

A meta-analysis of the effects of early developmental prevention programs in at-risk populations on non-health outcomes in adolescence

Author

Manning, Matthew, Homel, Ross, Smith, Christine

Published

2010

Journal Title

Children and Youth Services Review

DOI

[10.1016/j.chidyouth.2009.11.003](https://doi.org/10.1016/j.chidyouth.2009.11.003)

Rights statement

© 2010 Elsevier. This is the author-manuscript version of this paper. Reproduced in accordance with the copyright policy of the publisher. Please refer to the journal's website for access to the definitive, published version.

Downloaded from

<http://hdl.handle.net/10072/32011>

Griffith Research Online

<https://research-repository.griffith.edu.au>

A Meta-analysis of the Effects of Early Developmental Prevention Programs in At-Risk Populations on Non-Health Outcomes in Adolescence

Abstract

We present the results of a meta-analytic review of early developmental prevention programs (children aged 0-5: structured preschool programs, center-based developmental day care, home visitation, family support services and parental education) delivered to at-risk populations on non-health outcomes during adolescence (educational success, cognitive development, social-emotional development, deviance, social participation, involvement in criminal justice, and family well-being). This review improves on previous meta-analyses because it includes a more comprehensive set of adolescent outcomes, it focuses on measures that are psychometrically valid, and it includes a more detailed analysis of program moderator effects. Seventeen studies, based on eleven interventions (all U.S.-based) met the ten criteria for inclusion into the analysis. The mean effect size across all programs and outcomes was 0.313, equivalent to a 62% higher mean score for an intervention group than for a control group. The largest effect was for educational success during adolescence (effect size 0.53) followed by social deviance (0.48), social participation (0.37), cognitive development (0.34), involvement in criminal justice (0.24), family wellbeing (0.18), and social-emotional development (0.16). Programs that lasted longer than three years were associated with larger sample means than programs that were longer than one year but shorter than three years. More intense programs (those with more than 500 sessions per participant) also had larger means than less intense programs. There was a marginally significant trend for programs with a follow-through component into the early primary school years (e.g. preschool to Grade 3) to have more positive effects than programs without a follow-through. We conclude that the impact of well-conducted early development programs on quality of life in adolescence can be substantial for social policy purposes.

Keywords: early developmental prevention; meta-analytic review; at-risk populations; adolescent outcomes; quality of life.

1. Introduction

Longitudinal research has confirmed the benefits of well-designed and carefully implemented early developmental prevention (EDP) programs (Homel, 2005). Evidence from a series of interventions demonstrates that intervening with the child and the family early in the developmental pathway can minimize future health, educational, behavioral and crime-related problems (Farrington & Welsh, 2002). This is important given evidence demonstrating a deterioration in health (e.g. asthma, diabetes, obesity, intellectual disability, eating disorders, depression, attention deficit disorder, hyperactivity) and non-health outcomes (e.g. crime and educational success) over time for children, particularly from at-risk populations (Stanley, Richardson, & Prior, 2005).

This paper presents results of a meta-analytic review of the effectiveness in at-risk populations of EDPs involving children up to 5 years of age who do not have mental health or severe developmental problems. Outcomes encompass indicators of individual and family quality of life during adolescence, defined as ages 10-19 years (World Health Organisation, 1986). In the longitudinal literature we reviewed, at-risk populations were mainly socio-economically disadvantaged, operationalized as people with poor levels of education, living in areas of high unemployment, living in poverty according to local income standards, and perhaps isolated as a result of ethnicity and language.

Earlier systematic reviews and meta-analyses have focused on crime, educational, and family-related outcomes in the childhood and adolescent life phases. However, these studies are incomplete, as they do not include a full review of the salient outcomes in adolescence. Additionally, the methodology employed in earlier meta-analyses for defining quality of life outcome domains and their respective indicators in adolescence is undeveloped. To fill this gap, we carried out a meta-analysis that incorporated a structured and scientific method of identifying and analyzing important outcome domains and indicators. We focused on adolescence because of its importance both in terms of individual development and social investment. It is a phase of life when young people in developed countries are presented with many new opportunities for positive development, but also face “historically unprecedented

levels of physical, behavioral, and social health risks” (Lerner & Castellino, 2002, p. 122). In recognition of these risks and opportunities, much emphasis is placed by government on reducing negative outcomes during adolescence (e.g. school drop-out, formal contact with the criminal justice system). Our results are consistent with the argument that governments should invest more in the development and implementation of good EDPs to balance the more usual strategy of relying on costly remedial interventions (Manning, 2008).

The meta-analysis was carried out as the first part of a two stage process in a study conducted by Manning (2008). The second stage by Manning adapts the Analytical Hierarchy Process (Saaty, 1980) to develop a method for making complex multi-criteria decisions with respect to policy options for early childhood interventions. Nagin (2001) argued that a method was required that would permit analysts to identify common metric outcomes across competing and often disparate programs, such as home visitation to pregnant teenage women and centre-based developmental day care, with the goal of eliciting preferences and relative utility values. The second stage also provided an outline of how relative utility values, derived using the Analytical Hierarchy Process approach, may be used to identify the economic benefits of developmental prevention programs on non health-related quality of life outcomes in adolescence, as proposed by Nagin.

2. The Efficacy of Early Developmental Prevention Programs

Evaluations of the short- and long-term effects of EDPs including projects such as the Perry Preschool Project (Schweinhart, 2004), the Elmira Prenatal/Early Infancy Project (Eckenrode, et al., 1998; Olds, 2002) and the Seattle Social Development Project (Hawkins, Catalano, Kosterman, Abbot, & Hill, 1999) have confirmed a number of positive effects for children and their families. Effects are particularly evident for children who come from low-income backgrounds, with short-term gains in intellectual and academic achievement scores, and longer-term outcomes demonstrating successful educational performance, a reduction in behavioral problems and delinquency, and improved family wellbeing (Brooks-Gunn, Fuligni, & Berlin, 2003, p.5-9). Moreover, a systematic delivery of basic services to disadvantaged children and their families has demonstrated large reductions

in crime involvement amongst targeted groups (Reynolds, Ou, & Topitzes, 2004; Schweinhart, 2004; Schweinhart, Barnes, & Weikart, 1993; Yoshikawa, 1994; Zigler, Taussig, & Black, 1992). The evidence suggests that targeting interventions across multiple domains (e.g. families, schools and communities) results in improved educational outcomes, decreases in child maltreatment, reductions in child and youth antisocial behavior, lower levels of substance abuse, and increases in income and workforce participation (Brooks-Gunn, et al., 2003; Olds, 2002; Reynolds, Temple, Robertson, & Mann, 2001).

It is apparent that EDPs foster a wide range of benefits in childhood, adolescence, and beyond. However, previous meta-analyses of the benefits of EDPs for the entire adolescent life phase have been patchy, both in terms of the coverage of studies and the methods for identifying the salient domains and indicators in adolescence. In addition, previous studies have reported only limited analyses of the moderating effects of program characteristics such as duration and intensity. Knowing more about these effects is important for policy development and funding allocations.

A meta-analysis of preschool prevention programs for disadvantaged children and their families by Nelson, Westhues, Laurier and MacLeod (2003) focused on three broad outcomes: cognitive development, social-emotional development and parent/family well-being. Results revealed: "...cognitive impacts during the preschool period were greatest for those programs that had a direct teaching component in preschool", and "...cognitive impacts during the K-8 time period were greatest for those programs that had a follow-through educational component in elementary school" (Nelson, et al., 2003, p.1). This research also showed that the longer the child remained in the intervention, the greater the effect on cognitive and social-emotional development from Kindergarten through Grade 8; and the more intense the program, the greater the effects on cognitive outcomes during preschool and on family outcomes at Kindergarten through Grade 8. The largest effects on the three outcomes were for those programs that served predominantly African-American children.

The meta-analysis conducted by Farrington and Welsh (2003) reviewed the effectiveness of family-based prevention programs in reducing offending and antisocial

behavior by children and adolescents. Results demonstrated that programs that target family risk factors such as poor child-rearing, poor parental supervision, and inconsistent or harsh discipline reduced delinquency, particularly antisocial behavior. The least effective type of intervention with respect to delinquency was school-based programs. Home visiting, preschool, and multi-systemic therapy programs were generally effective at reducing offending and antisocial behavior.

A recent meta-analysis of early parent training programs conducted by Piquero, Farrington, Welsh, Tremblay and Jennings (2008) also focused exclusively on delinquency and anti-social behavior. Results showed that early family/parent training is effective in reducing behavioral problems in children through age five. Descriptive evidence provided by the authors also suggested that family/parent training is effective in reducing delinquency and crime into adolescence and through to early adulthood.

3. Aims

In this study we moved beyond previous meta-analytic reviews of EDPs by reanalyzing and extending the outcome domains identified by Nelson, Westhues, Laurier and MacLeod (2003), Farrington and Welsh (2003), and Piquero, Farrington, Welsh, Tremblay and Jennings (2008). We created seven comprehensive domains (cognitive development, educational/academic success, social-emotional development, deviance, social participation, criminal justice outcomes, and family well-being) that reflect the importance of both crime and non-crime adolescent outcomes for quality of life and individual developmental pathways. The seven outcome domains were selected by analyzing the psychometric properties of outcomes and their respective indicators related to individual or family quality of life during adolescence.

We incorporated results from five forms of EDP: structured preschool programs, center-based childcare/developmental day care, home visitation programs, family support services and parental education. We also included up-to-date results from current longitudinal research to review the impact of EDPs on those populations most at-risk throughout adolescence.

We aimed to answer two questions:

1. How much effect do EDPs have on the seven outcome domains during adolescence?
2. How do program characteristics moderate outcomes in adolescence?

Potential moderators include: type of program (e.g., the use of a structured preschool education component); program duration; program intensity (e.g., number of sessions); number of program components (e.g., home visiting plus enriched preschool education); and the use of a follow-through component (e.g., supplementary programs in the primary years to further support the family and child). In this article we investigate the effects of all these moderators except program type, which is the topic of a separate paper.

4. Method

In reviewing the effectiveness of EDPs on at-risk populations with respect to non health-related outcomes during adolescence, we employed a meta-analytic technique to synthesize research findings that incorporated similar outcome variables. We aimed to identify multiple dependent variables for the adolescent life phase and group these variables (based on an analysis of psychometric properties) into meaningful outcome domains, including their relevant indicators, and highlight the potential effectiveness of EDP programs on the seven resulting outcome domains using weighted effect sizes (d).

Given that no multi-attribute classification system of non health-related outcomes of EDPs exists, we drew (for the first time in the literature) on both psychometric test libraries and longitudinal research studies to identify salient domains, outcomes, and indicators. Data were then combined and analyzed for the adolescence life phase, including the transition from primary to high school and the transition from high school to post-secondary education or the workforce.

Ranges of unweighted effect sizes (d) for a given dependent variable are reported in Manning (2008). In this paper we removed outliers (defined as effect sizes at least two standard deviations from the mean), as proposed by Lipsey & Wilson (2000). Lipsey and Wilson argue

that extreme effect sizes may have disproportionate influences on statistics such as means and variances.

4.1. Selection of studies

Studies included in the meta-analysis satisfied the following criteria:

- a. The intervention began before the children started school;
- b. The focus of the intervention was on developing or enhancing child, parent-child, or family well-being;
- c. Intervention programs adopted either universal (e.g. all children in the at-risk population) or selected approaches (e.g. high-risk groups within an at-risk population);
- d. Intervention programs were not specifically aimed at treating children with mental health or severe developmental problems;
- e. A prospective design of at least Level 4 in design quality (i.e., a randomized or matched groups quasi-experimental design) according to the Maryland Scientific Methods Scale (Farrington, Gottfredson, Sherman & Welsh, 2002);
- f. The results of the research were reported in journal articles, books, or book chapters published in 2008 or earlier;
- g. Outcome measures collected and reported related to one or more of the seven outcome domains;
- h. There was at least one post-intervention follow-up, during the adolescent years;
- i. Effect sizes could be identified and calculated; and
- j. Interventions were directed at disadvantaged and/or low socio-economic populations.

Some well known and effective studies such as Healthy Families America (Daro & Harding, 1999), Parents as Teachers (Wagner & Clayton, 1999), the Incredible Years Program (Webster-Stratton, Rinaldi, & Reid, 2009), the Triple P Program (Sanders, Markie-Dadds, & Turner, 2003), the Home Instruction Program for Preschool Youngsters (HIPPY) (Baker, Piotkowski, & Brooks-Gunn, 1999), and the Early Start Program (Fergusson, Grant, Horwood, & Ridder, 2006) were not included in the analysis since they did not satisfy all the

criteria (many, for example, have not yet reported results for adolescence). Additionally, we did not include into this study papers relating to the economic evaluation of these programs. For example work conducted by the RAND corporation (e.g. Karoly, et al., 1998; Karoly, Kilburn, & Cannon, 2005; Kilburn & Karoly, 2008) and the work of James Heckman and colleagues (e.g. Heckman, Stixrud, & Urzua, 2006) who stressed the importance of the family in mediating the cognitive and social-emotional skills of children in their early years.

Studies that included results for a variety of risk factors and/or outcome domains were included in the meta-analysis. This is common practice for large meta-analyses, particularly those focused on primary and secondary intervention programs for children (Nelson, et al., 2003).

Our starting point was a manual search of key journals for the period 1970 to 2008: *Monographs of the Society for Research in Child Development, The Future of Children, American Educational Research Journal, Child Development, Applied Developmental Science, The Journal of Pediatrics, Journal of Early Intervention, Zero to Three, Journal of Applied Developmental Psychology, Developmental Psychology, The Elementary School Journal, Journal of the American Medical Association, Journal of Community Psychology, Evaluation Review, Journal of Consulting and Clinical Psychology, Early Childhood Research Quarterly, American Journal of Community Psychology, Trends and Issues: Australian Institute of Criminology, Canadian Journal of Community Mental Health, Child Abuse and Neglect, Journal of Primary Prevention, Pediatrics, Campbell Collaboration, and Journal of Prevention and Intervention in the Community.*

We then scanned relevant review articles, consulted the bibliographies of articles that met the selection criteria, and asked key researchers for assistance in identifying relevant articles. We searched ten electronic databases covering the years 1970 to 2008 (e.g. SAGE full text, CSA, Informat) using keywords such as prevention, early intervention, developmental intervention, preschool education, children, home visitation, and multi-component program. Efforts were made to track down any unpublished studies highlighted in the search. Authors of published articles were contacted to ascertain if additional results

relating to our key outcome domains were available. This significantly added to our list of potential studies, since some relevant programs could not be included in earlier meta-analyses due to a lack of critical published data. In all, over 5,200 abstracts were reviewed.

Our multi-faceted search methodology uncovered all of the studies identified in several narrative reviews of the literature (Brooks-Gunn, et al., 2003; Cohen & Radford, 1999; Hertzman & Wiens, 1996; Homel, 2005; Karoly, et al., 1998; MacMillan, MacMillan, Offord, Griffith, & MacMillan, 1994; Manning, 2004; McCain & Mustard, 1999; Mrazek & Brown, 2002; Nelson, Laurendeau, Chamberland, & Peirson, 2001; The Developmental Crime Prevention Consortium, 1999; Yoshikawa, 1994), as well as some not previously covered by any review. In addition, in some cases authors provided more detailed information about their data (e.g., standard deviations) so that effect sizes could be calculated.,

Eleven intervention programs (incorporating 17 follow-up studies) were identified which met all our criteria. In the few cases when both experimental and control groups received some form of preschool program only those studies where the intervention program participants received an additional component (e.g. a parenting program) were included in the analysis. Table 1 presents an overview of all interventions, highlighting relevant citations, study design, length and intensity of intervention, sample size for intervention [*En*] and control groups [*Cn*], and age of children at follow-ups.

Insert Table 1 about here

The eleven intervention programs in Table 1 represent a variety of forms of EDPs. These include structured preschool programs (SPP) that provide structured learning experiences through developmentally orientated programs in which children work with the preschool teacher to develop their language and cognitive skills (Coalition for evidence-based policy, 2005); center-based childcare/developmental day care programs (CBCC) that provide a stimulating and nurturing child care environment for infants and young children between the ages of 6 weeks to 5 years of age (Reynolds, 1994); home visitation programs (HV) that provide a variety of services to parents to reduce child abuse and neglect by parents who lack the skills to properly care for an infant or who do not have the social networks to support

them as new parents (Olds, Henderson, & Kitzman, 1994); family support services (FSS) that work in partnership with parents to aid in crisis situations to minimize the risk of child abuse and neglect and promote support to all family members with the goal of delivering the necessary support services to create strength, unity and improved functioning within the family (Children's Home Society and Family Services, 2007); and, parental education (PE), developed to improve core parenting skills (Sanders, 2003).

4.2. Outcome domains and coded variables

We created domains based on a review of the psychometric literature (e.g. Achenbach, 1991; Achenbach & Rescorla, 2000; Arnold, O'Leary, Wolff, & Acker, 1993; Caldwell & Bradley, 2001; Coopersmith, 1967, 1975; Corle, et al., 2001; Eyeberg & Pincus, 1999; Halpern, Baker, & Piotrkowski, 1993; McCarthy, 1972; Medley & Klein, 1957; Piotrkowski, Botsko, & Matthews, 2000; Richman & Graham, 1971) and longitudinal research (e.g. Campbell & Ramey, 1994, 1995; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002). This was done in order to include dependent variables that may have been defined and/or measured differently across studies and to provide a more reliable picture of the outcomes associated with the seven domains. Table 2 provides a description of the domains and their relevant indicators identified in the psychometric literature and longitudinal research. Manning (2008) provides more details of the analysis of the psychometric literature.

Insert Table 2 about here

Initially more fine-grained outcome domains were considered for adoption. However, because data were limited across interventions it was decided to use a smaller set of consolidated domains. The seven domains are sufficiently wide-ranging to cover the major outcomes reported in the literature, are conceptually distinct and meaningful, and provide the basis for a meta-analysis that is more comprehensive than any reported previously. Care was taken to ensure that dependent variables allocated to each domain remained consistent with the psychometric literature (Nelson, et al., 2003).

Fourteen variables were coded for each study. These included program characteristics (e.g. type of program, program components, length and intensity), participant characteristics (e.g. age of participant, ethno-racial background), and study characteristics (e.g. randomized design, number of outcome measures). The length of a program was coded in weeks (assuming 9 months of school per year and 4.3 weeks per month), and intensity was coded according to the intended number of sessions for both parents and children. This approach was adopted since most studies did not report actual attendances at each session. Thus intensity was computed as the total number of sessions available to a participant. A session was defined as a planned activity lasting up to half a day (e.g. structured preschool program) or a single home visit with a parent. The length of the intervention for the child was coded as more than or less than one year, while intensity was coded as less than or equal to 300 sessions, greater than 300 but less than or equal to 500 sessions, and greater than 500 sessions. Given that data came from longitudinal research, typically there was more than one paper for each intervention. Thus, multiple sources were used for coding each intervention.

4.3. Effect size protocols

We used a six-step approach, similar to a method proposed by Lipsey and Wilson (2000):

- (1) calculation of individual effect sizes (d) and corresponding variances for each variable from each study;
- (2) calculation of weighted mean effect sizes for each variable (\underline{d});
- (3) calculation of 95% confidence intervals (CIs) surrounding those effect sizes;
- (4) calculation of median effect sizes;
- (5) calculation of the Q statistic to assess the heterogeneity of the effect sizes for each outcome construct;
- (6) re-calculation of weighted mean effect sizes and 95% CIs based upon the extent of heterogeneity in the effect size distributions.

To determine the relative magnitude of the experimental effect, Cohen's d was used. d is more popular than alternative measures such as g (Hedge, 1981), and so is useful for

comparing a growing number of studies (Durlak & Lipsey, 1991). Moreover, based on Cohen's (1992) proposition that effect sizes of 0.20 are small, 0.50 are medium, and 0.80 are large, one can compare effect sizes to established benchmarks. However, it must be acknowledged that Cohen's descriptions of effect sizes may be misleading. Effect sizes in the human services and crime prevention literatures, for example, are often in the 0.20-0.30 range (e.g., Farrington & Welsh, 2003), yet these effects correspond to substantial and often cost-effective benefits in the population (Aos, Phipps, Barnoski & Lieb, 2001). Even widely used and highly cost-effective treatments in medicine, such as low dose aspirin, often have small effect sizes (e.g., a reduction of 21% in the risk of a cardiovascular event for patients with stable cardiovascular disease: Berger, Brown and Becker, 2008).

When means and standard deviations were reported in the reviewed studies, d was calculated by subtracting the mean score (\bar{X}_2) (post-test follow-up) for the comparison group from the mean score of the intervention group (\bar{X}_1) and dividing the result by the pooled standard deviation (S_{pooled}) (Equation 1). N_1 is the number of participants in the intervention group, N_2 is the number of comparison group members, SD_1 is the standard deviation of the score for the intervention group, and SD_2 is the standard deviation of the score for the comparison group.

$$ES_{\text{sm}} = \frac{\bar{X}_2 - \bar{X}_1}{S_{\text{pooled}}} \quad (1)$$

$$\text{where } S_{\text{pooled}} = \sqrt{\frac{(N_1 - 1)(SD_1)^2 + (N_2 - 1)(SD_2)^2}{N_1 - 1 + N_2 - 1}}$$

When means and standard deviations were not reported, effect sizes were calculated from test statistics (e.g. t-tests, F-ratio, frequencies, odds ratios and χ^2 tests for 2 x 2 contingency tables) using formulas outlined by Lipsey and Wilson (2000, p.199-202).

Studies with relatively small sample sizes (e.g. < 30 per group) generally have effect sizes that are upwardly biased (Durlak & Lipsey, 1991). Consequently, all d s were corrected for small sample bias across all outcome constructs using a formula provided by Hedges

(1981). In Equation 2, n is the total sample size and ES'_{sm} is the unbiased standardized mean difference.

$$ES'_{sm} = ES_{sm} \left[1 - \frac{3}{4n - 9} \right] \quad (2)$$

The variance (Var) of d was then calculated using Equation 3, where (n) are the sample sizes for each group and d is the mean:

$$Var d = \left[\frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)} \right] \quad (3)$$

Following the calculation of individual effect sizes and variances (extracted from the primary studies), a weighted mean effect size (\underline{d}) could be calculated using Hedges and Olkin's (1985) formula, where k is the number of findings, $w_i = 1/v_i$, and v_i is the variance of the individual effect size (Lipsey & Wilson, (2000, pp. 113-114) (Equation 4).

$$\underline{d} = \frac{\left(\sum_{i=1}^k w_i d_i \right)}{\left(\sum_{i=1}^k w_i \right)} \quad (4)$$

The variance of the weighted mean effect size could then be calculated using Equation 5:

$$Var(\underline{d}) = \frac{1}{\left(\sum_{i=1}^k w_i \right)} \quad (5)$$

This variance was used to calculate confidence intervals using Equation 6:

$$\left[\underline{d} - 1.96 \sqrt{Var(\underline{d})}, \underline{d} + 1.96 \sqrt{Var(\underline{d})} \right] \quad (6)$$

Another method proposed by Lipsey and Wilson (2000) to examine the reliability of \underline{d} is to calculate a median effect size. The general rule of thumb is that the closer the weighted \underline{d} and the median \underline{d} are to each other, the more reliable the effect size estimate. Nelson, Westhues and MacLeod (2003) argue that calculating a weighted mean effects size for each predictor variable provides a more reliable estimate of the effect size for a given outcome construct at a given period of time, as opposed to selecting one specific effect size.

Next, the homogeneity of the seven outcome constructs was tested. The Q-statistic allows for a statistical evaluation of the variation in the distribution of effect sizes for a given outcome construct (Nelson et al., 2003), and it also facilitates the identification of individual effect sizes that may be considered outliers. According to Lipsey and Wilson (2000), a particular finding should be considered an outlier if: (a) it is an extreme value (highest or lowest); (b) the Q statistic is significant; and, (c) if the single finding accounts for more than 50% of the value of the Q statistic.

A fixed effects model is normally used in the first stage of the analysis, and if Q is found to be significant the analysis is re-run using a random effects model. Lipsey and Wilson (2000) argue that a fixed model assumes that variability between effects sizes is due to sampling error, whilst a random effects model assumes that the variability between effects sizes is due to sampling error plus the variability in the population of effects (i.e. each study is estimating a slightly different population effect size). The fixed effects model weights each study by the inverse of the sampling variance (Equation 7),

$$W_i = \frac{1}{se_i^2} \quad (7)$$

while the random effects model weights each study by the inverse of the sampling variance plus a constant that represent cross-population effects variability (Equation 8),

$$W_i = \frac{1}{se_i^2 + \hat{v}_e} \quad (8)$$

where the random effects constant \hat{v}_e is calculated using Equation 9:

$$\hat{v}_e = \frac{Q_T - (k - 1)}{\sum w_i - \left(\frac{\sum w_i^2}{\sum w_i} \right)} \quad (9)$$

Equation 10 can then be used to calculate the Q-statistic for each outcome construct.

$$Q = \sum_{i=1}^k w_i (d_i - \underline{d})^2 \quad (10)$$

In Equations 9 and 10, k is the number of findings (effect sizes), w_i is the inverse variance weight of the individual effect size for each finding, d_i is the individual effect size,

and \bar{d} is the weighted mean effect size. The formula for Q estimates the sum of the squared deviations from \bar{d} , with each component weighted by its inverse variance. The significance of Q is assessed using the chi-square distribution with $k-1$ degrees of freedom (Nelson, et al., 2003).

If the Q -statistic is significant for a particular construct, the finding contributing the largest amount of variance to the overall Q -statistic can be removed. The weighted mean effect size can then be recalculated, and the Q -statistic recalculated based on the new \bar{d} . This process continues until the Q -statistic is no longer significant or the number of studies being used to calculate the variable falls below two, in which case it is removed from the meta-analysis.

Finally, provided the effect sizes were heterogeneous we examined possible moderating features of these programs (e.g. number and type of program components, program intensity, program duration, and the presence of follow-through programs, such as those that continued into early primary school). We analyzed these potential moderating factors using an analogue to ANOVA and weighted least squares regression (random effects). The analogue to the ANOVA uses Q to test the ability of a single categorical variable (e.g. intervention type) to explain excess variability in a distribution of effect sizes (Lipsey & Wilson, 2000). We used the software package Comprehensive Meta-Analysis ("Comprehensive Meta-Analysis," 1999) to run all moderator analyses.

4.4. Fail-safe n

Meta-analyses are vulnerable to bias given that studies that demonstrate non-significant findings rarely get published (Hedges & Olkin, 1985). Failure to include these unpublished studies in a meta-analysis may result in biased findings. To circumvent this dilemma, Rosenthal (1979) developed a method known as the 'fail-safe n .' The formula was developed to estimate the number of unpublished studies reporting null results that would be required to reduce the cumulative effect across the studies to a point of non-significance. Orwin (1983) adapted Rosenthal's method "...to the standardized mean effect size" (Lipsey

& Wilson, 2000, p.166). Orwin's method determines the number of studies required to reduce an effect size to a specified level (Equation 11),

$$k_o = k \left[\frac{\overline{ES}_k}{ES_c} - 1 \right] \quad (11)$$

where k_o is the number of effect sizes with the value of zero required to reduce the mean weighted effect size to ES_c , k is the number of studies, and \overline{ES}_k is the weighted mean effect size. We incorporated this measure into our study to eliminate, or at least reduce, the 'file drawer' dilemma, using a value of $ES_c = 0.20$ (i.e. a low effect size in Cohen's terms).

3. Results

Seventeen studies based on the 11 independent intervention projects in Table 1 met our selection criteria. Table 3 provides information on program characteristics. EDP programs that contained a structured preschool program (SPP) comprised 64% of the 17 studies. More than 45% of studies incorporated a home visitation (HV) component and a family support (FSS) component. More than a third (36%) included a center-based childcare/developmental day care component, while only 9% incorporated parent education (PE).

Insert Table 3 about here

Table 4 provides information about program participants and the study characteristics. Over 70% of participants who made up both control and experimental groups in EDP programs were from African-American backgrounds, about two thirds of the studies (65%) used a randomized design, and over 80% had sample sizes greater than 300.

Insert Table 4 about here

5.1. Mean effect sizes and Q statistics

Figure 1 presents the mean effect sizes. The mean weighted effect sizes (\underline{d}) for the outcome domains criminal justice (CJ) and family wellbeing (FW) were 0.243 and 0.178 respectively. Mean effect sizes for cognitive development (CD), social participation (SP) and deviance (D) were somewhat larger, at 0.339, 0.371 and 0.481 respectively. The largest effect size was for educational success (ES, 0.528), which in Cohen's terms represents a medium to

medium-high effect. The smallest effect was for social-emotional development (SED, 0.157). The overall weighted mean effect size across all domains was $\underline{d} = 0.313$, $P < 0.0001$. This is equivalent to a 62% higher mean for an intervention group than for a control group (Lipsey & Wilson, 2000, p.153, Table 8.1).

Insert Figure 1 about here

Effect sizes were adjusted for sample size. Three outliers (in the ES domain) were identified and removed from the analysis. The outliers came from the Chicago Child-Parent Center Study, Studies 4 and 5. Had these studies been included, the weighted mean effect size for the ES domain would have been reduced to $\underline{d} = 0.154$, which may have increased the standard errors or created sample bias. The overall Q test for homogeneity was not significant at $\alpha = .05$ for all outcome domains (Table 5), so in all cases the null hypothesis of homogeneity was not rejected.

Insert Table 5 about here

Figures 2-8 present results of the analyses for the seven outcome domains. Included are effect sizes (\underline{d}), lower and upper 95% confidence limits, p values, and Forrest plots. There were no outliers.

Insert Figures 2-8 about here

5.2. Moderator analyses

We were interested in understanding the effects of the number of components in an intervention, as well as the effects of varying levels of program intensity and duration. In addition we investigated whether programs that incorporated a follow-through component into the early primary school produced additional benefits in adolescence.

There were no significant differences ($Q = .129$; $P = .937$) between programs with one component ($k = 6$; $\underline{d} = 0.440$), and those with two components ($k = 23$; $\underline{d} = 0.441$) or three or more components ($k = 8$; $\underline{d} = 0.421$). Program intensity was coded according to the intended number of sessions for both parents and children. There was a significant difference ($Q = 11.883$; $P < 0.001$) between programs with fewer than 500 sessions ($k = 19$, $\underline{d} = 0.283$), and programs with more than 500 sessions ($k = 21$, $\underline{d} = 0.494$) (Figure 9). Specifically, for

the educational success indicators we found a significant difference ($Q = 9.505$; $P < 0.01$) between programs with more than 500 sessions ($k = 7$; $\underline{d}_i = 0.592$) and those with fewer than 500 sessions ($k = 4$; $\underline{d}_i = 0.214$).

Insert Figure 9 about here

EDP programs were divided into three categories with respect to duration: up to one year; longer than one year but less than three years; and longer than three years. Insufficient data were available to compare programs that operated for less than one year with other categories. A significant difference ($Q = 9.032$; $P < 0.01$) was found between programs longer than three years ($k = 20$; $\underline{d}_i = 0.479$) and those whose duration was greater than one year but less than three years ($k = 20$; $\underline{d}_i = 0.308$) (Figure 10). Moreover, we found a significant difference ($Q = 6.080$; $P < 0.01$) with respect to educational success between programs longer than three years ($k = 6$; $\underline{d}_i = 0.562$) and programs with duration longer than one year but less than three years ($k = 5$; $\underline{d}_i = 0.349$).

Insert Figure 10 about here

Finally, our moderator analysis revealed that programs that had a follow-through component (e.g. post intervention or supplementary programs to further support the family and child) tended to produce better results ($k = 8$; $\underline{d}_i = 0.516$) than programs that did not have a follow-through component ($n = 31$; $\underline{d}_i = 0.368$), but between class differences were only significant at the 10% level ($Q = 2.648$; $P = .104$). This partly reflects the small number of studies with a follow-through component.

5.3. Fail-safe n

The fail-safe n for 17 studies with an overall weighted mean of 0.313 is 13. This means that another 13 studies with non-significant findings would be required to reduce the weighted mean effect size to 0.20. Given the thoroughness of our methods (Section 4.1), we consider it unlikely that there are 13 missing longitudinal studies with null results.

6. Discussion

This meta-analysis demonstrates that early developmental prevention programs (EDPs) have positive effects on individual and family wellbeing into adolescence. The overall

size of the effects ($d = 0.313$, $P < 0.0001$) is in the small to medium range according to Cohen, but since this mean result is equivalent to a 62% higher mean for an intervention group than for a control group, the average or expected impact can be regarded as substantial. This overall effect is derived from a broad range of psychometrically valid indicators across seven major life domains, and reflects the impact of a mix of intervention strategies, including those with a substantial emphasis on parent or family support and training as well as on child-focused initiatives. This suggests, in contrast to the position of Cohen and Radford (1999), that programs directed towards families can have a wide range of beneficial effects on children, as well as those that provide direct developmental services to children (Freiberg et al., 2005; Homel et al., 2006).

Outcome domain effect sizes ranged from 0.53 for educational success to 0.16 for social-emotional development and 0.18 for family wellbeing. The substantial impact on educational success reflects, at least in part, the emphasis of many of the early EDP programs on preschool education and on relatively easy to measure indicators of academic achievement and cognitive development. Seven of the 11 programs in our sample incorporated a preschool component (the most common single characteristic), and nearly all studies included a range of outcomes related to school success and cognitive development. Perhaps even more important, however, is the fact that academic achievement in adolescence builds strongly on earlier success in the school system, and conversely early difficulties can become entrenched in the later years (Hertzman & Wiens, 1996; The Developmental Crime Prevention Consortium, 1999). The emphasis of many of the EDPs on school achievement combined with 'system continuity' help explain the large effect sizes for educational success, and similarly for cognitive development. Our mean effect size for cognitive development was 0.34, which is actually slightly higher than the figure of 0.30 reported by Nelson and colleagues (2003) for children from Kindergarten to Grade 8, suggesting that the effects of EDPs on cognitive development may persist from the primary years through adolescence.

In contrast to cognitive development and school success, discontinuities in environmental circumstances may - especially for disadvantaged populations - help to explain

the low effect sizes for social-emotional development and family wellbeing. Both these outcome domains are likely to be strongly influenced by changing family circumstances and proximal events, as well as by developmental processes in adolescence that diminish the effects of early in life programs unless reinforced by further preventive interventions in the primary or high school years. This is supported by the findings of Nelson and colleagues (2003), who reported a mean effect size of 0.30 for family wellbeing for families with children up to Grade 8 (compared with our figure of 0.18). There is evidence that a decline in family wellbeing during later adolescence is fairly common, reflecting such processes as changes in family cohesion and in adolescent adaptability to intimacy (e.g. Baer, 2004; Richmond & Stocker, 2007; Roming & Bakken, 1992).

Apart from adolescence-specific developmental processes, a likely explanation for our low figure for social-emotional development is that Nelson and colleagues operationalized this domain in broader terms, incorporating items such as parent and teacher ratings of children's behavior, social skills, self reported self-esteem, grade retention, placement in special education classes, teenage employment, educational success, and criminal behavior. Many of the indicators used by Nelson and colleagues had higher effect sizes in our study but appear in different outcome domains.

It is noteworthy that after educational success the EDPs in our study had their greatest effect on deviance (mean effect size 0.48), followed by the closely related domain of social participation (0.37). These results are broadly in line with the recent meta-analyses by Farrington and his colleagues (Farrington & Welsh, 2003; Piquero et al., 2008), and confirm the importance of EDPs for crime prevention policies (The Developmental Crime Prevention Consortium, 1999), notwithstanding the smaller effect size for criminal justice outcomes (0.24). This latter figure probably reflects the relative infrequency of formal contacts with police and the courts for adolescents, compared with the much more common self-reported delinquencies listed in Table 2.

6.1. Moderator analyses

The moderator analyses we have reported shed important light on how EDPs achieve their effects, although the small number of studies in our sample and the limited information available about program implementation (such as the number of sessions actually attended) make it difficult to draw firm conclusions. Results were clearest for program intensity and duration, where it seems that, in a nutshell, more is better. This finding is consistent with previous research by Nelson et al. (2003) and Ramey & Landesman Ramey (1998), and provides support for recent moves in the UK and Australia to commit substantial long-term government resources to community-based EDPs in disadvantaged communities (Edwards et al., 2009; Melhuish et al., 2008).

Unfortunately the evidence from our study for the ‘more is better’ argument is much shakier when it comes to the number of program components and the efficacy of follow-through programs in the years after preschool. Because so much happens to families over time, it is likely that the marginally significant trend that we found toward additional positive effects for programs with a follow-through component will become stronger when more studies report results for interventions that continue through the primary and even high school years. Nelson and his colleagues found that follow-through was a strong mediator of educational success from Kindergarten to Grade 8, and in narrative reviews McLoyd (1998) and Ramey & Landesman Ramey (1998) have argued that follow-through programs with an educational component can build upon preschool education.

Although our analysis did not find evidence to support the assertion that the more program components an EDP has, the greater the effect in the adolescent years, both developmental systems theory (Lerner & Castellino, 2002) and a large number of empirical studies point to the fundamental importance of multi-component programs that operate simultaneously in a range of developmentally relevant settings (e.g., home, school, community). As argued by Febbraro (1994), Reynolds et al., (2001), and Nelson et al. (2003), multi-component programs are critical in providing a network of support for families and children, and comprehensive, multi-component EDP programs have been shown to have positive effects on children’s social-emotional, educational, and cognitive development, and

on family well-being (Nelson, et al., 2001; Weissberg & Greenberg, 1998; Yoshikawa, 1994; Zigler, et al., 1992).

6.2. Limitations of this study

A major limitation of this study (shared with its predecessors) was the shortage of good quality longitudinal evaluations of EDPs with outcomes into adolescence. The small sample size reduced the power of many of the analyses, and particularly affected our ability to draw conclusions about the effects of program design and features of implementation, although it does seem clear that programs with extensive client contact over a period of at least three years produced the best results. It is important that future longitudinal evaluation studies report more information about program philosophies (e.g., a strengths orientation founded on respectful relationships and cultural sensitivity) and the complex patterns, duration, frequency, and quality of contacts with clients. This information, as well as more fine-grained data about the clients themselves, will facilitate a move from ‘black box’ models of impact to more nuanced and contextualized analyses of the kinds of resources and activities that best fit different population groups in a variety of situations.

For the international research and policy communities, a particular concern is that most of the high quality interventions and evaluations (and all of the programs in this study) emanate from the United States. The problem, as Kamerman (2000) puts it, is that the US does the best research but has amongst the worst child policies in the developed world. Moreover, in the context of very high income inequality the US has amongst the worst outcomes for children and young people (Wilkinson & Pickett, 2009). These particular features of US society and culture make it risky to apply the results of the EDPs in this study to other countries, although the recent positive evaluations of Sure Start in the UK and Communities for Children in Australia (Edwards et al., 2009; Melhuish et al., 2008) suggest that generalization to other English-speaking countries may be quite feasible.

7. Conclusion

In this paper we have attempted to inform the choices faced by decision makers interested in moving government policies in the direction of primary, long term prevention by

presenting a rigorous analysis of the best available evidence on the effects of early developmental programs on adolescent outcomes. We have gone beyond previous meta-analyses in that we have developed a more comprehensive set of adolescent outcomes, we have focused on measures that are psychometrically valid, and we have included a more detailed analysis of program moderator effects. In this way we have attempted to simulate the policy environment as realistically as the evidence permits.

Claims that EDP programs are good evidence-based policy are, based on our findings, quite legitimate. Not only do the results of this analysis and previous meta-analyses provide clear evidence of many individual and societal benefits, EDPs have also been demonstrated to be a good investment from an economic perspective (see Greenwood, et al., 2001; Karoly, et al., 1998; Olds, Henderson, Phelps, Kitzman, & Hanks, 1993). What is needed now is an enlarged evidence-base that includes more long-term experiments as well as rigorous evaluations and cost-benefit analyses of large-scale programs in countries outside the United States, and that also incorporates detailed information about program philosophies, modes of implementation, and characteristics of client populations.

References

- Achenbach, T. M. (1991). *Manual for the child behaviour checklist/4-18 and 1991 profile*. Burlington, VT: Dept. of Psychiatry, University of Vermont.
- Achenbach, T. M. and Rescorla, L. A. (2000) *Child Behavior Checklist for ages 6-18*, Psychological Association Resources, Inc., Florida.
- Arnold, D. S., O'Leary, S. G., Wolff, L. S. and Acker, M. M. (1993) The Parenting Scale: A measure of dysfunctional parenting in discipline situations, *Psychological Assessment*, 5, 137-144.
- Baer, J. (2004) Is family cohesion a risk of protective factor during adolescent development?, *Journal of Marriage and Family*, 64, 668-675.
- Baker, A. J. L., Piotkowski, C. S. and Brooks-Gunn, J. (1999) The Home Instruction Program for Preschool Youngsters (HIPPY), *The Future of Children*, 9, 116-133.
- Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, S. W., Epstein, A. S. and Weikart, D. P. (1984) *Changes lives: The effects of the Perry Preschool Program on youths through age 19*, High/Scope Educational Research Foundation, Ypsilanti, MI.
- Brooks-Gunn, J., Fuligni, A. S. and Berlin, L. J. (2003). *Early child development in the 21st century : Profiles of current research initiatives*. New York; London: Teachers College Press.
- Caldwell, B. M. and Bradley, R. H. (2001). *HOME Inventory administration manual*. Arkansas: University of Arkansas.
- Campbell, F. A. and Ramey, C. T. (1994) Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families, *Child Development*, 65, 684-698.
- Campbell, F. A. and Ramey, C. T. (1995) Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention, *American Educational Research Journal*, 32, 743-772.
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J. and Miller-Johnson, S. (2002) Early childhood education: Young adult outcomes from the Abecedarian Project, *Applied Developmental Science*, 6, 42-57.
- Children's Home Society and Family Services. (2007). Child and family support services. Retrieved 20 April 2007, from http://www.chsfs.org/Child_Family_Support_Services.html
- Coalition for evidence-based policy. (2005). Perry Preschool Project (High quality preschool for children from disadvantaged backgrounds). 2006, from <http://www.evidencebasedprograms.org/Default.aspx?tabid=32>
- Cohen, N. and Radford, J. (1999). *The impact of early childhood intervention on later life*. Toronto: Hincks-Dellcrest Institute.
- . Comprehensive Meta-Analysis (Version 1.0.25). (1999). Englewood, New Jersey: Biostat, Inc.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco: Freeman, W.H.
- Coopersmith, S. (1975). *Coopersmith Self-Esteem Inventory: Technical manual*. Palo Alto, CA: Consulting Psychologists Press.
- Corle, D. K., Sharbaugh, C., Mateski, D. J., Coyne, T., Paskett, E. D., Cahill, J., et al. (2001) Self-rated quality of life measures: Effects of change to a low-fat, high-

- fiber, fruit and vegetable enriched diet, *Annals of Behavioural Medicine*, 23, 198-207.
- Daro, D. A. and Harding, K. A. (1999) Health Families America: Using research to enhance practice, *The Future of Children*, 9, 152-176.
- Durlak, J. A. and Lipsey, M. W. (1991) A practitioner's guide to meta-analysis, *American Journal of Community Psychology*, 19, 291-332.
- Eckenrode, J., Ganzel, B., Henderson, C. R., Smith, E., Olds, D., Powers, J., et al. (2000) Preventing child abuse and neglect with a program of nurse home visitation, *JAMA*, 284, 1385-1391.
- Eckenrode, J., Olds, D., Henderson, C. R., Kitzman, H., Luckey, D., Pettitt, L. M., et al. (1998) Long-term effects of nurse home visitation on children's criminal and anti-social behaviour, *Journal of the American Medical Association*, 280, 1302.
- Eyeberg, S. and Pincus, D. (1999). *Eyeberg Child Behavior Inventory & Sutter-Eyberg Student Behavior Inventory-Revised*. Odessa, FL: Psychological Assessment Resources.
- Farrington, D. and Welsh, B. C. (2002) Family-based crime prevention, In *Evidence-based crime prevention*, (Eds) L. W. Sherman, D. Farrington, B. C. Welsh and D. Layton MacKenzie, Routledge, London, pp. 22-55.
- Farrington, D. and Welsh, B. C. (2003) Family-based prevention of offending: A meta-analysis, *The Australian and New Zealand Journal of Criminology*, 36, 127-151.
- Febbraro, A. (1994) Single mothers "at risk" for child maltreatment: An appraisal of person centred interventions and a call for emancipatory action, *Canadian Journal of Community Mental Health*, 13, 47-60.
- Fergusson, D., Grant, H., Horwood, L. and Ridder, E. (2006) Randomized trial of the Early Start Program of home visitation: Parent and Family Outcomes, *Pediatrics*, 117, 781-784.
- Gray, S. W. and Klaus, R. A. (1970) The Early Training Project: A seventh-year report, *Child Development*, 41, 909-924.
- Greenwood, P., Karoly, L., Everingham, S., Hoube, J., Kilburn, M., Rydell, C., et al. (2001) Estimating the costs and benefits of early childhood interventions, In *Costs and benefits of preventing crime*, (Eds) B. C. Welsh, D. Farrington and L. Sherman, Westview Press, Colorado, pp. 123-148.
- Halpern, R., Baker, A. J. L. and Piotrkowski, C. S. (1993). *The Child Classroom Adaptation Inventory*. New York: National Council of Jewish Women.
- Hawkins, D. J., Catalano, R., Kosterman, R., Abbot, R. and Hill, K. (1999) Preventing adolescent health-risk behaviours by strengthening protection during childhood, *Archives of Pediatrics and Adolescent Medicine*, 153, 226-234.
- Heckman, J., Stixrud, J. and Urzua, S. (2006) The effects of cognitive and non-cognitive abilities on labour market outcomes and social behaviour, *Journal of Labour Economics*, 24, 411-482.
- Hedges, L. and Olkin, I. (1985). *Statistical methods for meta-analysis*. Toronto: Academic Press, Inc.
- Hedges, L. V. (1981) Distribution theory for galss's estimator of effect size and related estimators, *Journal of Educational Statistics*, 6, 107-128.
- Hertzman, C. and Wiens, M. (1996) Child development and long-term outcomes, *Social Science and Medicine*, 43, 1083-1095.

- Homel, R. (2005) Developmental crime prevention, In *Handbook of crime prevention and community safety*, (Eds) N. Tilley, Willan Publishing, Cullompton, Devon, U.K, pp. 71-106.
- Johnson, D. L. (2006) Parent-child development center follow-up project: Child behaviour problem results, *The Journal of Primary Prevention*, 27, 391-407.
- Johnson, D. L. and Blumenthal, J. B. (2004) The parent-child development centres and school achievement: A follow-up, *The Journal of Primary Prevention*, 25, 195-209.
- Karoly, L., Greenwood, P., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., et al. (1998) *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*, RAND, Santa Monica, CA.
- Karoly, L., Kilburn, M. and Cannon, J. (2005) *Early childhood interventions: Proven results, future promise*, RAND.
- Kilburn, M. and Karoly, L. (2008) *The economics of early childhood policy: What the dismal science has to say about investing in children*, RAND Labor and Population, Pittsburgh, PA.
- Lally, R. J., Mangione, P. L. and Honig, A. S. (1988) The Syracuse University Family Development Research Program: Long-range impact on an early intervention with low-income children and their families, In *Parent education as early childhood intervention: Emerging directions in theory, research and practice*, (Eds) D. R. Powell, Ablex Publishing Corporation, New Jersey, pp. 79-104.
- Lazar, I. and Darlington, R. (1982) Lasting effects of early education: A report from the consortium for longitudinal studies, *Monographs of the Society for Research in Child Development*, 47, 1-151.
- Levenstein, P., Levenstein, S., Shiminski, J. A. and Stolzberg, J. E. (1998) Long-term impact of a verbal interaction program for at-risk toddlers: An exploratory study of high school outcomes in a replication of the Mother-Child Home Program, *Journal of Applied Developmental Psychology*, 19, 267-285.
- Lipsey, M. W. and Wilson, D. B. (2000). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: SAGE Publications.
- MacMillan, H. L., MacMillan, J. H., Offord, D. R., Griffith, L. and MacMillan, A. (1994) Primary prevention of child physical abuse and neglect: A critical review, *Journal of Child Psychology and Psychiatry*, 35, 835-856.
- Manning, M. (2004). *Measuring the costs of community-based developmental prevention programs in Australia*. Unpublished Masters (Hons), Griffith University, Brisbane.
- Manning, M. (2008). *Economic evaluation of the effects of early childhood intervention on adolescent outcomes*. Unpublished PhD, Griffith University, Brisbane.
- McCain, M. N. and Mustard, F. J. (1999) *Early years study: Reversing the real brain drain*, Publications Ontario, Toronto.
- McCarthy, D. A. (1972). *Manual for the McCarthy scales of children's abilities*. New York: Psychological Corporation.
- McLoyd, V. C. (1998) Children in poverty: Development, public policy, and practice, In *Handbook of child psychology*, (Eds) W. Damon, I. E. Sigel and K. A. Renninger, John Wiley and Sons, New York
- Medley, D. M. and Klein, A. A. (1957) Measuring classroom behaviour with a pupil-reaction inventory, *The Elementary School Journal*, 57, 315-319.

- Meyer, L. A. (1984) Long-term academic effects of the Direct Instruction Project Follow Through, *The Elementary School Journal*, 84, 380-394.
- Miller, L. and Bizzell, R. P. (1983) Long-term effects of four preschool programs: Sixth, seventh, and eighth grades, *Child Development*, 54, 727-741.
- Mrazek, P. J. and Brown, C. H. (2002) *Final report: An evidence-based literature review regarding outcomes in psychosocial prevention and early intervention in young children*, Invest in Kids Foundation, Toronto.
- Nagin, D. (2001) Measuring the economic benefits of developmental prevention programs, In *Crime and justice: A review of research*, (Eds) M. Tonry, The University of Chicago Press, Chicago, pp. 347-384.
- Nelson, G., Laurendeau, M., Chamberland, C. and Peirson, L. (2001) A review and analysis of programs to promote family wellness and prevent the maltreatment of preschool and elementary school aged children, In *Promoting family wellness and preventing child maltreatment: Fundamentals for thinking and action*, (Eds) I. Prilleltensky, G. Nelson and L. Peirson, University of Toronto Press, Toronto, pp. 220-272.
- Nelson, G., Westhues, A., Laurier, W. and MacLeod, J. (2003) A meta-analysis of longitudinal research on preschool prevention programs for children, *Prevention and Treatment*, 6, 1-35.
- Olds, D. (2002) Prenatal and infancy home visiting by nurses: From randomized trials to community replication, *Prevention Science*, 3, 153-172.
- Olds, D., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., et al. (1998) Long-term effects of nurse home visitation on children's criminal and antisocial behaviour, *Journal of the American Medical Association*, 280, 1238-1244.
- Olds, D., Henderson, C. R. and Kitzman, H. (1994) Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life?, *Pediatrics*, 93, 89-98.
- Olds, D., Henderson, C. R., Phelps, C., Kitzman, H. and Hanks, C. (1993) Effects of prenatal and infancy nurse home visitation on government spending, *Medical Care*, 31, 155-174.
- Orwin, R. G. (1983) A fail-safe N for effect size in meta-analysis, *Journal of Educational Statistics*, 8, 157-159.
- Piotrkowski, C. S., Botsko, M. and Matthews, E. (2000) Parents' and teachers' beliefs about childrens school readiness in a high-need community, *Early Childhood Research Quarterly*, 15, 537-558.
- Piquero, A. R., Farrington, D., Welsh, B., Tremblay, R. E. and Jennings, W. G. (2008) *Effects of early family/parent training programs on antisocial behaviour and delinquency*, U.S Department of Justice, Washington.
- Ramey, C. T. and Landesman Ramey, S. (1998) Early intervention and early experience, *American Psychologist*, 53, 109-120.
- Reynolds, A. J. (1994) Effects of a preschool plus follow-on intervention for children at-risk, *Developmental Psychology*, 30, 787-804.
- Reynolds, A. J., Ou, S.-R. and Topitzes, J. W. (2004) Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centres, *Child Development*, 75, 1299-1328.
- Reynolds, A. J., Temple, J. A., Robertson, D. L. and Mann, E. A. (2001) Long-term effects of an early childhood intervention on educational achievement and

- juvenile arrest: A 15 year follow-up of low income children in public school, *Journal of the American Medical Association*, 285
- Richman, N. and Graham, P. J. (1971) A behavioural screening questionnaire for use with three-year-old children, *Journal of Child Psychology and Psychiatry*, 12, 5-33.
- Richmond, M. and Stocker, C. (2007) Associations between family cohesion and adolescent siblings' externalizing behaviour., *Journal of Family Psychology*, 20, 663-669.
- Roming, C. and Bakken, L. (1992) Intimacy development in middle adolescence: Its relationship to gender and family cohesion and adaptability, *Journal of Youth and Adolescence*, 21, 325-338.
- Rosenthal, R. (1979) The "file drawer problem" and tolerance for null results, *Psychological bulletin*, 86, 638-641.
- Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.
- Sanders, M. (2003) Triple P - Positive Parenting Program: A population approach to promoting competent parenting, *Australian e-Journal for the Advancement of Mental Health*, 2, 1-17.
- Sanders, M., Markie-Dadds, C. and Turner, K. (2003) Theoretical, scientific and clinical foundations of the TripleP - Positive Parenting Program: A population approach to the promotion of parenting competence, *Parenting Research and Practice Monograph No. 1*.
- Schweinhart, L. J. (2004). *The High/Scope Perry Preschool Study through age 40: Summary, conclusions, and frequently asked questions*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Schweinhart, L. J., Barnes, H. V. and Weikart, D. P. (1993) *Significant benefits: The High/Scope Perry Preschool Study through age 27*, High/Scope Press, Ypsilanti, MI.
- Sprigle, J. E. and Schaefer, L. (1985) Longitudinal evaluation of the effects of two compensatory preschool programs on fourth- through sixth-grade students, *Developmental Psychology*, 21, 702-708.
- Stanley, F., Richardson, S. and Prior, M. (2005). *Children of the lucky country? How Australia has turned its back on children and why children matter*. Sydney: Pan Macmillan.
- The Developmental Crime Prevention Consortium. (1999) *Pathways to prevention: Developmental and early intervention approaches to crime in Australia*, National Crime Prevention: Commonwealth Attorney-General's Department,, Brisbane.
- Tremblay, R. E., Masse, B., Perron, D., LeBlanc, M., Schwartzman, A. E. and Ledingham, J. E. (1992) Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analyses, *Journal of Consulting and Clinical Psychology*, 60, 64-72.
- Wagner, M. M. and Clayton, S. L. (1999) The Parents as Teachers program: Results from two demonstrations, *The Future of Children*, 9, 91-115.
- Webster-Stratton, C., Rinaldi, J. and Reid, M. (2009). Long-term outcomes of Incredible Years Parenting Program: Predictors of adolescent adjustment. Retrieved from <http://www.incredibleyears.com/library/paper.asp?nMode=1&nLibraryID=598>

- Weissberg, R. P. and Greenberg, M. T. (1998) School and community competence-enhancement and prevention programs, In *Handbook of child psychology*, (Eds) W. Damon, I. E. Sigel and K. A. Renninger, Wiley, New York
- Yoshikawa, H. (1994) Prevention as cumulative protection: Effects of early family support and education on chronic delinquency and its risks, *Psychological bulletin*, 115, 28-54.
- Zigler, E., Taussig, C. and Black, K. (1992) Early childhood intervention: A promising preventative for juvenile delinquency, *American Psychologist*, 47, 997-1006.

Table 1: Early developmental prevention programs included in the meta-analysis

Intervention Citation/Country of study	Study Design (randomised design – yes/no)	Program length	Sample Size Intervention (En)	Sample Size Control (Cn)	Child's age at follow-up	Outcomes
Abecedarian Project (Campbell & Ramey, 1994, 1995; Campbell, et al., 2002) United States of America	EE group (8 years of intervention – 5 yrs kindergarten + 3 yrs primary school) and EC group (5 yrs intervention – preschool only) (yes)	8 years (EE group) 5 years (EC group)	En =53	Cn =51	Age 12, 15 and 20	Academic achievement, cognitive, adult cognitive outcomes, adult reading grade equivalent, adult math grade equivalent, school success (post-secondary academic), completed school years, high school graduation, adult employment, socio-economic success (self-supporting), teen pregnancy reduction, social responsibility (misdemeanour, felony, incarceration, drug use)
Parent-Child Development Centres (PCDs) (Johnson, 2006; Johnson & Blumenthal, 2004) United States of America	Matched control (no)	5 Years	En =84	Cn =160	Age 13-16	Academic achievement/school performance, mother and family development/family functioning, child behaviour problems
Chicago Child-Parent Centre (Reynolds, 1994; Reynolds, et al., 2001) United States of America	CPC preschool vs. Comparison group (no)	2 years and 4-6 extended	En = 989	Cn =550	Age 12, 16 and 20 years	Cognitive, social emotional, school success (special education, high school graduation, school drop out, grade retention), social responsibility (juvenile arrest, multiple arrests by 18 yrs, rates of violent and non-violent arrest)
Early Training Project (Gray & Klaus, 1970; Lazar & Darlington, 1982) United States of America	Home visitation and preschool vs. Control (yes)	2 & 3 years	En = 61	Cn =27	End of Preschool age 9-10 years, 16 years	Child cognitive and language development, personal behaviour, social/emotional
Elmira Prenatal/ Early Infancy Project (Eckenrode, et al., 2000; Olds, et al., 1998) United States of America	Intervention vs. Control (yes)	3 years	En(*) =184 En(**) =38	Cn(*) =116 Cn(**) =62	Age 15 years	Social/emotional Criminal and antisocial behaviour
Learning to Learn (Sprigle & Schaefer, 1985) United States of America	Learning to Learn vs. Head Start (no)	3 years	En = 44	Cn =39	Age 12	Cognitive, social-emotional outcomes
Louisville Experiment (Miller & Bizzell, 1983) United States of America	Preschool interventions vs. Control (yes)	1 Year	En = 114	Cn =36	Age 13	Cognitive outcomes
Mother-Child Home Program (Levenstein, Levenstein, Shiminski, & Stolzberg, 1998)	Home-based intervention with mothers vs. Control (yes)	>1 year-2 years	Full 2 year En = 70	Less than 2 year Cn =38 No	Age 13, 17 and 22 years	Social-emotional outcomes, cognitive, high school graduation, school drop out

Intervention Citation/Country of study	Study Design (randomised design – yes/no)	Program length	Sample Size Intervention (En)	Sample Size Control (Cn)	Child's age at follow-up	Outcomes
United States of America				program <i>n</i> = 15		
Perry Preschool Program (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984) United States of America	Perry Preschool vs. Control (yes)	2 years	Preschool <i>En</i> = 68 Age 13 <i>En</i> = 68 Age 18 <i>En</i> = 55	Preschool <i>Cn</i> = 65 Age 13 <i>Cn</i> = 65 Age 18 <i>Cn</i> = 62	Preschool, age 13, 18	Social-emotional, personal behaviour, social development, academic skills, personal behaviour, school success, cognitive outcomes, effects on deviance and social patterns (e.g. delinquent behaviour, threatened or injured another person, employment, self-confidence), special education, high school graduation, school drop-out, post secondary academic and vocational training, social responsibility (e.g. juvenile arrest, multiple arrests by 18 yrs, adult arrests), effects on socioeconomic success (e.g. employment/unemployment, annual income, self-supporting), effects on health, family, and children in mid-life.
The Syracuse Family Research Development Program (FDRP) (Lally, Mangione, & Honig, 1988) United States of America	Multi-component vs. comparison (no)	5 years	Preschool <i>En</i> = 82	Preschool <i>Cn</i> = 74	Preschool, 13-15 years of age	Cognitive, social-emotional, social responsibility (e.g. juvenile arrests, violent arrest)
Direct Instruction project (Meyer, 1984) United States of America	DISTAR follow through vs. comparison (no)	3-4 years	<i>En</i> = 65	<i>Cn</i> = 100	18-20 years of age	Educational success (e.g. school graduation, retention, school drop-out, accepted for college). Cognitive development (e.g. ninth grade reading and math scores)

Table 2: Outcome domains and operationalizations

Outcome domains	Operationalizations
Educational success (ES)	Special education; feeling of belonging at school; graduation; school drop-out; long-term school suspension; grade retention; completed years of education; school attendance (e.g. >20 absent days from school per year); learning problems.
Cognitive development (CD)	IQ; achievement tests; school grades; rating of academic skill and performance; school failure.
Social-emotional development (SED)	Parent/teacher rating of problem behavior; social skills; self-esteem; self-confidence; obsessive-compulsive behavior.
Deviance (D)	Rates of delinquent behavior; drug use (e.g. marijuana and alcohol); lying about age (e.g. false Id); running away from home; caught breaking the law; gang involvement.
Social participation (SP)	Casual employment in teen years; socio-economic success; engaged in skilled jobs (e.g. electrician); makes active response to problems.
Criminal justice (CJ)	Rates of juvenile arrest; rates of violent and non-violent arrest; incarceration; petition requests to juvenile court; adjudicated as a person in need of supervision due to incorrigible behavior (PINS).
Family wellbeing (FW)	Child maltreatment; parent-adolescent relationship; family functioning; parental mental health; parental employment/education; parental social support; quality of parenting; adolescent influence in family decisions; single-parent families; parental involvement in schooling; discrepancy between mother's occupational aspirations for child and child's own aspirations; child abuse and neglect; feeling of family unity.

Table 3: Program characteristics

Program Characteristics (n=11)	Number	% of total programs
Number of programs with a home visiting component (HV) (visiting parents in their home, including pediatric follow-up services)	6	54.5
Number of programs with a parent training/child management and/or educational strategies component (PE) (e.g. groups exclusively for parents such as parent training or individual parent training)	1	9.1
Number of programs with a preschool program component (SPP) (specifically, teachers and staff adopting educational strategies)	7	63.6
Number of programs with family/parenting support and education, guidance, case management, and referrals to other agencies component (FSS) (e.g. health and other human services)	5	45.5
Number of programs with center-based childcare/developmental day care component (CBCC)	4	36.4
1 -2 program components	8	72.7
3 or more program components	3	27.3
≤ 1 year duration of intervention	1	9.1
>1<3 years duration of intervention	6	5.5
>3 years duration of intervention	4	36.4
Intensity of intervention (Child)*		
≤ 300 sessions	3	27.3
>300 sessions ≤500	3	27.3
>500 sessions	5	45.5
Follow-through intervention	4	36.4

Table 4: Study Characteristics

Characteristics of program participants (n=11)	Number	% of total programs
• Predominately African-American ethno-racial background (≥ 50% participants African-American)	8	72.3
• Other ethno-racial backgrounds (<50% participants African-American e.g. Mexican-American)	3	27.3
Study Characteristics (n=17)	Number	% of total programs
• Randomized design	11	64.7
• Non-randomized (e.g. matched-pairs)	6	35.3
Target population clearly described	17	100
• Sample size		
• <300	14	82.4
• >300≤500	1	5.9
• >500	2	11.8
Year published		
• Before 1980	1	5.9
• Before 1990	6	35.3
• Before 2000	5	29.4
• After 2000	5	29.4

Table 5: Test of homogeneity for seven outcome domains

	k	d.	Lower C.I (95%)	Upper C.I	t-Value	P-Value	Q-Value	df (Q)
Educational success	10	.528	.4001	.6794	7.5850	.0000	9.3795	9
Cognitive development	14	.339	.2475	.4400	7.0046	.0000	7.9839	13
Social-emotional development	6	.157	.0459	.2580	2.8102	.0050	6.2648	5
Deviance	3	.481	.2664	.7070	4.3459	.0000	1.7636	2
Social participation	3	.371	.1806	.5672	3.8025	.0002	.3670	2
Criminal justice	5	.243	.1461	.3424	4.8819	.0000	3.2027	4
Family wellbeing	5	.178	.0641	.2782	3.1354	.0018	6.9570	4

Figure 1: Weighted average effect sizes (d) corrected for sample size for seven adolescent outcomes

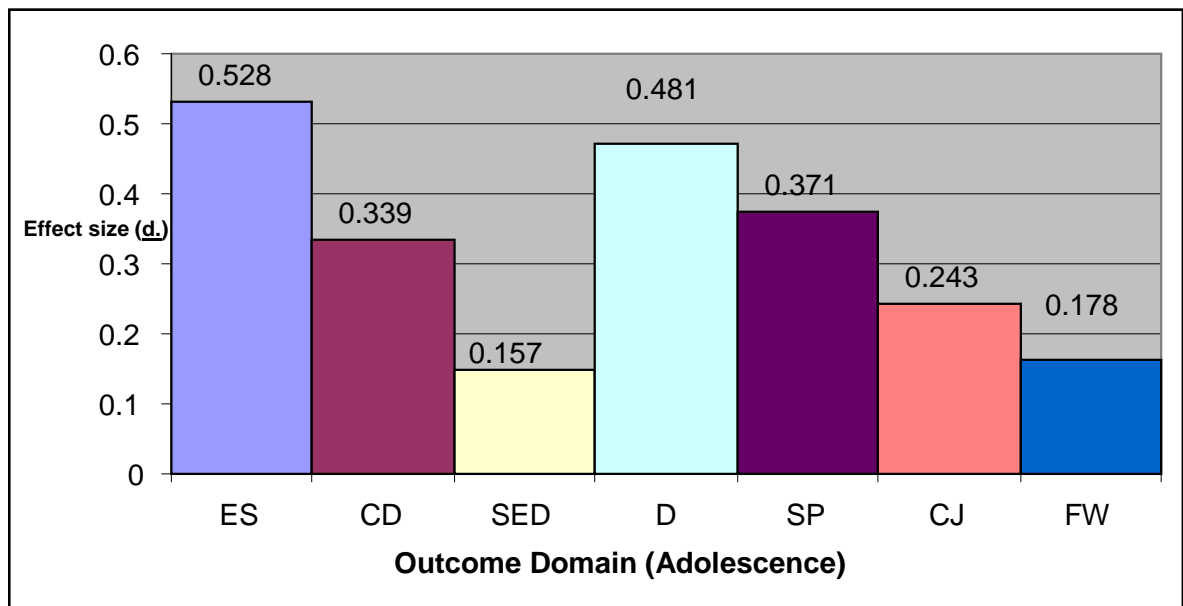


Figure 2: Educational Success

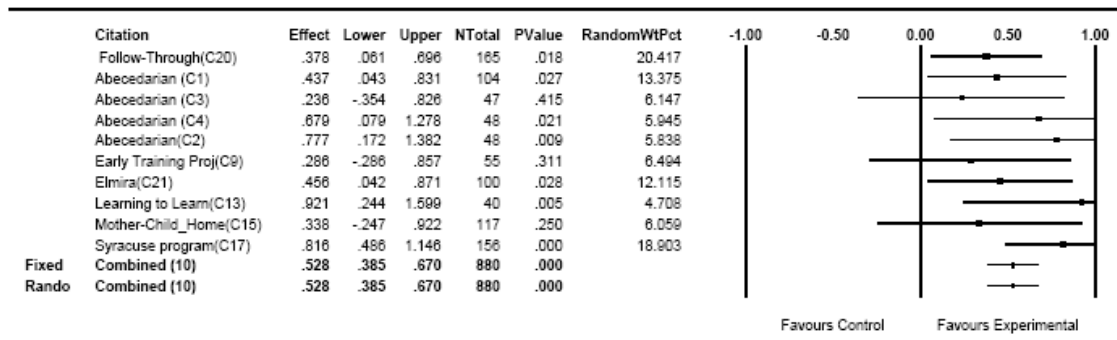


Figure 3: Cognitive Development

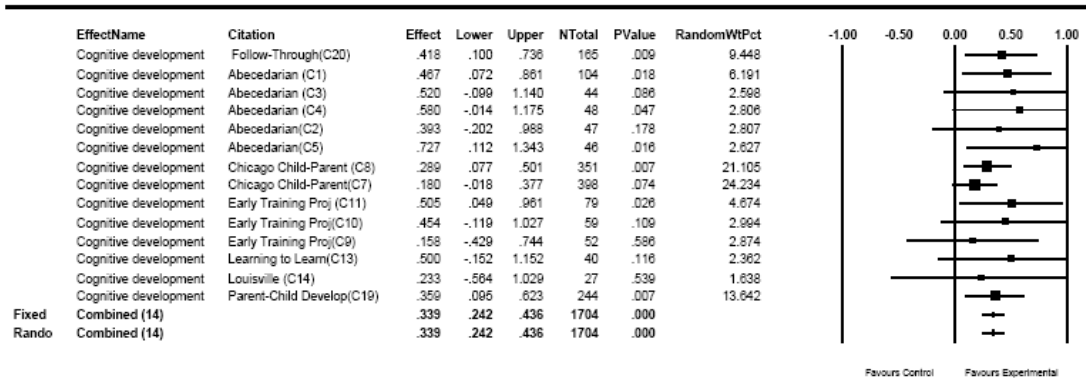


Figure 4: Social-Emotional Development

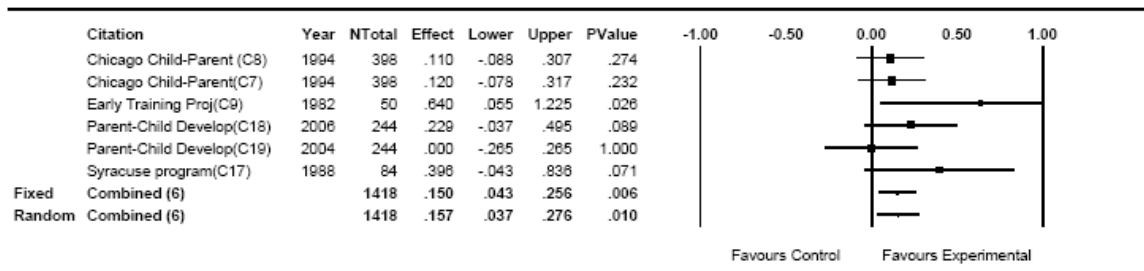


Figure 5: Deviance

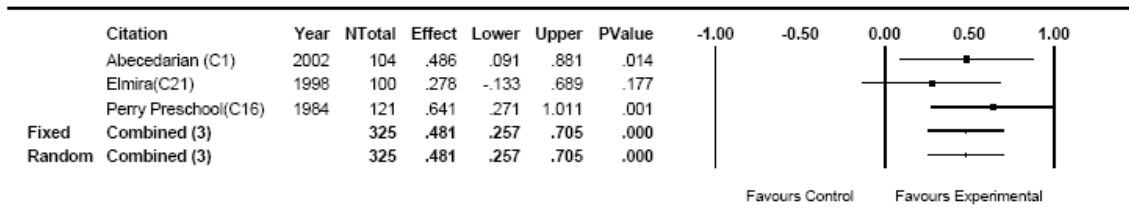


Figure 6: Social Participation

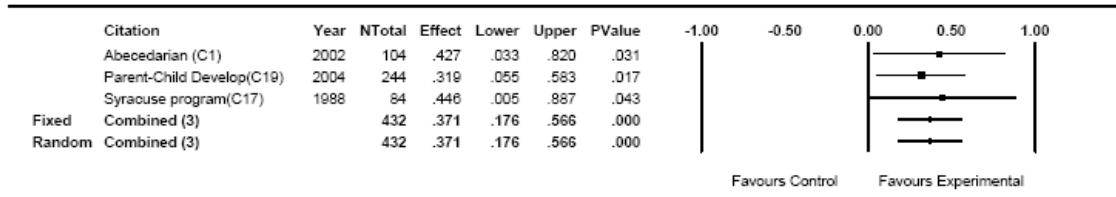


Figure 7: Criminal Justice

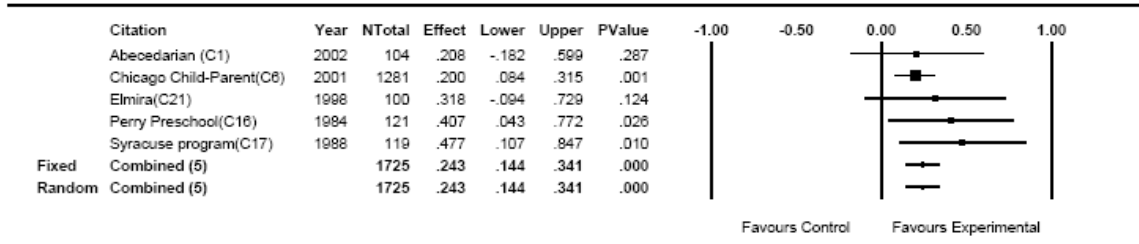


Figure 8: Family Wellbeing

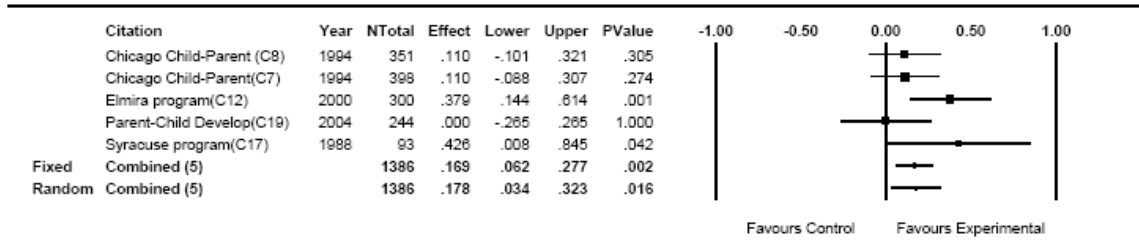
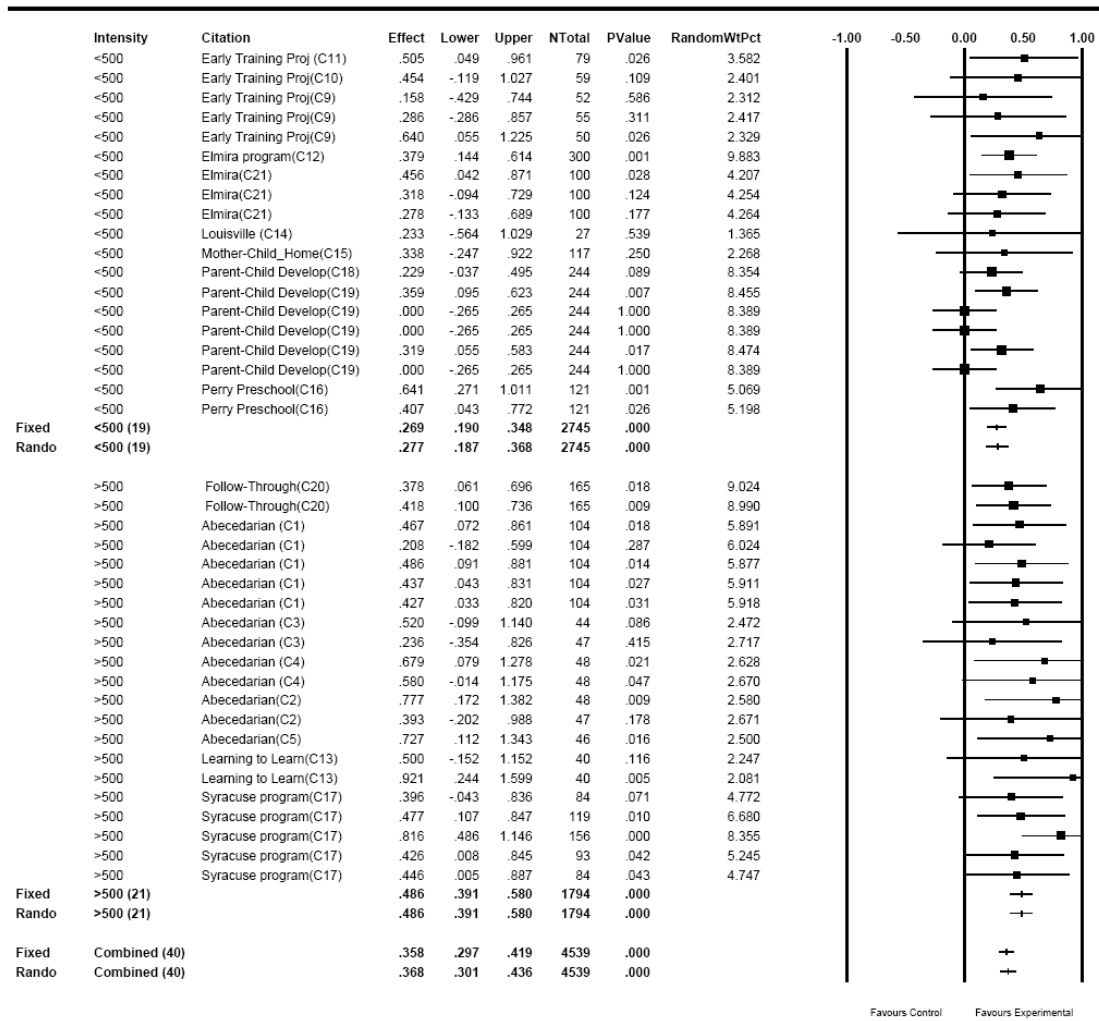
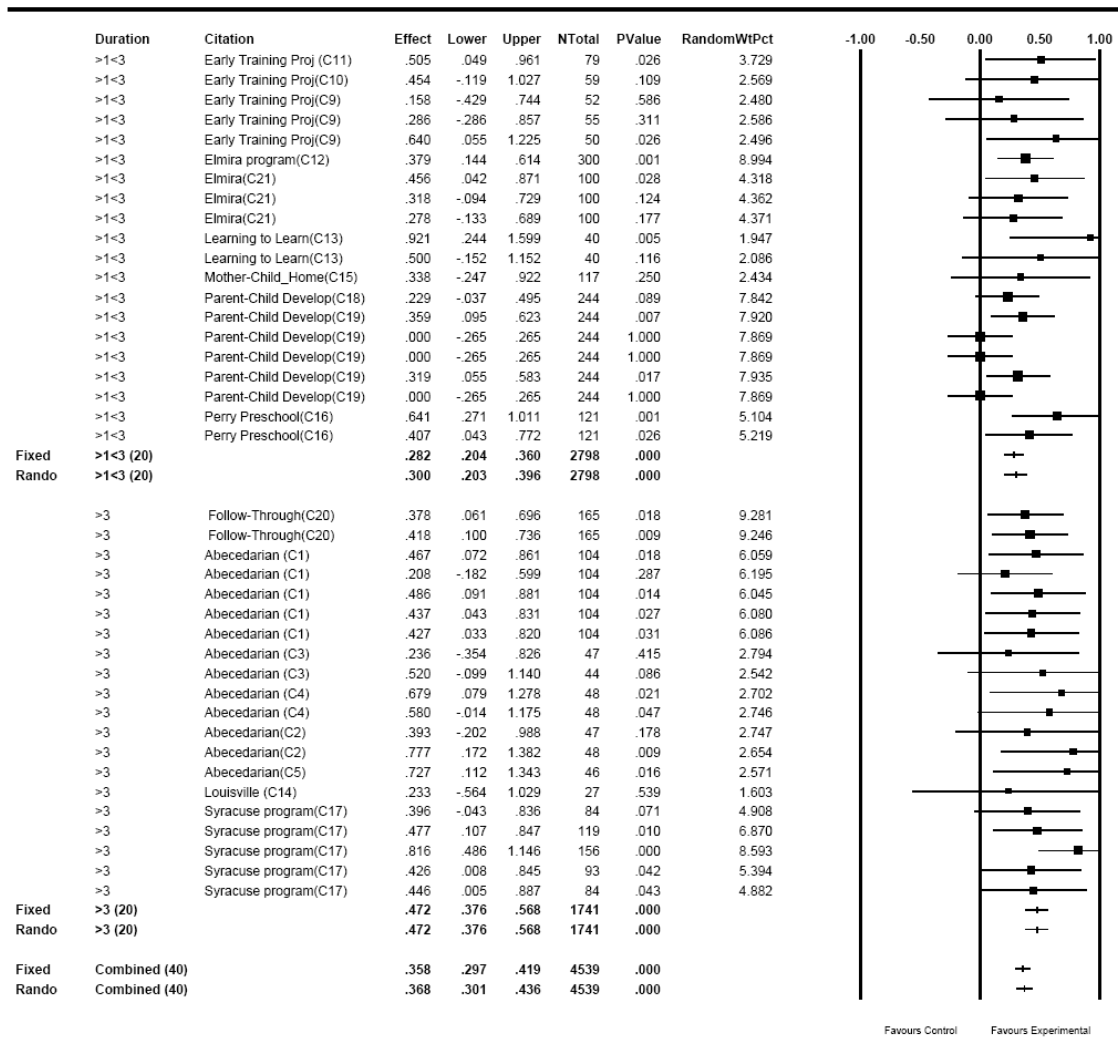


Figure 9: Moderator analysis: Program Intensity



Note: Effect sizes are represented in the diagrams by Hedges g. Corresponding d for <500 = 0.283; >500=0.494 based on a random effects model.

Figure 10: Moderator analysis: Program Duration



Note: Effect sizes are represented in the diagrams by Hedges *g*. Corresponding *d* for >1<3 = 0.308; >3=0.479 based on a random effects model.