower by 31.6 percentage points and highly statistically significant).

Extending the analysis at 36 months to estimate the impact of EYEP on the percentage of children falling below higher thresholds, the bottom 17 per cent and 25 per cent of population, there is stronger evidence of effects of the program. For example, using the main model, the estimated impact of EYEP is to lower the proportion of children in the bottom 25 per cent by 30.6 percentage points, with the effect being highly statistically significant.

It is also possible to estimate the effect of EYEP at 36 months using just the CBCL measure, since all except one child were assessed using that measure at 36 months. With the CBCL measure, EYEP is found to have significantly improved children's social and emotional development at 36 months. The impact of EYEP is to lower children's CBCL overall score by 6.2 points (equal to 0.6 of a standard deviation), an effect which is highly statistically significant. Lower scores on the separate components of externalising

Table 7: Impact of enrolment in EYEP for 36 months – Robustness analysis

Outcomes	(1) IQ	(2) Language	(3) Resilience	(4) Number of observations (EYEP/Control)
Main model	7.646 (0.076)	6.834 (0.119)	2.113 (0.361)	43/42 (IQ and Language) 40/36 (Resilience)
Exclude children whose baseline IQ was assessed prior to 6 months of age	9.684 (0.030)	9.347 (0.040)		36/36
Exclude children who were born premature and whose baseline IQ was assessed prior to 6 months of age	7.633 (0.040)	7.190 (0.069)		42/42
Exclude children who were born premature and whose baseline IQ was assessed prior to 12 months of age	8.386 (0.025)	8.202 (0.034)		42/38
Exclude children whose baseline IQ was assessed prior to 6 months of age and children who were born premature and whose baseline IQ was assessed prior to 12 months of age	10.912 (0.017)	10.843 (0.010)		36/32
Include dummy variable for WPPSI (cf. Bayley scales)	7.626 (0.052)	7.178 (0.079)		43/42
Include dummy variable for DESSA (cf. DECA)			2.009 (0.390)	40/36

Note:
1] p-values are in parentheses;
2] Premature = Born at 36 weeks or earlier.

behaviour and internalising behaviour are found, by 5.5 and 5.0 points respectively, both statistically significant. The estimated impacts on CBCL are robust to the alternative estimation methods and samples.

Large and significant impacts of EYEP are therefore found for most estimates based on the percentage of children below a threshold level of social and emotional development, and when the continuous CBCL measure is used. The exception is the estimated impact of EYEP on the proportion of children below the 10th percentile. We interpret this latter result as mainly reflecting the extra variability in estimated impacts when a categorical-type outcome measure is used: Comparing the results at 24 months and 36 months we find that a small number of children in the intervention group have crossed from just above to just below the 10th percentile level, with the opposite happening for the control group. This is why the estimated impacts of EYEP on the proportions of children below the 17th and 25th percentiles are large and significant, whereas the impact on the 10th percentile measure becomes smaller.

Table 8: Impact of enrolment in EYEP – 12, 24 and 36 months

	Outcomes	(1) 12 months	(2) 24 months	(3) 36 months
	IQ			
(1)	Main model (24 months report)		7.06 (0.017)	7.65 (0.076)
(2)	Alternative model (12 months report)	3.81 (0.065)	5.01 (0.068)	6.82 (0.042)
	Language			
(1)	Main model (24 months report)		5.15 (0.072)	6.83 (0.119)
(2)	Alternative model (12 months report)	0.56 (0.425)	2.87 (0.225)	6.61 (0.068)
	Resilience			
(1)	Main model (24 months report)		3.34 (0.059)	2.11 (0.361)
(2)	Alternative model (12 months report)	1.95 (0.163)	3.29 (0.047)	2.68 (0.204)
	Social-emotional development: Bottom 10 per cent			
(1)	Main model (24 months report)		-29.2 (0.003)	-16.4 (0.329)
(2)	Alternative model (12 months report)	-12.2 (0.107)	-31.6 (0.01)	-4.6 (0.423)
	Parent psychological distress (K6)			
(1)	Main model (24 months report)		-1.77 (0.098)	-0.06 (0.483)
(2)	Alternative model (12 months report)	-1.01 (0.187)	-1.65 (0.098)	-0.05 (0.434)

Note:
1] p-values are in parentheses;
2] Estimated impacts of EYEP reported are for the full samples available at 12 months, 24 months and 36 month; and
3] Estimated impacts in top row at 24 months and 36 months are from 'Main model' (second model used in 24 months report). Estimated impacts in bottom row at 12 months, 24 months and 36 months are from 'Alternative model' (model used in 12 months report).

Parenting daily hassles No evidence is found of an impact of EYEP on parenting daily hassles at 36 months. Using the main model, the estimated impact of EYEP is a small decrease in frequency of hassles (0.5 point on zero to 80 scale) and slight increase in intensity of hassles (1.1 points on zero to 100 scale). Neither estimated impact is close to being statistically significant. A range of small negative and positive estimated impacts is found from the robustness checks, none of which are close to being statistically significant.

Psychological distress of primary caregivers (K6) No evidence of an overall impact of EYEP on the psychological distress of primary caregivers is found at 36 months. Using the main model, the estimated impact is to lower the K6 measure by 0.1 point (on 6 to 30 scale), an impact which is not close to being statistically significant. This is a weaker effect than the estimated impacts at 12 months (lower by 1.0 point but not statistically significant) and at 24 months (lower by 1.7 points and statistically significant).

Impacts by gender

Estimates of the impact of EYEP are reported separately for boys and girls in Table 9. The evolution of effects over time in the trial for boys and girls are shown in Table 10. Quite distinct patterns of impact are found.

IQ/Language: The estimated impacts of EYEP on IQ and language at 36 months are much larger for boys than girls; and the impact is statistically significant for boys but not girls. The differences between the impacts for girls and boys are not significant, although that is likely to be explained by the small sample size. The larger impact of EYEP on IQ for boys than girls is consistent with the

findings at 12 and 24 months. As well, for both IQ and language, a large jump in the size of impact for boys compared to girls happens from 24 to 36 months.

Resilience/Social and emotional development:
The estimated impact of EYEP on resilience is slightly larger for boys than girls at 36 months, but the estimates for both groups are not statistically significant. This contrasts with the pattern at 24 months when there was a much larger and highly significant impact for boys and zero effect for girls. There is mixed evidence on the relative impact of EYEP on social and emotional development for boys and girls. Using the percentage of children below the 10th percentile in the population distribution, the impact is larger for girls than boys, continuing the pattern from 24 months. But the size of impact for girls, and difference from boys, are not as large or significant at 36 months. Using the 17th percentile threshold finds a larger and significant impact for girls; but with a 25th percentile threshold that pattern reverses, with a larger and significant impact for boys. The estimated impact on social and

Table 9: Impact of enrolment in EYEP for 36 months – By gender

	Outcomes	(1) Main model	(2) Girls	(3) Boys	(4) Girls-Boys	Sample size {Girls(T/C); Boys(T/C)}
	Children's outcomes					
(1)	IQ	7.646 (0.076)	4.732 (0.260)	10.801 (0.066)	-6.069 (0.305)	{(23/21); (20/21)}
(2)	Language	6.834 (0.119)	2.520 (0.357)	10.691 (0.072)	-8.171 (0.216)	{(23/21); (20/21)}
(3)	Resilience	2.113 (0.361)	3.410 (0.270)	1.200 (0.476)	2.210 (0.301)	{(20/18); (20/18)}
(4a)	Social and emotional development (Per cent below bottom 10th percentile population norm)	-16.4 (0.329)	-20.3 (0.214)	-6.10 (0.302)	-10.5 (0.371)	{(19/20); (19/18)}
(4b)	Social and emotional development (Per cent below bottom 17th percentile population norm)	-26.4 (0.077)	-34.1 (0.046)	-21.4 (0.261)	-11.6 (0.242)	{(19/19); (19/18)}
(4c)	Social and emotional development (Per cent below bottom 25th percentile population norm)	-30.6 (0.015)	-13.6 (0.188)	-39.9 (0.008)	26.3 (0.232)	{(19/19); (19/18)}
(4d)	Child behaviour problems	-6.193 (0.035)	-8.733 (0.083)	-4.329 (0.223)	-4.404 (0.273)	{(18/19); (19/18)}
(4e)	Child – Externalising problems	-5.448 (0.072)	-6.863 (0.142)	-4.408 (0.255)	-2.455 (0.342)	{(18/19); (19/18)}
(4f)	Child – Internalising problems	-5.019 (0.078)	-4.904 (0.232)	-5.104 (0.163)	-0.199 (0.498)	{(18/19); (19/18)}
	Parents' outcomes					
(5)	Parenting daily hassles – Frequency	-0.493 (0.432)	-1.155 (0.324)	-0.009 (0.329)	-1.147 (0.326)	{(16/17); (18/20)}
(6)	Parenting daily hassles –	1.067 (0.402)	6.994 (0.297)	-0.968 (0.476)	7.961 (0.359)	{(16/17); (18/20)}
	Intensity					
(7)	Psychological distress (k6)	-0.061 (0.483)	-2.416 (0.167)	1.593 (0.258)	-4.008 (0.135)	{(16/17); (18/20)}

Note:
1] p-values in parentheses;
2] Column (1) reports estimates from regression adjusted matching using the main model – see Table 6. Columns (2) and (3) report estimates from regression analysis on separate samples respectively of girls and boys. Column (4) reports estimates from a weighted regression model estimated on the matched sample with a variable interacting a dummy variable for being in the intervention group with a dummy variable for gender.

emotional development using the CBCL measure shows a larger impact for girls than boys (entirely due to the externalising problems scale), with the impact for girls statistically significant.

Impacts on outcomes for primary caregivers: For the K6 measure, the impact of EYEP is to lower psychological distress by a larger amount for primary caregivers of girls than boys, with the difference being close to statistically significant. This is the same pattern as was estimated at 24 months. There is no evidence of a difference between primary caregivers of girls and boys in the impact of EYEP on parenting daily hassles.

Table 10: Impact of enrolment in EYEP – 12, 24 and 36 months

Outcomes	(1) 12 months	(2) 24 months	(3) 36 months
IQ			
Boys		7.27	10.80
		(0.039)	(0.066)
	5.81	6.39	
	(0.038)	(0.056)	
Girls		6.70	4.73
		(0.053)	(0.260)
	1.62	3.11	
	(0.347)	(0.226)	
Language			
Boys		0.36	10.69
		(0.470)	(0.072)
	0.93	4.67	
	(0.405)	(0.146)	
Girls		4.25	1.20
		(0.190)	(0.476)
	-2.46	5.30	
	(0.257)	(0.470)	
Resilience			
Boys		5.27	1.20
		(0.015)	(0.476)
Girls		-0.002	3.41
		(0.500)	(0.270)
Social-emotional development: Bottom 10 per cent			
Boys		-10.4	-5.2
		(0.240)	(0.411)
Girls		-52.7	-20.5
		(0.001)	(0.254)
Parent psychological distress (K6)			
Boys		-1.03	1.59
		(0.265)	(0.255)
Girls		-2.97	-2.42
		(0.076)	(0.167)

Note:
1] p-values are in parentheses;
2] Estimated impacts of EYEP reported are for the full samples available at 12 months, 24 months and 36 month; and
3] Estimated impacts in top row at 24 months and 36 months are from 'Main model' (second model used in 24 months report). Estimated impacts in bottom row at 12 months and 24 months and from 'Alternative model' (model used in 12 months report).

Impacts by IQ/Language at time of entry to trial

Estimates of the impact of EYEP separately for children whose IQ and language scores at the time of entry to the trial were 90 points and below or above 90 points are reported in Table 11. Much larger impacts on IQ and language are found for children whose score at the time of entry to the trial was 90 points or below, effects which are highly statistically significant. The size of increase in IQ is sufficient to lift children with a low starting IQ to the population norm at 36 months. From an average at the time of entry to the trial of 85 points, by 36 months children who started with an IQ of less than or equal to 90 points had improved to have an average IQ of 98.6 points. For language skills, the improvement in average score is from 86.5 points to 98.2 points. It is not possible to reject the hypothesis that the IQ and language scores of children whose IQ was 90 points or below at entry to the trial have increased at 36 months to equal the population norm average of 100 (at levels of significance above 50 per cent).

Summary

At 36 months, EYEP had important impacts on children's cognitive and non-cognitive development:

1] Children's cognitive development: Children's IQ and language development are sufficiently improved by EYEP that the program's objective to make participants developmentally equal to their peers is achieved. EYEP achieved a large increase in children's IQ in the first 12 months of the program, with that impact being sustained, and some evidence of further improvement, in the next 24 months. The impact on language grew throughout the trial, and has become statistically significant at 36 months. Larger and more significant impacts of EYEP on cognitive development are estimated for boys than girls. Much larger and more significant impacts on cognitive development are found for children with low IQ and language skills at the time of entry to trial. The improvement for these children is so large as to raise

Table 11: Impact of enrolment in EYEP for 36 months – By IQ/Language score at time of entry to trial

	Outcomes	(1) Main model	(2) Score <=90	(3) Score>90	Sample sizes{(<=90(T/C); >90(T/C))}
(1)	IQ	7.646 (0.076)	14.769 (0.021)	2.216 (0.405)	{(21/27); (22/15)}
(2)	Language	6.834 (0.119)	14.283 (0.044)	2.159 (0.412)	{(17/16); (26/26)}

Note: p-values in parentheses.

their average levels of IQ and language development to the same level as for the general population.

2] Children's non-cognitive development: The estimated impact on children's protective factors related to resilience is relatively small and not statistically significant. However, throughout the trial a small impact on resilience has been sustained, and that impact was highly significant at 24 months. Hence, we conclude there is likely to be a small positive average impact of EYEP on resilience, but with relatively high variability between children. Stronger evidence exists of an impact of EYEP on children's social and emotional development. Large and significant impacts on the CBCL measure and on the percentage of children with scores below the 17th and 25th percentile thresholds are found. While the impact of EYEP on the percentage of children below the 10th percentile is smaller and less significant than at 24 months, this is likely to be mainly explained by the extra variability introduced into estimated impacts when categorical outcome measures are used. Overall, it seems reasonable to conclude that by 24 months EYEP had improved children's social and emotional development and that this impact has been sustained to 36 months. The majority of evidence suggests that impacts on non-cognitive development have been larger for girls than boys.

3] Primary caregivers: Only limited evidence exists of impacts of EYEP on the well-being of primary caregivers. A significant impact of EYEP in reducing psychological distress of primary caregivers is found for girls, but apart from that no statistically significant impact is detected.

10. Interpretation

Interpretation 1 – Why is EYEP affecting outcomes?

The impact of EYEP on IQ might derive from several sources: the attributes and quality of the EYEP; children in EYEP receiving an increased quantity of early years education and care; improvement to the financial situation of families in EYEP due to receiving free early years education and care; or an increased quantity and quality of parental inputs from primary caregivers in EYEP.

It is not possible to closely evaluate whether improvements to a family's financial position or to parental inputs have contributed to the impact of EYEP. But it is known that a relatively small proportion of families in the control group faced out of pocket costs for childcare (26.5 per cent). As well, EYEP was not intended to directly affect the behaviour of primary caregivers; nor do we find significant impacts of EYEP on the psychological well-being of primary caregivers. Hence, the most likely reasons for the impact of EYEP on outcomes is differences in the quality and quantity of formal education and care received by the intervention group compared to the control group.

The attributes of EYEP – and how the program differs from usual care – have been documented in section 3 (as well as described in detail in Jordan and Kennedy, 2019). The remainder of this sub-section addresses how the quantity of formal education and care differed between the intervention and control groups. It is concluded that the size of difference in the amount of formal education and care received by the groups is sufficient to suggest that this may also partly explain the impact of EYEP, in addition to the quality and attributes of the program.

Outcomes for children in the intervention group at 36 months will reflect the accumulated effects of their participation in EYEP over the whole duration of the trial. Hence, to interpret the estimated impacts of EYEP, we compare the types and amount of formal education and care received by children in the intervention and control groups over the 36 months of the trial. We also present information for the 12 months preceding the 36 months data collection to complete the picture (together with the reports from 12 and 24 months) of how usage of formal education and care evolved during the trial.

Data on children's participation in formal education and care outside EYEP was obtained from primary caregivers at the annual data collection. For the control group this data source is used as to estimate their total usage of formal education and care. For children who commenced school prior to the 36 months data collection, total school hours are also included. For the intervention group, the total amount of formal education and care is defined to be equal to time enrolled at EYEP plus usage of other formal services as recorded from the annual data collection. Information on EYEP attendance is for the 12 months prior to the data collection with the primary caregiver, to ensure that information is comparable with control group.

The data collection on children's participation in formal education and care involved questions regarding aggregate experiences in the previous 12 months, rather than collecting diary-type information. This has implications for the types of descriptive information on participation in early years education and care that can be presented. First, the set of types of education and care services used by a child in the previous 12 months can be listed, but the exact timing of receipt of those services is not known. Second, information on the total amount of education and care received is derived from responses to questions on weeks of care and hours per week of care received over the previous 12 months.

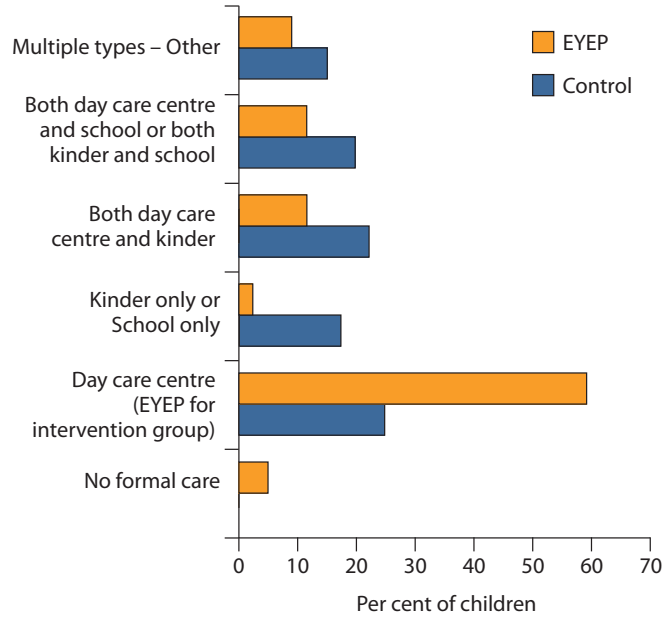
The information on formal education and care in the preceding 12 months is from the sample of children used to estimate the impact of EYEP on IQ and for whom data on usage of education and care services were collected at 36 months. The information on formal education and care over the whole 36 months is based on the sample of children used to estimate the impact on IQ and for whom data on usage of education and care services were collected at 12, 24 and 36 months. Note that these samples for the intervention group differ from the sample for whom attendance data were reported in Table 4.

In interpreting information on use of formal education and care 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 their care plan'. This is likely to cause higher levels of usage of education and care services in the control group than would otherwise be the case.

Figure 3 shows the types of formal education and care services received by the intervention and control groups in the 12 months preceding the 36 months data collection. Almost all children had some type of formal education and care during the year. Children in the intervention group were most likely to have day care centre only (EYEP) as their main activity, with small proportions also having been at a kindergarten program outside EYEP or at school at some time during the year. Children in the control group are evenly spread between day care centre only, kindergarten or school only, and day care centre with either kindergarten or school at some time during the year. The scope for children in the control group to attend kindergarten and school appears to have caused an increase in their likelihood of receiving formal education and care and a large change in the type of formal education and care they received. In the first 24 months of the trial about 20 per cent of the control group had no formal care and about 50 per cent had day centre care only. Funding for 15 hours per week of kindergarten for 4 year-old children was gradually introduced across Victoria from 2013 coinciding with the control group reaching the age at which they could access that service.

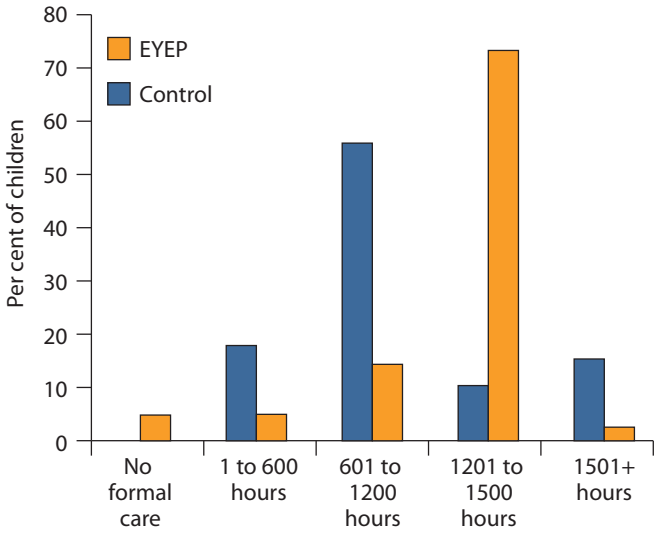
Figures 4a and 4b compare the distributions of hours of formal education and care between children in the

Figure 3: Type of formal education and care attended by children in the EYEP and control groups in previous 12 months prior to data collection at 36 months



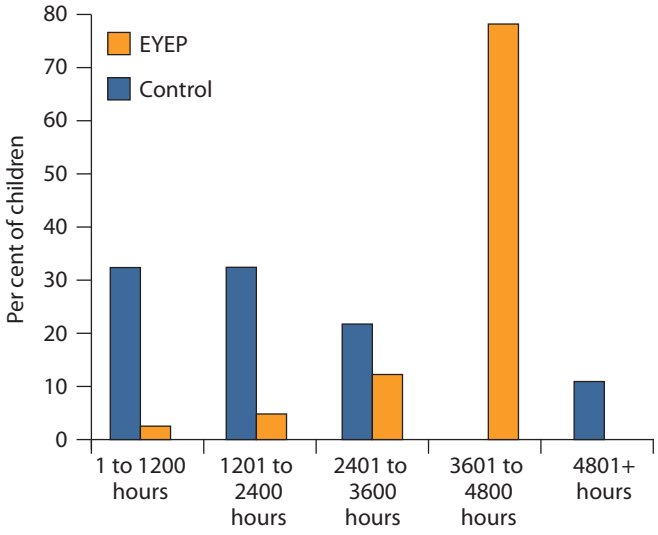
Note: Sample is children who provided data on IQ at 36 months data collection.

Figure 4a: Distribution of annual hours in formal education and care in past 12 months at 36 months in EYEP trial



Note: Sample is: (i) children who provided data on IQ at 36 months; (ii) children in the EYEP group who had attended for at least 60 days in the first 36 months.

Figure 4b: Distribution of total hours in formal education and care in 36 months of EYEP trial



Note: Sample is children who provided data on IQ at 12, 24 and 36 months; and (ii) children in the EYEP group who had attended for at least 60 days in the first 36 months.

Table 12: Average hours in formal education and care in EYEP trial			
	All	Girls	Boys
1] Past 12 months (to 36 months)			
EYEP	1130.0	1106.0	1156.0
Control	1037.0	1108.0	954.0
Number of observations – EYEP/Control	40/39	20/21	20/18
2] 36 months			
EYEP	3556.0	3630.3	3481.7
Control	2103.1	1592.9	2576.9
Number of observations – EYEP/Control	40/27	20/13	20/14

Note:
Sample for past 12 months at 36 months is: (i) children who provided data on IQ at 36 months; (ii) children in the EYEP group who had attended for at least 60 days in the first 36 months. Sample over 36 months is children who provided data on IQ at 12, 24 and 36 months; and (ii) children in the EYEP group who had attended for at least 60 days in the first 36 months.

intervention and control groups. Table 12 shows average hours. In the 12 months preceding the data collection at 36 months, annual hours of formal education and care are similar for the intervention and control groups, with average annual hours respectively of 1130 and 1037. Looking across the whole 36 months there is a larger difference in average total hours, 3556 hours for the intervention group and 2103 hours for the control group. The larger difference over 36 months than in the final 12 months reflects the large differences found between the intervention and control groups in the first 24 months of the trial. The scope to attend kindergarten and school for most children in the control group in the last 12 months of the trial then meant that time in formal education and care was similar for the groups over that period.

We conclude that the most likely explanations for the impact of EYEP on children's development are first, the attributes and quality of EYEP compared to the formal early years education and care received by the control group; and second, that children in the intervention group received a larger amount of early years education and care over the first 24 months of the trial. There is not strong evidence that the impact of EYEP has derived from an effect on the budgets of families in the intervention group, or on inputs from their primary caregivers.

Interpretation 2 – Why does the impact of EYEP vary by gender?

Our finding of larger impacts of EYEP on the cognitive development of boys than girls is largely consistent with studies of early years programs in the United States. That the pattern of impact by gender becomes mixed once non-cognitive development is considered is also not entirely at odds with the previous studies. Overall, however, it would have to be said that our findings on the relative size of impacts by gender are more mixed than the previous literature.

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 in these studies for why there is a larger impact for boys than girls is that boys' development is more sensitive to the quality and level of stress in 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, a larger impact on social and emotional development for girls than boys is found in a recent re-evaluation of Abecedarian project (Garcia et al., 2018).

This difference between our findings and previous studies may simply be due to features of the sample of participants in the EYEP trial; for example, girls in the EYEP trial had a higher average IQ at the time of entry to the trial and hence had less scope to improve. Alternatively, this analysis of EYEP may be yielding new insights into impacts by gender due to the wide range of outcome variables being considered. For that reason, understanding more about the sources of gender differences in the impact of EYEP is an important future task.

Interpretation 3 – Why does the impact of EYEP vary by IQ/Language at the time of entry to the trial?

The difference in the impact of EYEP by children's IQ and language development at their time of entry to the trial is striking. The effect of EYEP is highly concentrated on children whose initial score on IQ/language was 90 points or less. By 36 months, the average IQ and language scores of those children had increased substantially towards the population norm average of 100 points; whereas impacts on children with initial IQ and language scores above 90 points are much smaller and not statistically significant. The difference in impacts by starting IQ and language is consistent with EYEP being 'compensatory', making up for deficiencies in development of children who 'start behind'. However, the evidence of gradual improvement in IQ throughout the trial, and the jump in impact of EYEP on language between 24 and 36 months, also suggests a role for dynamic complementarity, where an improvement in skill development at a point in time begets later improvement in skills.

Interpretation 4 – Timing of impacts of EYEP on outcomes for children

Distinct patterns of timing in the impact of EYEP are observed, with the patterns differing between outcomes. A significant impact on IQ emerges at 12 months, and there is some indication that the size of impact increases further at 24 and 36 months. The impact on language jumps in size and becomes significant between 24 and 36 months. The impact on children's protective factors related to resilience is relatively steady throughout the trial, but peaks in size and is only significant at 24 months. The effect of EYEP on social and emotional development is large and significant at 24 months; and the balance of evidence is that the same impact remains at 36 months.

There are several potential explanations for the timing of impacts of EYEP on the different outcomes, although at this stage our discussion is speculative. It is also worth noting that the timing of impacts of early years interventions on outcomes is not a topic that has been much addressed in existing international studies of early years programs.

Concentration of impacts of EYEP at 24 months onwards for most outcomes for children may demonstrate that this is the amount of time necessary for the program to have an impact on the development of children from highly disadvantaged and stressful environments. But then an explanation would also be needed for why there was a significant impact prior to 24 months on IQ. A possibility is that different amounts of EYEP are required to affect different outcomes, which might reflect a natural progression of skill development for children.

Another dimension to the pattern of timing of impacts – relating to the final 12 months of the trial - may be a slowing in the 'falling behind' by children in the control group. Participation of control group children in formal early years education and care increased substantially in their final 12 months of the trial, mainly it seems due to the scope for them to attend kindergarten and school. The narrowing in the difference in the amount of formal education and care received by children in intervention and control groups might explain why the estimated impact of EYEP on children's non-cognitive development appears to stabilise between 24 and 36 months. (Manning et al., 2017, provide evidence that teacher-led programs achieve better outcomes for children). The difficulty with this explanation however is that impacts of EYEP on children's cognitive development increased over the same period.

With many unresolved questions, learning more about the ordering of impacts of EYEP on outcomes is an important future task for research. This could include investigating how changes during the trial in relative usage of early years education and care by the intervention and control groups have affected the evolution of impacts of EYEP.

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, (2021), *Child Protection Australia 2019-20*, Child Welfare series no.74, Catalogue no. CWS 78 (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. 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.

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.

Manning, M., S. Garvis, C. Fleming and G.T.W. Wong (2017), 'The relationship between teacher qualification and the quality of early childhood education and care environment', A Campbell Systematic Review 2017/1.

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.

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

Tseng, Y., B. Jordan, J. Borland, N. Coombs, K. Cotter, M. Guillou, A. Hill A. Kennedy and J. Sheehan (2019), *Changing the Life Trajectories of Australia's Most Vulnerable Children – Report no.4: The first 24 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/0003/3085770/EYERP-Report-4-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

Children's Protection Society/Kids First Assistant to Research/Governance Committee

Margaret Farquharson

Children's Protection Society/Kids First 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

University of Melbourne Research staff

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

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):	
Main model (24 months report)	**Alternative model (12 months report)**
i] Variables included in both matching and regression stages: Gender; Age at 36 months IQ assessment; Duration between IQ assessments at entry to trial and 36 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 language other than English is main language spoken at home; DV for whether both parents were in household at time of consent meeting; IQ score at time of entry to trial; Language score at time of entry to trial; DV for whether alcohol/substance abuse was a risk factor at referral; DV for whether primary caregiver/child/sibling had complex medical issues at referral; DV for whether family violence (current or past) was a risk factor at referral; ii] Variables included in only regression stage: DV for whether primary caregiver is immigrant; DV for multiple children in family in trial from time of referral.	Main model modified by excluding 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 primary caregiver/child/sibling had complex medical issues at referral; DV for whether family violence (current or past) was a risk factor at referral.

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
Report no. 5	October 2022	36 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. 5
36 months in the Early Years Education Program: Assessment of the impact on children and their primary caregivers